INVESTIGATION OF THE METABOLIC CHANGES IN THE ADAPTIVE EVOLUTION OF S. CEREVISIAE STRAINS BY USING GENOME-SCALE METABOLIC MODELS

by

Handan Çetin

B.S., Molecular Biology and Genetics,
Canakkale Onsekiz Mart University, 2018

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Computational Science and Engineering
Boğaziçi University
2021

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my thesis supervisor Prof. Ş. Kutlu Ülgen and thesis co-supervisor Prof. Zeynep Petek Çakar for their invaluable help, support, and patience that cannot be underestimated throughout my studies. Their guidance, valuable support and motivation throughout this study were invaluable.

I would like to thank my thesis committee members Assoc. Prof. Arzucan Özgur, Dr. Betül Uralcan and Dr. Pınar Pir, for devoting their valuable time and effort to my thesis and providing guidance and recommendations.

I would also like to thank everyone I have been working together in the Ulgen Research Group for their collaborations and helpful advice in this work environment.

Finally, I would like to extend my sincere thanks to all my family members and friends for their substantial support through tough times of COVID-19.

I gratefully acknowledge the financial support of Boğaziçi University Research Fund (no. 16743P) throughout my master's education.

Thank you.

ABSTRACT

INVESTIGATION OF THE METABOLIC CHANGES IN THE ADAPTIVE EVOLUTION OF S. CEREVISIAE STRAINS BY USING GENOME-SCALE METABOLIC MODELS

Increasing an organisms' tolerance level to various stress conditions is one of the most common applications of adaptive laboratory evolution (ALE) experiments which provides high amounts of genomic and experimental data that usually requires interdisciplinary research to analyze.

This study integrates enzyme kinetics and expression profiles of multiple *S. cerevisiae* strains that have evolved in 8 stress conditions, ethanol, caffeine, coniferylaldehyde, iron, nickel, phenylethanol, and silver, into the genome-scale metabolic model of yeast to simulate evolved strains' behaviors. Traditional analyses such as FBA, FVA, MOMA, PhPP, robustness, survivability, sensitivity, and random sampling are conducted to identify the most common and divergent points in cellular metabolism. Reconstructed models were able to simulate batch conditions where efficient protein allocation is the main goal for cells to survive under stressful environments. Simulation results were investigated both individually and comparatively, on both small-scale and large scale. The most common and varying targets of reactions, metabolites, and enzymes that are assumed to be responsible for the adaptation mechanics were identified and discussed across mutant strains.

The study presented here successfully used metabolic modeling techniques on ALE data while providing a systematic view of the general yeast metabolism.

ÖZET

S. CEREVISIAE SUŞLARININ ADAPTİF EVRİMİNDEKİ METABOLİK DEĞİŞİKLİKLERİN GENOM ÖLÇEKLİ METABOLİK MODELLERLE İNCELENMESİ

Bir organizmanın çeşitli stres koşullarına tolerans düzeyini artırmak, genellikle analiz etmek için disiplinler arası araştırma gerektiren yüksek miktarda genomik ve deneysel veri sağlayan adaptif laboratuvar evrimi (ALE) deneylerinin en yaygın uygulamalarından biridir.

Bu çalışmada, etanol, kafein, koniferilaldehit, demir, nikel, feniletanol ve gümüş stresi koşullarında evrimleşen *S. cerevisiae* suşlarının ekspresyon profilleri ve enzim kinetiği, evrimleşmiş suşların davranışlarını simüle etmek için genom ölçekli metabolik model ile bütünleştirilmiştir. Akı denge ve değişkenlik analizi, dayanıklılık analizi, duyarlılık, hassaslık ve rastgele örneklendirme gibi geleneksel metotlar, hücresel metabolizmadaki en ortak ve farklılaşmış noktaları belirlemek için uygulanmıştır. Yeniden yapılandırılmış modeller, stresli ortamlarda hayatta kalabilmek için ana hedefini verimli protein pay ayırma olarak düzenlemiş hücrelerin deney koşullarını simüle edebilmiştir. Simülasyon sonuçları, hem küçük hem de büyük ölçekli olarak, hem bireysel hem de karşılaştırmalı olarak incelenmiş, adaptasyon mekaniğinden sorumlu olduğu öngörülen reaksiyonlar, metabolitler ve/veya enzimlerden en ortak ve değişken davranış gösterenleri mutant suşlar arasında tanımlanmış vetartışılmıştır.

Yapılan çalışmalar, genel maya metabolizması hakkında sistematik bir bakış açısı sağlarken, genom ölçeğinde metabolik modelleme tekniklerini ALE deneylerinin verileri üzerinde başarıyla kullanmıştır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

h	Hour
S	Stoichiometric Coefficient Matrix
v	Flux vector
α	Alpha
β	Beta
Δ	Delta

LIST OF ACRONYMS/ABBREVIATIONS

ALE	Adaptive Laboratory Evolution
DIY	Do-It-Yourself
EMS	Ethyl Methane Sulfonate
FBA	Flux Balance Analysis
FVA	Flux Variability Analysis
GPR	Gene-Protein-Reaction
GSMM	Genome-Scale Metabolic Model
GUR	Glucose Uptake Rate
HRB	Hit-and-Run on a Box
LP	Linear Programming
MR	Metabolic Regulation
MOMA	Minimization of Metabolic Adjustment
PhPP	Phenotype Phase Plane
PR	Post-Transcriptional Regulation
QP	Quadratic Programming
TR	Transcriptional Regulation
YMM	Yeast Minimal Media
YPD	Yeast Extract-Peptone-Dextrose

1. INTRODUCTION

Evolution has been intensively studied since the publication of *On the Origin of Species* in order to untie the mysteries behind fascinating machineries within living systems. In a broader sense, most of the studies on evolutionary biology have two main goals: (1) To document history of life through evolutionary point of view, and (2) to understand causal mechanisms responsible for the biological diversity that we have on Earth [1, 2]. In addition to these goals, the idea of using pre-existing living systems as cell factories for industrial purposes gained close attention in the last decades [3].

Evolutionary engineering studies, adaptive laboratory evolution in particular, refers to the experiments in which the environmental conditions are altered gradually to obtain adapted populations in the laboratories [4]. Accumulation of mutations obtained due to environmental alterations through generations, where the favored individuals (mainly the ones with increased fitness) are selected to become parents of the next generation, result a population with advantageous traits compared to starting population. Since one of the main challenges in the evolutionary engineering field is being able to analyze experimental to answer fundamental questions, adaptive evolution studies usually require interdisciplinary research in order to answer fundamental questions. What is the genetic basis for adaptation? Can we evolve any organism to any condition? If so, how we can achieve? These questions and many more are asked every day, and the corresponding answers potentially give rise to more questions.

The purpose of this thesis was to enlighten metabolic changes in the adaptation of the yeast *S. cerevisiae* using computational methods, specifically genome-scale metabolic models. Evolved yeast strains (such as ethanol tolerant, multi-stress resistant strains) were analyzed comparatively to the unevolved strains using intracellular flux distributions obtained with the integration of transcriptomics data. This study will also contribute to the global understanding of metabolic regulations in yeast, and will be further expandable into metabolic engineering studies.

2. THEORETICAL BACKGROUND

2.1. Adaptive Laboratory Evolution

Since the very first laboratory evolution experiment was published in the late 19th century by William Dallinger, technological advancements allowed researchers to employ fully-controlled experiments for strain engineering to achieve desired traits [5]. Challenges in the experimental design of evolutionary studies, such as maintenance of the generations, controlling the environment, and feasibility to perform data analysis make the use of microorganisms in evolutionary studies more suited, especially *Escherichia coli* and *Saccharomyces cerevisiae* given their extensive characterization [6].

Commonly used approaches of adaptive laboratory evolution (ALE) include chemostat cultures and serial batch or colony transfers (Figure 2.1). Serial transfer experiments in which the initial population is aliquoted and transferred into a new medium are simpler and cheaper to set up, however the possibility of genetic drift is higher due to random sampling of the population. On the other hand, continuous systems such as chemostat experiments in bioreactors have advantages on maintaining constant growth rates and population sizes, but with increasing experimental cost [7].

In addition to the methodological choice, decision of the selection criteria and the time span for the experiments are critical factors in ALE [5]. Growth rates, survivability in stressful environmental conditions and biomass yields are the most common fitness criteria for a population to be selected. Throughout the selection process, reaching >500 generations may take a few weeks or a few months depending on the organism, therefore, detailed planning is essential [8].

Capturing genomic changes in the dynamic evolution process is the main goal of an ALE study, therefore frequent data collection in longitudinal manner is a must to unveil molecular details. Next-generation technologies empower researchers to catch even single-nucleotide mutations through genome sequencing and also provide an understanding on broader regulatory changes through equation expression profiles [8]. Focus of the studying the differences between evolved and unevolved strains may be on the individual protein level or as a system-level trait. Interpreting the system as a network enables various ways to investigate causality behind adaptive mechanisms and dynamics of evolution [9, 10].

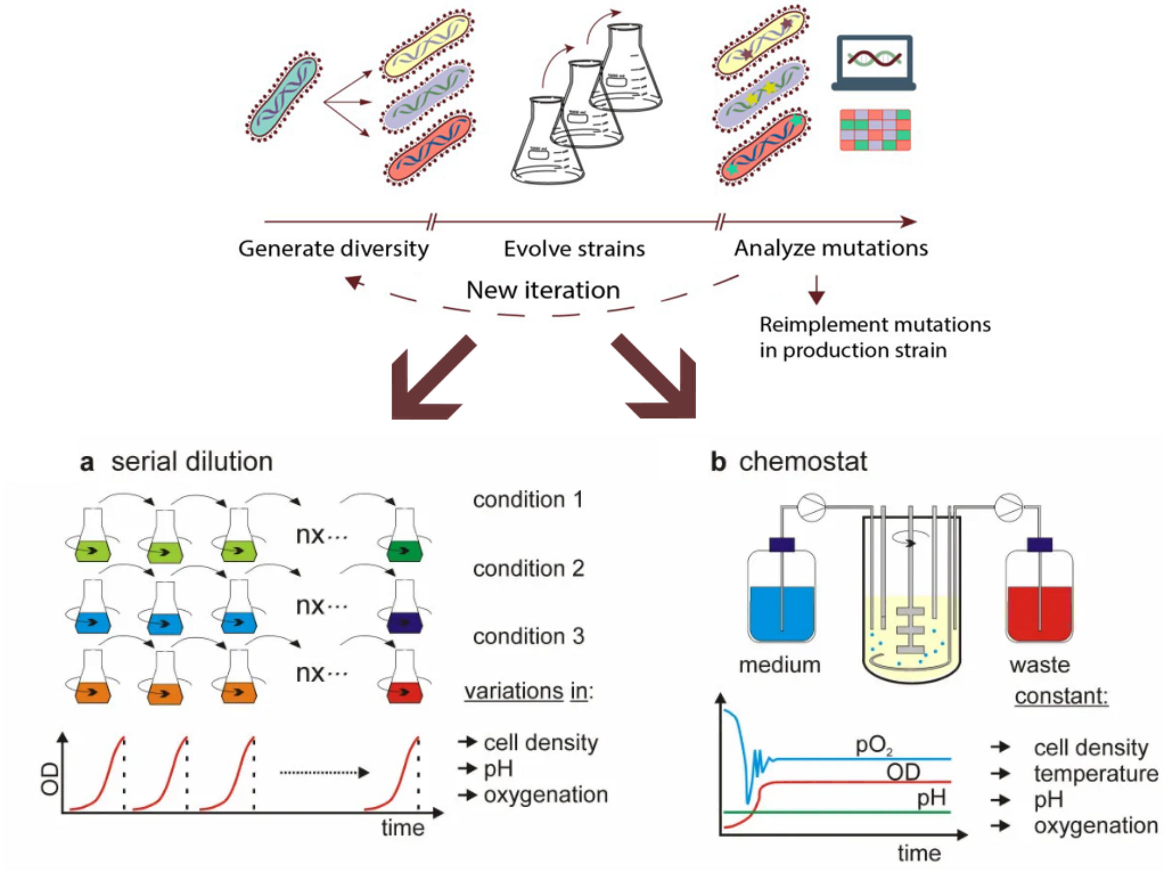


Figure 2.1. Workflow of ALE experiments and the two iteration mo (a) sequential serial passages in shake flasks, (b) chemostat cultures [5, 11].

2.1.1. Computational Tools for ALE Experiments

Adaptive laboratory evolution, in laboratories, uses natural selection to capture microbial adaptations for decades. Although the results of these experiments would provide the ideal selection pressure to be imposed on organisms to obtain favored strains, the difference between the starting and the evolved strain is not easy to find. Throughout the years, various additional methods and computational tools are invested to investigate findings. With decreasing cost of whole genome sequencing and resequencing, identification of mutations can reveal the mystery behind the adaptations. Due to complexity of cellular systems and missing points in the metabolic regulations, the interpretation of the mutation effects is still the main challenge in the field [12].

A web-based database platform, ALEdb (aledb.org) [13], has been created last year. Platform collects and reports on ALE acquired mutations and their experimental conditions and provides several features such as searching for specific mutations within reports, exporting of mutation query results for custom analysis, and many more. With these useful features and being freely available for non-commercial uses, ALEdb becomes a resourceful tool in the field (Figure 2.2).

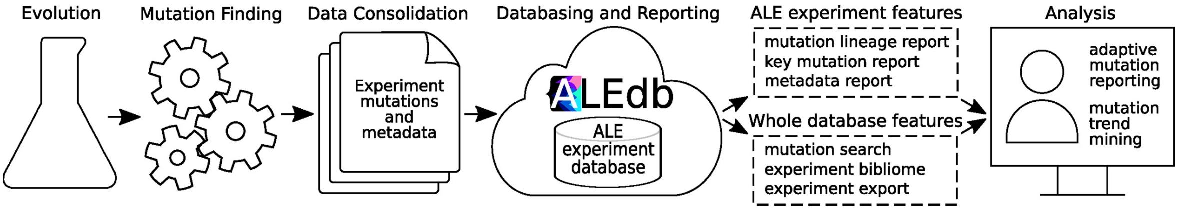


Figure 2.2. Pipeline flow of ALE data from beginning (experimental design) to results reporting in ALEdb [13].

eVOLVER [14], an open-source do-it-yourself (DIY) framework that promises a solution to the main challenge of balancing the tradeoff between controllable (such as bioreactor designs) and throughput (such as batch designs) experiments in the ALE. Framework comes with both software and hardware, and it offers to define the various experimental parameters e.g. temperature, density, media, in a virtually automated fashion. It simultaneously controls multiple experimental designs, collects, and measures data in real-time. eVOLVER is also designed to be scalable to high-throughput, therefore, it is a promising tool to carry out automated multidimensional growth and selection experiments.

In order to demonstrate the potential of eVOLVER framework, they designed a device to conduct automated sexual reproduction for adapting yeast populations under different selections. In the experiment, instead of manual sampling at arbitrary time points, the sampling process and mating carried out automatically within the device at the times when the cultures reach the defined points such as growth-rates. Media selection, vial-to-vial transfers, and deceive cleaning routines were all included in the framework. This experiment illustrated the potential of eVOLVER framework in terms of automatizing the custom build and controllable microfluidics for continuous cultures.

Another simulator, ALEsim [15], was developed for the optimization of ALE experiments and promises to improve ALE results. It is a model built on the basic principles of growth in order to predict growth rates in batch cultures of ALE experiments. By evaluating various passage sizes, ALEsim can calculate how and when to deploy resources to shorten experiment timelines and to achieve desired results. ALEsim provides a way to design better experiments and quantify them, to obtain the desired outcome with the resources available on hand. For a simulation using a developed model, experimental (i.e., culture size, passage size, passage optical density), statistical (i.e., number of experiments), and biological (i.e., beneficial mutation rate) parameters must be set for results to be biologically meaningful. These parameters can differ for different strains and conditions, and it is crucial to define the parameters for models to be constrained. ALEsim provides an advantage by allowing automated simulations, with the cells being allowed to grow under the user-defined conditions so that the experimental and biological parameters can be strictly controlled over the course of an experiment.

2.2. Systems Biology

With the increasing availability of the computational tools and the development of high throughput techniques in the omics field, systems biology has shown a strong emergence in the last few years as a key multi-disciplinary field for integrating the multilayer complexity of biological systems, particularly in the areas of transcriptomics, proteomics, metabolomics and fluxomics [16]. This amount of available data allows researchers to investigate molecular cell processes in a large scales, applying theoretical, experimental and computational methods.

Biological systems based on complex interactions between various molecular components. The relations between these components are often obey nonlinear kinetics, for example, most of the reactions are regulated by one or more feedback or feed-forward loops with incomprehensible behaviors. When considered, cell structure and compartmentalization are also often introduce complexities to the unexpected behavior of the entire biological system [17]. Mathematical modeling with these factors taken into consideration is used as a general approach to encompass existing knowledge in biological systems, and to gather information by analyzing these models to acquire a better understanding [18].

A mathematical model of a cell can be approached by two different in either a bottom-up or top-down directionality (Figure 2.3) [19, 20]. Top-down approach is an experimental oriented approach, it starts from the whole picture and aims to characterize biological mechanisms closer to the smaller parts and their interactions in the network. In the bottom-up approach, collected data from biological knowledge is used as a starting point, a subsystem is generated to deduce the functional properties of smaller points in the network. Combination of the pathway level models (bottom-up) into a model for the entire system level (top-down) is the ultimate goal in the systems biology therefore these approaches are complementary.

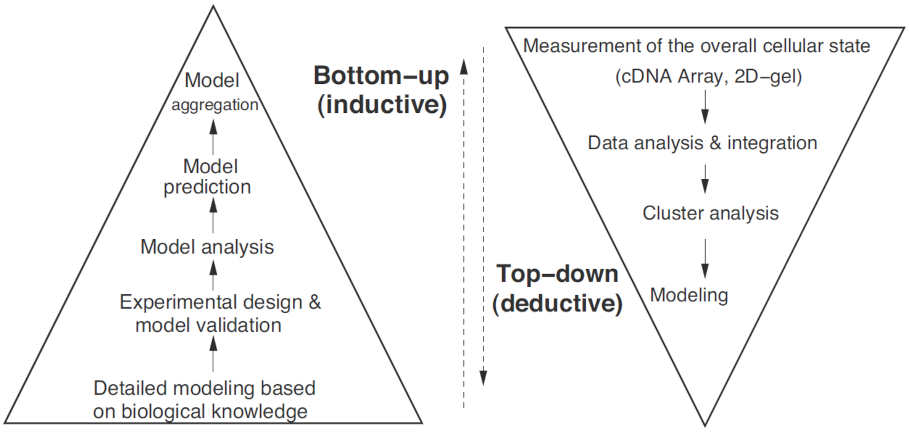


Figure 2.3. Systems biology approaches. Bottom-up approach on the left, top-down approach on the right triangle [18].

2.2.1. Metabolic Networks

In the context of systems biology, metabolic network reconstructions have become a common interest for the researchers over the past 20 years [21]. Organism-specific metabolic network analyses allow scientists to design experiments and even obtain beforehand predictions. These networks are the main sources of the mathematical models which can simulate metabolic fluxes reflecting the experimental reality [22].

Before the improvement of genome sequencing or annotation technologies, initial core metabolic networks were based on the accessible information of biochemical pathways [23] [24]. In the last decade, larger genome-scale metabolic models (GSMMs) have been able to be developed rapidly with the help of databases for annotated genomes, providing information on substrates and products of each enzyme and each bioreaction [25]. Growing biochemical databases provide automatization processes for the metabolic network reconstructions. As a result, genome-scale metabolic networks are available today for almost all organisms with an annotated genome available in the literature [26, 27]. From the first genome-scale metabolic model of *Escherichia coli* to other organisms, the steps are required for GSMM development remained the same regardless of the biological diversity.

A generally applicable protocol is defined by the Palsson group [21,25] for the reconstruction of biochemical networks described in the Figure 2.4 [28]. Briefly, genomic data for the biochemical reactions of an organism are identified from the databases. Extraction and processing of the gene-protein-reaction relationship (GPR) of the genomic data results a draft reconstruction. GPR associations in the draft model should be reviewed by the researchers and manually curated if the identifying process is achieved with the help of automated computational algorithms [26]. Since the genomic data is the least representative of the biological phenotypes, available transcriptomic, proteomic, metabolomic and/or subcellular localization data are also used to further curate the model. Once the final metabolic network is reconstructed with bibliographic information, it is translated into a mathematical model.

Once a metabolic network is reconstructed, a rational link between a genome sequence, the proteins encoded in the genome, and the reactions catalyzed by the proteins allowing to investigate the relationships between genotype and phenotype is achieved [29]. As the final step, GSMM needs to be validated by the new experimental data sets. GSMM validation process for various experimental conditions require detailed cultivation data from experiments. For example, information on the biomass composition of the specific organism leads more accurate biomass equation in the model, that is one of the key factors in the GSMM optimization and validation [30]. Even tough multiple steps in the GSMM reconstruction can be achieved with the automated softwares available, it is usually necessary to curate the obtained model manually.

Approaches for analyzing metabolic networks are mainly categorized as dynamic or structural approaches. Even though the former is promising more realistic approach, its implementation in the literature is obstructed due to the unavailability of kinetic parameters for the majority of enzymes within a metabolic network [31,32]. Because of the lack of kinetic parameters, structural metabolic modeling has been widely used for analyzing cellular metabolism at a steady-state assumption as a kind of snapshots taken at specific times.

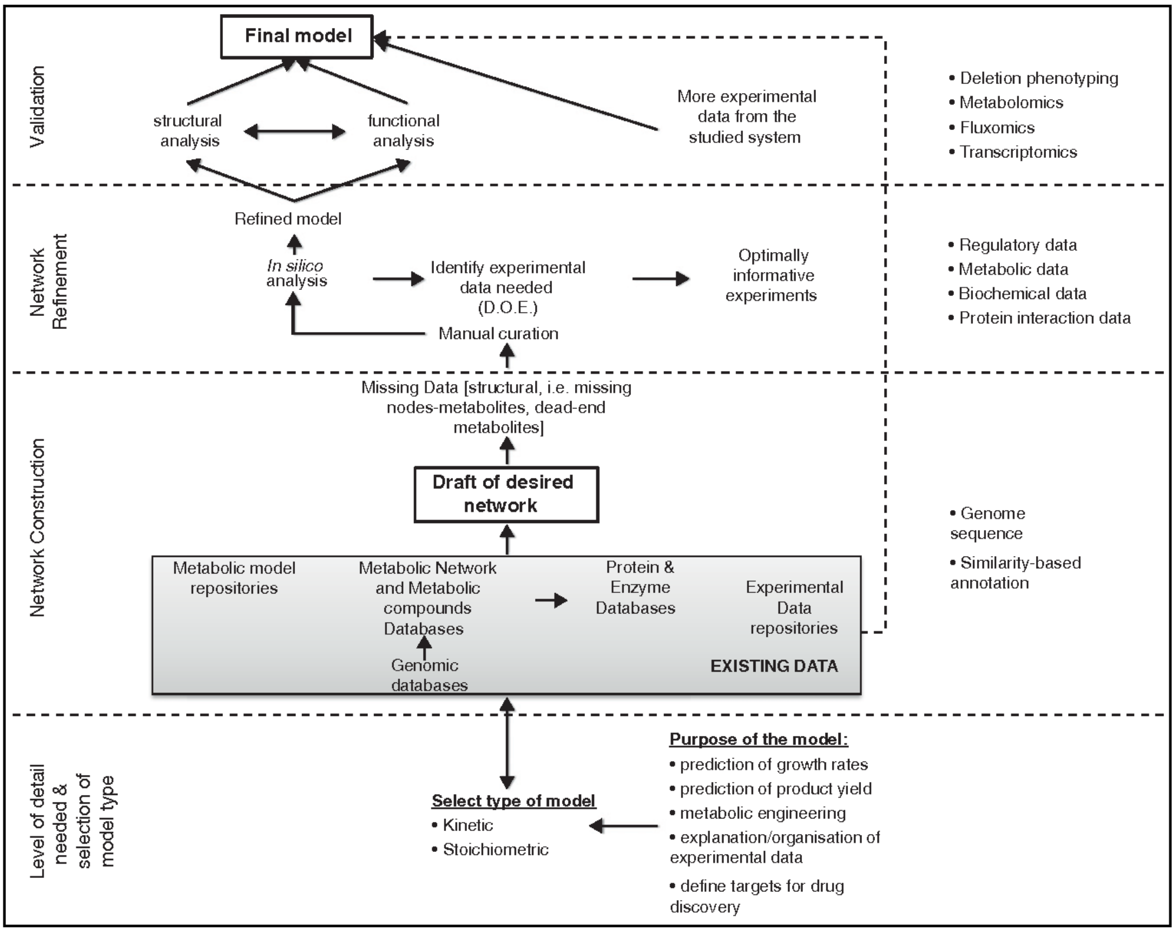


Figure 2.4. Overview of metabolic network reconstruction protocol [28].

GSMMs are one of the most useful tools in systems biology, especially in metabolic engineering studies [33]. In 1998, with the publication of *Metabolic Engineering: Principles and Methodologies*, the term metabolic engineering is defined as the optimization of natural processes within cells to increase the production of certain substances [34]. Hence, studies of metabolic engineering can be considered as genetic engineering in strain development. However, while metabolic engineering manipulates strains by altering flux distributions in the pathways; genetic engineering modifies specific genes, proteins and/or enzymes of interest [35].

Although GSMMs are mainly used in metabolic engineering strategies, other applications both for descriptive and predictive purposes can be found in the literature [36]. The ultimate goal of the GSMM reconstruction is to predict flux distribution profiles as close *in silico* as they are *in vivo*. Hence, GSMMs are in continuous research to improve predictability of organism-specific models.

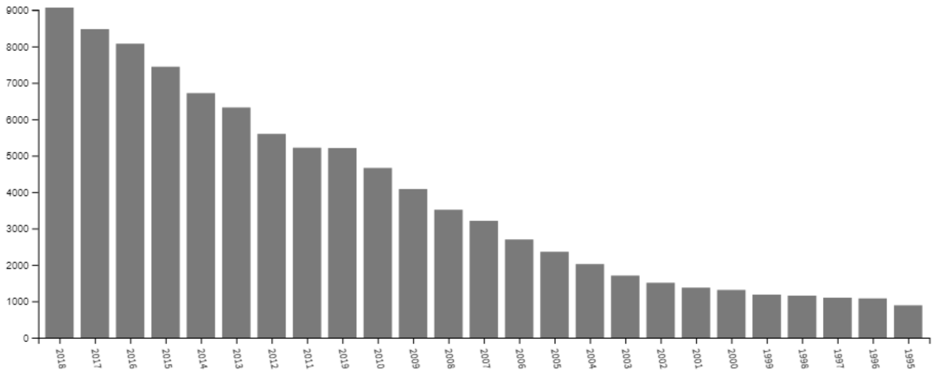


Figure 2.5. Web of Science article counts on the query "metabolic model" keyword.

2.2.2. Mathematical Representation of Metabolic Networks

Basic applications of metabolic network modelling, although its framework is developed by engineers or mathematicians, is used by biologists with various mathematical backgrounds [37]. In order to speak with one voice, fundamental concepts of mathematical representation of metabolic networks in GSMM reconstructions will be provided in this section.

One can propose a steady-state model for the correlation between the metabolites exist in the network. This model claims that the production and consumption of a metabolites must balance each other [38]. For example, consider the metabolite "A" in the toy network in Figure 2.6. It can be taken into cell with the rate of b1, and can be converted into B or C with the rates of v1 and v2 respectively; at the same time, metabolite C can be converted into A with the rate of v3. Note that enzyme kinetics are not considered, and these equations are formed only considering mass balances.

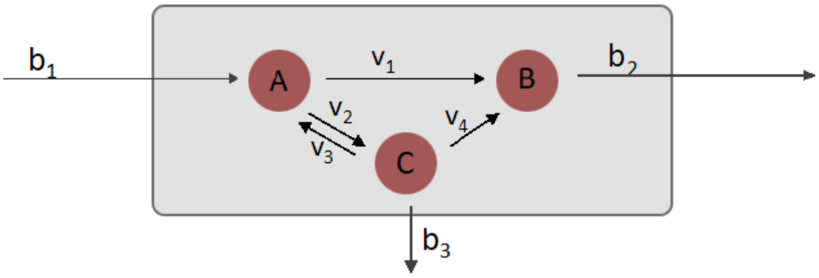


Figure 2.6. Toy network for the steady-state assumption.

If we accept the steady-state assumption, and since we know all the possible reactions passing through metabolite A, we can write following differential equation

$$\frac{d(A)}{dt} = -v_1 - v_2 + v_3 + b_1 = 0 \tag{2.1}$$

meaning that production rate of metabolite A is equal to its consumption rate. In other words, there will be no accumulation of metabolite A in the cell. Considering we have 3 metabolites and 7 reactions; this representation of reactions can be written as a system of linear equations. When written, A matrix ”S” which consists of stoichiometric coefficients of the reactions appears. A single column of S matrix, representing a single reaction, provides information about the connections between the metabolites participating in that reaction; and a single row of S matrix, representing a metabolite, provides information about the connection of all the reactions in which that metabolite participates. Next to S matrix, a vector ”v” is written. Vector ”v” is called the flux vector, and it contains the change in concentration of metabolites. In other words, flux vector represents the flow rates or ”fluxes” of each metabolite over time.

The solution space of S matrix, in the steady-state assumption, is generated by the null space base vectors of S. Studying on the null space vectors, we can define dead-end reactions (zero rows which mean reactions cannot carry flux), enzyme subsets (rows of scalar multiplies of each other, possibly chain reactions) and independent components (diagonal block structures in the null space, reactions that are independent from the network). Due to the large number of reactions in biological networks, the system is underdetermined and have multiple solutions in a convex flux cone space, referring to multiple steady-states of the cell.

2.2.3. Constraint-Based Modelling

As explained in the previous section, solution of a metabolic network is not unique, meaning that the model can be found in multiple states of flux distributions. In order to obtain more biologically relevant solutions, v vector can be calculated within a set of constraints [39]. The rationale behind constraining the flux vector is that biological systems must obey general principles in nature, such as basic rules of chemistry and the laws of thermodynamics.

In terms of physico-chemical properties of a biological system, mass and charge balance must be conserved. These conservation laws are given into mathematical problem as "hard" constraints since these rules are inviolable [40]. Reaction reversibility is a hard constraint, determined by the laws of thermodynamics. Some reactions are irreversible by nature under certain conditions. These reactions are mathematically bounded in the flux vector, so they can have only positive ($v_i \geq 0$, forward directionality) or negative ($v_i \leq 0$, backwards directionality) values. Depending on the methodology, by reversing the equations of backwards-directed reactions, and splitting the reversible reactions into two forward and reverse reactions, one can obtain a vector v where each element is greater than or equal to 0.

Since the experiments are carried out in a defined medium, environmental constraints must be added *in silico* simulations as well. These constraints include nutrient availability (such as carbon, nitrogen, or oxygen sources), the pH value, temperature, or any other experiment-specific conditions. Additionally, spatial constraints can be given to the system to limit substrate and enzyme availability for a specific compartment, especially for the metabolites which are transferred between compartments by facilitated diffusion. There are several methods used in constraint-based modelling (Figure 2.7) and additional constraints are introduced into the field every day.

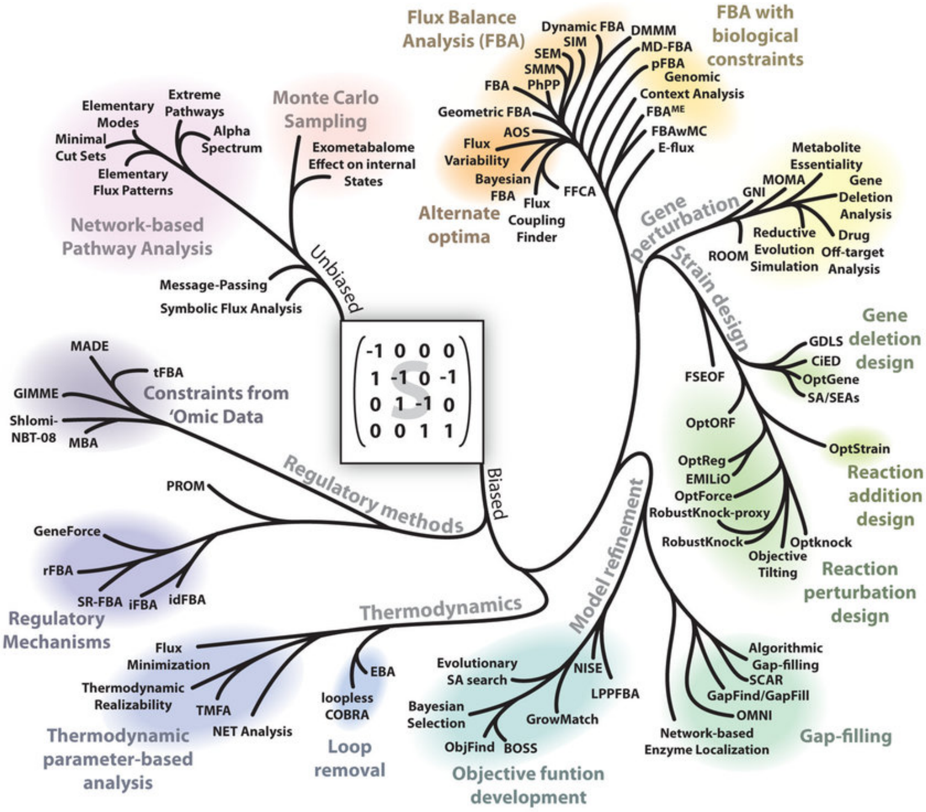


Figure 2.7. Various constraint-based modeling methods [41].

Additional to the previously mentioned methods, more recently, regulatory constraints have started to be used to simulate experimental observations on cellular regulatory mechanisms, such as variable amounts of gene products (transcriptional, translational regulations) and their activities (enzymatic regulations). Integration of omics data, mainly transcriptome data, is used to capture these regulations *in silico*.

2.3. *Saccharomyces cerevisiae*

The species "yeast" includes a range of eukaryotic single-celled microorganisms, although it is commonly used to describe *Saccharomyces cerevisiae.* Also known as the baker's yeast, *S. cerevisiae* is one of the extensively used microorganisms for alcoholic fermentation of beverages, bio-ethanol production, and processing various foods since ancient times [42].

Yeast was the first eukaryotic organism whose genome was fully sequenced and annotated [43], and besides its benefits in the industry, it is used as a model system for other eukaryotic cells including humans [44, 45].

2.3.1. Central Carbon Metabolism of *S. cerevisiae*

From the end of the eighteenth century, mainly after the fermentation is defined as respiration without oxygen, the metabolism of *S. cerevisiae* has been studied extensively [46, 47]. Its capability to produce ethanol is one of the most characterized microbial processes due to industrial utilization.

The set of anabolic and catabolic reactions in the cell are referred as the metabolism. Glycolysis, pentose-phosphate pathway (PPP), tricarboxylic acid cycle (TCA) or Krebs cycle, the glyoxylate cycle and the electron transport chain are the main pathways in central carbon metabolism.

Glycolysis is a cytosolic pathway where 1 molecule of glucose breaks down into 2 molecules of pyruvates while 1 molecules of ATP on the substrate level and 2 molecules of NADH. Within the glycolysis, important precursors for biomass such as 3-phosphoglycerate, 3PG, PEP are generated. The final product of glycolysis, pyruvate, is an important branch point metabolite in the central carbon metabolism.

The end product of glycolysis, pyruvate, can be converted into acetaldehyde through pyruvate decarboxylase (PDC) and, to ethanol with consumption of 1 NADH molecule by alcohol dehydrogenase (ADH). Acetaldehyde is also an important branch point metabolite and it can be converted into acetate by the acetaldehyde dehydrogenase (AcDH) with reduction of 1 NADP+ to NADPH. Acetate can be converted into acetyl-CoA by the acetate synthase (ACS) in the cytoplasm, and used as precursor for biosynthesis of secondary metabolites and in lysine biosynthesis. The stoichiometric reaction for the glycolysis can be written as

$$\begin{aligned} \text{Glucose} + 2\text{ADP} + 2\text{P} + 2\text{NAD+} \rightarrow \\ 2\text{Pyruvate} + 2ATP + 2NADH + 2H2O + 2H+ \end{aligned} \tag{2.2}$$

where glucose, ADP, phospate and NAD+ are consumed to produce pyruvate, ATP, NADH, water and hydrogen.

Some enzymes in the glycolytic pathway are regulated allosterically. For example, presence of ATP in the cytoplasm inhibits the hexokinase (HXK) which deactivates the first step of glycolysis [48]. It is also known that during low concentrations of glucose-6-phosphate (G6P) and fructose 6-phosphate (F6P), mitochondrial oxidative phosphorylation enzymes are activated and it is antagonized by the fructose 1,6- bisphosphate (F1-6bP) [49]. This phenomenon shows that the metabolite F1-6bP inhibits the respiratory flux rates.

To determine whether the regulation of metabolism was taking place at transcriptional and translational level or through metabolic regulation, microarray studies and obtained transcriptomic data are analyzed by ter Kuile and Westerhoff [50]. By using mathematical modeling, they concluded that regulation is almost never completely at translational level and metabolic regulation points must be investigated. Daran-Lapujade, et al., found the correlation between transcript levels, flux profiles and enzyme activity is very little *in vivo* [51].

Gluconeogenesis, as a form of a reversed glycolysis, generates glucose-6-phosphate (G6P) from amino acids (except lysine and leucine) and, C2 and C3 compounds. C2 compounds can also be converted into C4 dicarboxylic acids bypassing oxidative decarboxylation through glyoxylate cycle. During fully respiratory conditions, if the available glucose ratio is above 1%, synthesis of the enzymes for glyoxylate cycle is repressed while oxidative decarboxylation is bypassed. Although this pathway is considered as a modification of the TCA cycle, the localization of this pathway is still unknown.

Krebs cycle or the citric acid cycle (TCA cycle), forms the compounds NADH, CO_2 and $FADH_2$. For 1 mole of pyruvate oxidized in TCA cycle, 4 mol of NADH, 1 mol of $FADH_2$, 3 mol of CO_2 and 1 mol of GTP are generated. The overall stoichiometric reaction for the TCA cycle can be written as

$$\begin{split} Pyruvate + 3H_2O + GDP + 2P + FAD + 4NAD+ \rightarrow \\ 3CO_2 + GTP + 4NADH + FADH_2 + 4H+ \end{split} \tag{2.3}$$

where pyruvate, water, GDP, phosphate, FAD and NAD+ are consumed to produce GTP, NADH, $FADH_2$, hydrogen and carbon dioxide.

The TCA cycle is mostly regulated at transcriptional level and it is known that the cycle is repressed by high glucose concentrations [52]. The main regulatory points in the TCA cycle are the citrate synthase, isocitrate dehydrogenase and α-ketoglutarate dehydrogenase. The cycle is also regulated by feedback inhibition mechanisms and regulation is highly controlled at enzymatic level. For example, the isocitrate dehydrogenase (IDH) is increases its activity with the presence of AMP and inhibited by the presence of ATP [53]; while the low level of NADH/NAD+ molecules activates the majority of enzymes in the cycle [54]. The TCA cycle is a major pathway in mitochondria and it is used to generate important precursor metabolites such as citrate, oxaloacetate, and succinyl-coA.

The reduced cofactors generated from metabolic pathways in the mitochondrial matrix, such as NADH molecules, are re-oxidized to H2O molecules by NADH:ubiquinone oxidoreductase (also called internal NADH dehydrogenase) which is localized on the inner membrane on mitochondria [55]. This oxidation process is called mitochondrial respiration, or oxidative phosphorylation, and it provides a way to the cell to generate energy in the form of ATP molecules. NADH molecules generated in the cytosol can also be oxidized in yeast by external NADH dehydrogenase, or it can be taken inside mitochondria via glycerol-3-phosphate shuttle to be coupled with the respiratory chain, where the ubiquinone pool donates [56] its electrons to the cytochrome-c through the bc1-complex, and oxidation of cytochrome-c by is catalyzed by cytochrome-c oxidase [57]. Similar to NADH molecules, produced $FADH_2$ molecules in the TCA cycle (during oxidation of succinate into fumarate) act as electron donors for ubiquinase [58].

Overall, the central carbon metabolism is finely tuned throughout the organic evolution to supply required energy and to produce building blocks for a yeast cell. All the pathways mentioned are interconnected through metabolites (reaction products or substrates) and/or cofactors (such as ATP, NAD, and NADP) within the cell. These pathways are activated or down-regulated by a tight regulation in order to satisfy what is required according to environmental conditions.

The most characterized regulatory effects of yeasts can be listed as 1) Pasteur Effect, the inhibition of fermentation in the presence of oxygen with a decreased affinity for sugar uptake [59], 2) Kluyver Effect, for the species *Candida utilis*, *Kluyveromyces wickerhamii* and *Debaryomyces yamadae*, being able to ferment glucose anaerobically but not being able to ferment maltose, lactose, and sucrose [60], 3) Custer Effect, the species *Dekkera* and *Brettanomyces* ferment glucose to ethanol faster under aerobic conditions [61], and 4) Crabtree Effect, the suppression of respiration by high glucose condition.

2.3.2. Adaptive Evolution Studies on *S. cerevisiae*

Various strategies for evolutionary engineering in *S. cerevisiae* are developed throughout the years. These strategies can be collected as 1) serial transfers in shake flasks, 2) sequential batch reactor cultivations, 3) continuous cultivations, 4) dynamic selection regimes, and 5) synthetic selection circuits [62].

Ho, Ping-Wei, et al. successfully obtained a faster glycerol utilization in two evolved derivatives by serial transfer in synthetic medium with glycerol as sole carbon source. After 55 generations, evolved strains were able to grow in synthetic glycerol medium, and evolved strains showed a higher growth rate on glycerol [63].

New *S. cerevisiae* strains that grew as fast in the absence of biotin as in its presence are obtained by Bracher, Jasmine M., et al [64]. The evolved strains were obtained by sequential batch reactor cultivation on synthetic medium in the absence of biotin, and showed 32-fold increased maximum specific growth rates.

Smith, Van Rensburg, and Görgens improved xylose fermentation in *S. cerevisiae* strain D5A+ in the presence of lignocellulosic inhibitors by ethyl methanesulfonate (EMS) mutagenesis and anaerobic chemostat cultivation after 100 generations [65]. The study showed that the laboratory evolution profoundly altered the yeast metabolism, provided that evolved strains achieved 7.5-fold reduction of lag phase and were able to remove HMF, furfural and acetic acid completely in 24 h.

Ethanol tolerance is the most important trait of industrial yeasts, and the ability to grow in high levels of ethanol is an industrially desired feature without question. Voordeckers, Karin, et al. improved ethanol tolerance in S288c strain FY5 by aerobic turbidostat cultivation on complex medium with gradually increasing concentrations of ethanol (from 6% up to 12%). After 200 generations, up to 2.5-fold increased fitness was achieved in medium containing 9% ethanol and growth in the presence of 12% ethanol [66].

Acid stress is another important factor in productive purposes of industrial yeasts. Fletcher et. al. set up an ALE experiment where three parallel lines were exposed to different levels of pH stress to elucidate long term adaptation to acidic environments in CEN.PK strain [67]. The pH of the colonies were steadily increased or decreased, depending on the line, and the cultures were serially cultivated after 3-7 generations. The evolved populations were obtained after approximately 290 generations, and the maximum specific growth rates were increased 31.3% (pH decreased by HCl), 200% (pH increased by L-lactic acid) and 11.5% (pH decreased by lactic acid).

Salt stress tolerance is a well-characterized trait for cells, and genomic studies previously proved that *S. cerevisiae* contains multiple conserved pathways in the salt stress response [68]. Considering that, Dhar et. al. exposed yeast strain BY4741 to high sodium chloride (NaCl) concentrations to initiate and investigate salt stress response. In the study where the grown cultures were serially transferred each day, after 300 generations they have found that the high salt medium which also introduces osmolarity stress, reduces the growth rates 11%, whereas obtained evolved strains were able to grow 8–12% faster in salty media.

In addition to the above-mentioned examples, Prof. Z. Petek Cakar's lab at Istanbul Technical University has been a pioneer in evolutionary engineering (ALE) studies in Turkey, producing many evolved *S. cerevisiae* strains that are resistant to diverse stress types [69–75].

2.3.3. Metabolic Models of *S. cerevisiae*

After the first *S. cerevisiae* genome sequence is published, the first cDNA spotted microarray exploring metabolic gene regulation in 1997 [76], and the first commercial platform for oligonucleotide microarray data (Affymetrix) to investigate cellular regulations were reported in 1998 [77]. Existing genome data is integrated with the extensive annotation based on microarray data and biochemical knowledge from literature, leading of the publication of the first GSMM of *S. cerevisiae* in 2003 [78].

The first genome-scale model of S. cerevisiae, iFF708, contained a total of 1175 reactions, 708 genes and 584 metabolites compartmentalized between the cytosol and mitochondrion. The model was validated and used in *in-silico* gene deletion studies [79] and the calculation of physiological parameters [80].

From the first model reconstruction, more biologically accurate data were integrated into models such as the logical relationship between genes (GPR associations), the comprehensive conservation of mass and charge through elementally and charge-balanced reactions (proton balance), and the full compartmentalization of metabolites and proteins. iND750, the second genome-scale metabolic model of *S. cerevisiae* contained 1149 unique reactions, 750 genes, 646 metabolites and five additional compartments (peroxisome, nucleus, golgi apparatus, endoplasmic reticulum, and vacuole) [81].

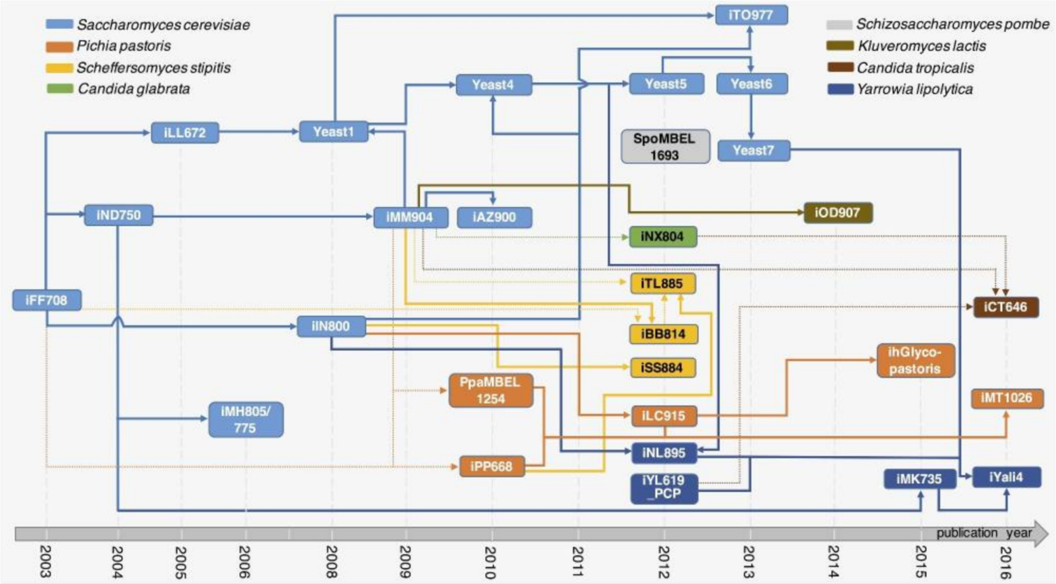


Figure 2.8. Evolutionary timeline of different yeast species genome-scale metabolic models [82].

Afterwards, the model iLL672 was reconstructed based on the older iFF708 with improved connectivity of the network [83]. Previous models were reconciled by deleting duplicates and multiple dead-end reactions, resulting a model with a total of 636 metabolites and 1038 reactions. Improvements of the metabolic models are followed by the iIN800 [84], which tRNA synthesis, transport processes and a detailed lipid metabolism included.

Later, an improved version of the iND750 model, named iMM904, is reconstructed introducing a new database-independent nomenclature for 1226 metabolites, 1577 reactions, and 905 genes such as SMILES or InChI strings [85]. The idea of the reconstruction of the iMM904 was to provide a common model that serves as a base for the systems biology of yeast. Additionally, authors claimed that the integration of the exometabolomic data into the model allowed to essentiality predictions using GSMMs.

Corrections to the iMM904 model formed the next model, named iAZ900 containing 900 genes, 1240 metabolites and 1602 reactions distributed among seven compartments [86]. In addition to the corrections on reversibility of the reactions, GPR associations and biomass definition, revision of the previous model included addition/removal of reactions, metabolites and/or genes.

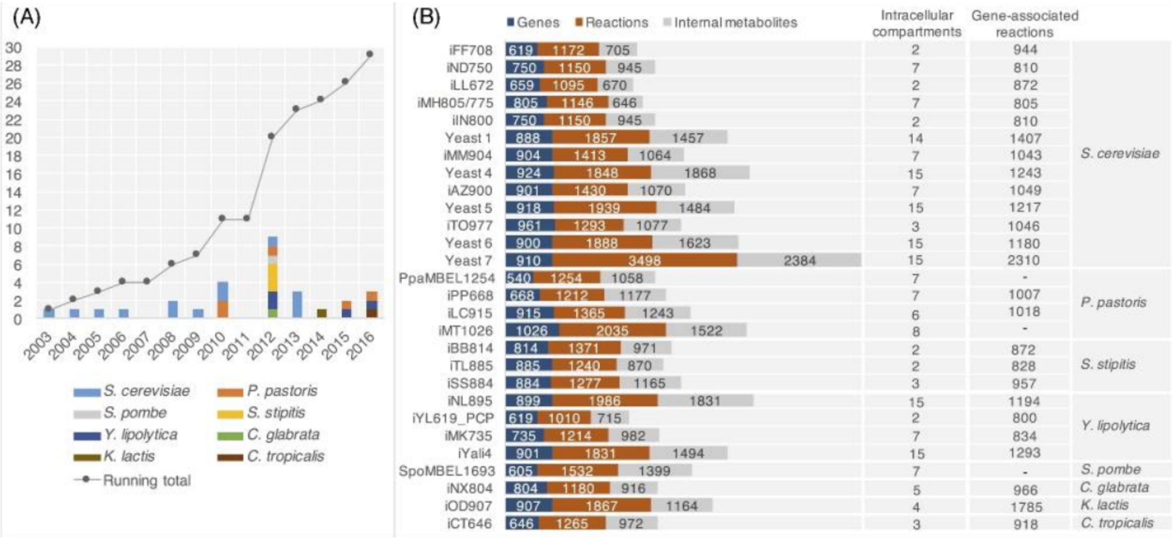


Figure 2.9. Published genome-scale models of yeast in gene, reaction, metabolite and compartment numbers [82].

Yeast1, the first collaboratively constructed consensus genome-scale metabolic network was built using standardized identifiers [87]. The consensus model was cumulatively improved despite its inability to perform computational simulations. With Yeast4 [88], which includes metabolite transport reactions, an improved lipid metabolism (based on iIN800) and improved network connectivity, a consensus model was able perform simulations. Following updates on the consensus model included expansions on the lipid metabolism with Yeast 5 [89]; improved predictions with Yeast 6 [90]; and major changes in the fatty acid, glycerolipid, and glycerophospholipid metabolisms with Yeast 7 [91] models. The latest version of the consensus model of *S. cerevisiae*, Yeast8 [92], is released in the last year and continues to be updated by the community.

2.3.4. Applications of *S. cerevisiae* GSMMs

Several genome-scale metabolic models of *S. cerevisiae* have been used in metabolic engineering applications. These studies were carried out for the development of yeast cell factories to increase yield industrially. Despite the high number of GSMM availability, only a selection of the models has been successfully validated experimentally.

For improved bioethanol production, Bro et al. simulated a number of strategies that increased ethanol yields under anaerobic conditions using the model iFF708 [93]. FBA analyses showed the alterations on the redox metabolism, and the resulting strain had reduced glycerol yield 40% on glucose and increased ethanol yield (+3%) with the same maximum specific growth rate.

iMM904 model is successfully used with Optknock framework [94] (an FBA-based algorithm to predict gene deletion strategies) and new industrially preferable strains, which has improved 2,3-butanediol production, were designed. This successful implementation was the first report about increased 2,3-butanediol yield, and designed strains were validated experimentally [95].

Another report was the producing fumaric acid by direct fermentation in yeast, using the model iND750 [96]. This prediction was also based on literature mining, i.e. the target gene was selected after intense research, and FBA analyses on the metabolic model showed that deletion of the selected gene can lead to fumaric acid production without any significant change in growth rates.

The same iMM904 model was also used by Sun et al., to obtain a strain with 8- to 10-fold increased amorphadiene product yield compared to the wild type [97]. Results gathered by identifying the knockout positions with a potential to promote terpenoid production in the cell. Predictions were validated by the constructed single mutant and the study indicated that the metabolic models are powerful tools in the engineering of knockout strains before performing any molecular manipulations.

3. MATERIALS AND METHODS

3.1. Consensus *S. cerevisiae* Metabolic Model

Variety of *S. cerevisiae* genome-scale metabolic models have been used since 2003, and each reconstructed model introduced more manual curations, increasing gene numbers from annotations and better predictions regarding the previous ones [82]. A consensus genome-scale metabolic model of *S. cerevisiae*, Yeast8, is presented in an open-source, version-controlled maintainable way in 2019, claiming that the model can be represented and investigated in a systematic way using Git and GitHub as a hosting service for the model repository [92]. Systematic way of Yeast8 enables to study simultaneously in collaborative studies, provides record keeping of model changes, version updates, where each version of it can be released periodically and accessible all the time.

Yeast8 model can be considered as an updated version of Yeast7 [91] with additional corrections based on the annotations available in KEGG and ChEBI, and several gene inclusions from the model iSce926 [98]. Final version of Yeast8 to date, version 8.3.4 released on July 28, has 3991 reactions, 2691 metabolites, 1149 genes and 14 intracellular compartments. Additional statistical analysis on its stoichiometric matrix can be seen in Figure 3.1.

All simulations in the Methods section are done on the model Yeast8 v8.3.4 which is hosted in Github. Optimization problems were solved with the COBRA Toolbox v3.0.6 and the RAVEN Toolbox v2.3.1 in MATLAB (version 9.7.0.1216025 (R2019b)), using Gurobi solver (version 8.1.1).

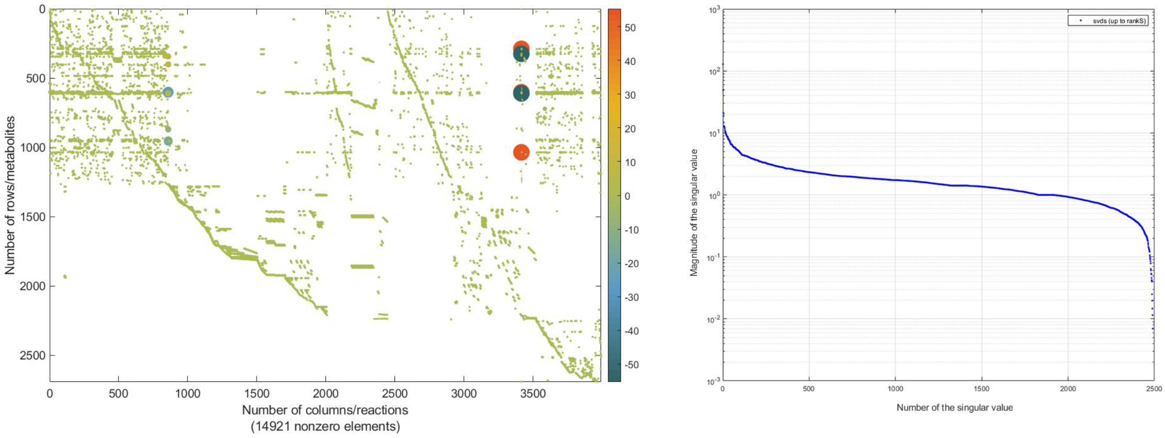


Figure 3.1. Coefficients and singular values of the stoichiometric matrix of Yeast8.

3.2. Experimental Data Acquisition

Extracellular metabolomics data are obtained from Cakar's Lab [99]. Briefly, they performed ethyl methane sulfonate (EMS) mutagenesis on the prototrophic *Saccharomyces cerevisiae* strain CEN.PK 113-7D (MATa, MAL2-8c, SUC2) as reference, and the cells were successively batch cultivated under gradually increasing levels of stress factors as a selection strategy by evolutionary engineering. After selection, the evolved strains were inoculated in 2% Yeast Minimal Media (YMM), and the extracellular concentrations of glucose, ethanol, glycerol and acetate were measured at different time points by using HPLC. OD_{600} values were determined by a spectrophotometer. Additionally, cell dry weight analysis was conducted to determine biomass production. Measured OD_{600} values and dry weights of the reference strains (without mutagenesis) of all ALE experiments were used in this study, and are collected in Table 3.2. Experimental extracellular metabolite concentrations are also collected in Table 3.1, and compared with the computational simulation results in the discussions.

In order to determine uptake and secretion rates of the metabolites, the steady-state assumption is applied in three hours intervals as the shortest measured time-points. Missing data on cell dry weights are estimated from the OD_{600} values, and these cell dry weight data is used to calculate fluxes (in the unit of mmol/(gDW)$^{-1}$).

Measurement of the cell dry weight at the 3rd hour was crucial for the steady-state assumption, however data was not available from the experiments. Curve trend of the OD_{600} plot is used as a guide to estimate cell dry weight (Figure 3.2).

Table 3.1. Measurements of extracellular concentrations [70].

Time(h)	**Glucose(g/L)**	**Ethanol(g/L)**	**Glycerol(g/L)**	**Acetate(g/L)**
0	19.99	0	0	1.08
3	17.98	0.58	0.02	1.24
6	15.85	1.2	0.06	1.16
9	12.21	3.39	0.18	1.37
12	9.18	7.97	0.61	2.45
15	0.4	8.17	0.69	2.46
27	0	8.28	0.76	2.6
46	0	8	0.77	2.45
50	0	6.62	0.64	2.02
54	0	5.74	0.55	1.73
58	0	5.46	0.54	1.74
72	0	3.72	0.49	1.33

As the slope in the curve of $\ln\mathrm{OD}_{600}$ as a function of time gives the growth rates of cells, natural logarithm of OD_{600} values were calculated to obtain specific growth rates by using the Equation (3.1)

$$\mu = \frac{\Delta \ln OD_{600}}{\Delta t} \tag{3.1}$$

where OD_{600} is the optical density of a sample measured at a wavelength of 600 nm, and t is the time.

Table 3.2. Measured OD_{600} and cell dry weight values of reference strain [70].

Time (h)	**OD600**	**ln(OD600)**	**Cell DW (g/L)**
0	0.21	-1.560647748	-
3	0.53	-0.634878272	-
6	1.76	0.565313809	0.9
7.5	2.66	0.978326123	-
9	4.46	1.495148766	1.9
12	5.31	1.669591835	-
15	5.88	1.771556762	-
18	5.83	1.763017	2.32
21	6.07	1.803358605	-
24	5.87	1.769854634	-
30	6.14	1.814824742	2.26

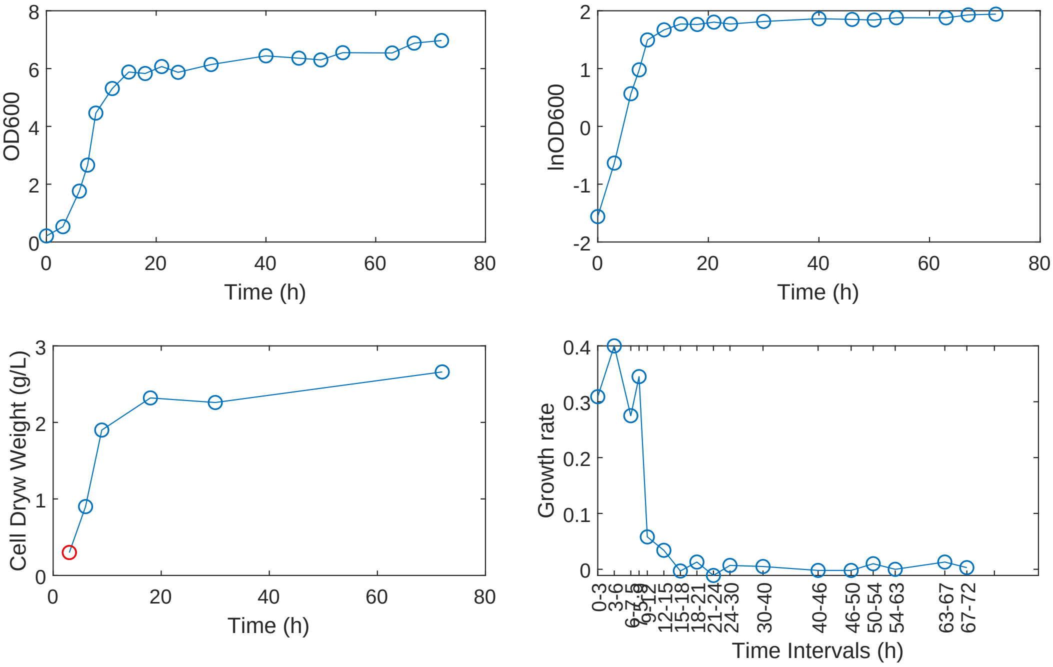


Figure 3.2. OD_{600}, $\ln OD_{600}$, cell dry weights and growth rates graphs. Estimated missing cell dry weight data is shown in red color.

3.3. Flux Balance Analysis

Flux balance analysis (FBA) assumes that the living cells act as they optimized their lives towards some goal, and as if they were at steady state. To be more clear, steady-state assumption indicates that the metabolites are both produced and consumed at the same rate in a cell, without an accumulation. Therefore, in this system, metabolites are constrained by only the stoichiometric coefficients arise from mass balance of metabolites.

FBA solves a set of ordinary differential equations regarding to

$$S_{m\times n}\cdot v=0 \tag{3.2}$$

where S is the matrix of the stoichiometric reaction coefficients with m number of metabolites (as rows) and n number of reactions (as columns), and v is the vector of all associated reaction fluxes (mmol/(gDW)$^{-1}$). Because the matrix S usually has more reactions than metabolites, the system can result multiple solutions, and being called an underdetermined system.

To solve it for an optimal solution, additional constraints or an objective function is required. A "growth reaction" is usually included in the reactions of the system to represent the "goal" in the definition of living systems. Growth reactions act as the final consumption of metabolites necessary for the biomass production or cell replication.

Additional to the growth, several exchange reactions (uptake or secretion of metabolites from or into extracellular space) are also included. Since the concentrations of extracellular metabolites are measurable experimentally, constraints can be applied to exchange reaction fluxes to shrink solution space.

The more constraints introduced into the system, such as reversibility of reactions or known rate values, the more smaller solution space results. The growth reaction is usually used as an objective function to determine a unique solution from this solution space. The linear problem appears as

$$\begin{aligned} \max_v \quad & c^T \cdot v \\ \text{subject to} \quad & S_{m\times n} \cdot v = 0 \\ & v_{lb} \leq v \leq v_{ub} \end{aligned} \tag{3.3}$$

where c is the objective function vector, v is the vector of fluxes, S is the stoichiometric matrix as in above equation. Subscripts lb and ub are the lower and upper boundaries on v. These constraints define a feasible region of the problem. The coefficients of the biomass constituents are defined as the same as the batch conditions in the reference article [100], for the reason that detailed knowledge is not available in the acquired experimental data (see section 3.2). The coefficients for the final biomass equation can be found in the Table 3.3.

There are multiple available command-line functions to solve a linear programming problem such as *linprog()* function of MATLAB, *optimizeCbModel()* function of the COBRA Toolbox, or *solveLP()* function of the RAVEN Toolbox with using variety of solvers such as GLPK, GUROBI or IBM cplex.

In order to simulate batch conditions where the YMM is used as media, all the exchange reactions in the model are first blocked (lower bounds are set to 0). Then, only the exchange reactions of the ions that are available to the cells in the experimental design (ammonium, phosphate, sulphate, iron^{2+}, H^{+}, water, chloride, Mn^{2+}, Zn^{2+}, Mg^{2+}, sodium, Cu2^{2+}, Ca^{2+}, potassium) are set free (lower bounds are set to -1000), meaning that cells can uptake then as needed.

Table 3.3. Biomass formulation that are used in the Yeast8.

Products	**Components**
Biomass	lipid + protein + carbohydrate + RNA + DNA + cofactor + ion
Lipid	lipid backbone + lipid chain
Lipid Backbone	0.0063196 1-phosphatidyl-1D-myo-inositol backbone + 0.024311 ergosterol + 0.0062241 ergosterol ester backbone + 0.0013536 fatty acid backbone + 0.0054422 phosphatidyl-L-serine backbone + 0.023579 phosphatidylcholine backbone + 0.006337 phosphatidylethanolamine backbone + 0.006271 triglyceride backbone
Lipid Chain	0.0073947 C16:0 chain + 0.021702 C16:1 chain + 0.0020726 C18:0 chain + 0.0079624 C18:1 chain
Protein	0.57284 Ala-tRNA(Ala) + 0.20064 Arg-tRNA(Arg) + 0.12698 Asn-tRNA(Asn) + 0.37145 Asp-tRNA(Asp) + 0.0082405 Cys-tRNA(Cys) + 0.1316 Gln-tRNA(Gln) + 0.37682 Glu-tRNA(Glu) + 0.36258 Gly-tRNA(Gly) + 0.08278 His-tRNA(His) + 0.2406 Ile-tRNA(Ile) + 0.37007 Leu-tRNA(Leu) + 0.35734 Lys-tRNA(Lys) + 0.063302 Met-tRNA(Met) + 0.16718 Phe-tRNA(Phe) + 0.20564 Pro-tRNA(Pro) + 0.23148 Ser-tRNA(Ser) + 0.23897 Thr-tRNA(Thr) + 0.035459 Trp-tRNA(Trp) + 0.12735 Tyr-tRNA(Tyr) + 0.33037 Val-tRNA(Val)
Carbohydrate	0.68454 (1->3)-beta-D-glucan + 0.22871 (1->6)-beta-D-glucan + 0.33052 glycogen + 0.65017 mannan + 0.12646 trehalose
RNA	0.044535 AMP + 0.043276 CMP + 0.044535 GMP + 0.057992 UMP
DNA	0.0036 dAMP + 0.0024 dCMP + 0.0024 dGMP + 0.0036 dTMP
Cofactors	0.00019 coenzyme A + 1e-05 FAD + 0.00265 NAD + 0.00015 NADH + 0.00057 NADP(+) + 0.0027 NADPH + 0.00099 riboflavin + 1.2e-06 TDP + 6.34e-05 THF + 1e-06 heme a
Ions	3.04e-05 iron(2+) + 0.00363 potassium + 0.00397 sodium + 0.02 sulphate + 0.00129 chloride + 0.00273 Mn(2+) + 0.000748 Zn(2+) + 0.000217 Ca(2+) + 0.0012425 Mg(2+) + 0.000659 Cu2(+)

3.4. Integration of Enzymatic Constraints

Solution space of the yeast model can be further constrained by limiting fluxes of reactions with integration of enzymatic information by using GECKO (GEM with Enzymatic Constraints using Kinetic and Omics data) method [101]. In GECKO methodology, if a reaction has an annotated enzyme, enzyme entity itself is included in the same reaction equation. Stoichiometric coefficient of this entity is defined by its kinetic information and abundance in the cell (Figure 3.3). S matrix of the system is expanded with the addition of new rows for the enzymes and new columns for each enzyme's usage.

Metabolites A and B can be written in equations such as

$$A \xrightarrow{E_i} B \tag{3.4}$$

$$A + n_{ij}E_i \rightarrow B \tag{3.5}$$

where E_i is the enzyme catalyzing the convertion of A to B. The idea behind this methodology arises from the biochemical constraints, such that the maximum rate of a reaction is smaller or equal to its enzyme's turnover number and the concentration (mmol/gDW), and the equation

$$v_j \leq k_{cat}^{ij} \cdot [E_i] \tag{3.6}$$

shows the condition for the reaction flux depending on enzyme concentration. Enzymes in the reaction equations must be considered as pseudo-metabolites because they do not affect the mass balance of the reaction, i.e., enzymes are not consumed in the reactions but they are occupied for a short period of time.

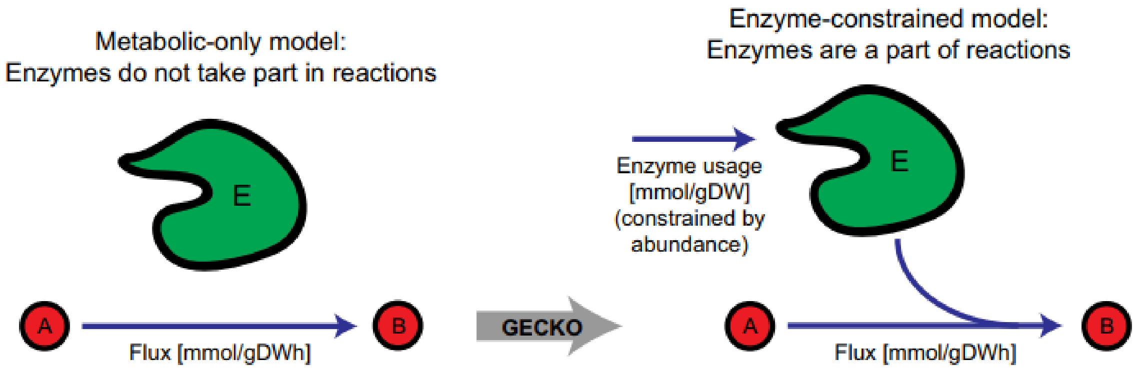


Figure 3.3. GECKO methodology to integrate enzymatic constraints [101].

To maintain mass balances of enzymes, an overall enzyme usage pseudo-reaction is introduced into system in order to supplement enzymes into reactions, similar to exchange reactions, as the equation

$$EU_i : \quad \rightarrow E_i \tag{3.7}$$

shows the draw reaction of an individual enzyme. If the flux carried by this reaction is e_i, we can say that e_i ranges from 0 (i.e., no enzyme available) to a maximum value of $[E_i]$ (i.e., all the enzyme is used in the reaction)

$$0 \leq e_i \leq [E_i] \tag{3.8}$$

then a mass balance for the enzyme E_i can be defined under steady state assumption

$$-n_{ij} \cdot v_j + e_i = 0 \tag{3.9}$$

arises. By rearranging the Equation (3.8) and Equation (3.9), we obtain the equation

$$v_j \leq \frac{1}{n_{ij}} \cdot [E_i] \tag{3.10}$$

and, if we compare this equation with the Equation (3.6), we obtain

$$n_{ij} = \frac{1}{k_{cat}^{ij}} \tag{3.11}$$

where we can finally calculate the stoichiometric coefficient for the enzyme. However, the Equation (3.6) does not hold in all cases since reactions can be more complex. For isozymes in which multiple enzymes catalyze the same reaction, the equation becomes

$$v_j \leq \sum_i k_{cat}^{ij} \cdot [E_i] \tag{3.12}$$

and, since the reaction can be catalyzed by all the isozymes equally, new reactions can be defined for each isozyme (note that each enzyme can have different k_{cat} values). For example, if we think that the Equation (3.4) had two isozymes, $A \xrightarrow{E_1 or E_2} B$, the new reactions

$$\begin{aligned} R_{j/1}: &\quad \frac{1}{k_{cat}^{1j}} E_1 + A \rightarrow B \\ R_{j/2}: &\quad \frac{1}{k_{cat}^{2j}} E_2 + A \rightarrow B \end{aligned} \tag{3.13}$$

would be defined. These additional reactions distrupt the boundaries for the initial reaction. In order to maintain original boundaries, another "arm-reaction"

$$\begin{aligned} R_{j/arm}: &\quad A \rightarrow M_{int} \\ R_{j/1}: &\quad \frac{1}{k_{cat}^{1j}} E_1 + M_{int} \rightarrow B \\ R_{j/2}: &\quad \frac{1}{k_{cat}^{2j}} E_2 + M_{int} \rightarrow B \end{aligned} \tag{3.14}$$

is also introduced to the system where it acts as an intermediate metabolite. For an enzyme complex catalyzing a single reaction, since the reaction is catalyzed by all the enzymes which share the same k_{cat} value, the Equation (3.6) becomes

$$v_j \leq k_{cat}^{ij} \cdot \min_k \frac{[U_{ik}]}{S_{ik}} \tag{3.15}$$

where $[U_{ik}]$ is the concentration of the subunit of the catalyzing enzyme E_i, and S_{ik} is the stoichiometry of the subunit. For the reaction, if we think that the Equation (3.4) was catalyzed by an enzyme complex of two subunits, namely E_1 and E_2, the equation

$$R_j : \quad \frac{S_1}{k_{cat}^{ij}} E_1 + \frac{S_2}{k_{cat}^{ij}} E_2 + A \rightarrow B \tag{3.16}$$

where two of the subunits for the enzyme complex included in.

Finally, for the reversible reactions, two reactions must be defined for both forward and backward direction reactions with the same catalyzing enzyme but possibly with different k_{cat} values. Assume the Equation (3.4) is reversible, the equations

$$R_{j/f} : \quad \frac{1}{k_{cat}^{ij/f}} E_i + A \rightarrow B \tag{3.17}$$

$$R_{j/b} : \quad \frac{1}{k_{cat}^{ij/b}} E_i + B \rightarrow A \tag{3.18}$$

would be introduced to the system.

3.5. Differential Expression Analysis and Integration

Normalized gene expression profiles of the GSE series (Table 3.4) obtained from the Gene Expression Omnibus are first differentially analyzed using Biobase [102], GEOquery [103] and limma [104] packages in R environment. In the analysis briefly, design matrices are constructed for each ALE experiment separately. Contrast matrices are obtained to make pair-wise comparison between the reference and the evolved strains. Linear models for each gene are fitted to the expression data for each contrast matrix, and empirical Bayes method was used on the fitted models for the assessment of differential expressions. Genes with the p-values of <0.05 are selected to be integrated into the constrained-based model.

Table 3.4. GSEQuery codes of used expression data.

Experiment	**Query**	**Strain Names**	**Publication DOI**
ethanol	GSE78759	REF vs. B2 vs. B8	[69]
caffeine	GSE124452	REF vs. Caf905-2	[70]
coniferylaldehyde	GSE119240	REF vs. BH13	[71]
iron	GSE61317	REF vs. M8FE	[72]
nickel	GSE50985	REF vs. M9	[73]
phenylethanol	GSE59353	REF vs. C9	[74]
silver	GSE143335	REF vs. 2E	[75]

Obtained relative fold changes of the genes are integrated into metabolic model as the multipliers of the enzyme coefficients that are used to draw enzymes from the enzyme pool as in Equation (3.7). In that sense, the gene products, i.e., the proteins with higher fold-change values (> 1) are considered over-expressed and they are easily drawn from the pool, while proteins with lower fold change values (< 1) are considered under-expressed and the reactions that are using these enzymes are become more costly to the system.

3.6. Phenotype Phase Plane and Robustness Analysis

As previously mentioned, there is no single solution to the linear problem of the metabolic model. Phenotype phase planes (PhPP) are used to describe all the feasible metabolic states in two or three dimensional surfaces, depending on the number of metabolites chosen to see how they affect the objective function [105]. Commonly for aerobic models, various levels of glucose and oxygen availability through their uptake reactions are used to generate PhPP surfaces in three dimensions with the objective function of growth. Phenotype phase plane construction refers to a double robustness analysis on the model for selected reactions. Fundamentally, a phenotype phase plane is a two-dimensional plane defined by axes of the exchange rates of the selected nutrients. Since there are finite number of points in this plane, flux balance analysis finds the optimal flux distribution with the maximum or the minimum value of the objective function from all the existing points. The robustness of the network is determined by calculating the optimal flux distribution over iteratively altering the selected nutrient exchange reaction boundaries from maximum (100%) to minimum (0%). Briefly, robustness analysis shows how the objective changes as the flux through the selected reaction varies in magnitude [105].

In Figure 3.4, an example of a phenotypic phase plane for the specific growth rate as a function of two substrates (oxygen and glucose uptake rates) is shown. Four sections in the figures represent the four phases in the cell, i.e., four different metabolic states. In the first section (at low glucose uptake rates), after the investigation on the flux vector, it can be concluded that glucose is fully consumed and converted into carbon dioxide through TCA cycle. In the following sections, several byproducts in the form of weak acids are secreted (acetic acid in the second phase; acetic and formic acid in the third phase). In the fourth phase, ethanol is secreted additional to the acetic and formic acid. Line of optimality (LO) in the phenotype phase planes shows the optimal boundary where the carbon source (glucose) is converted into biomass at a maximum yield.

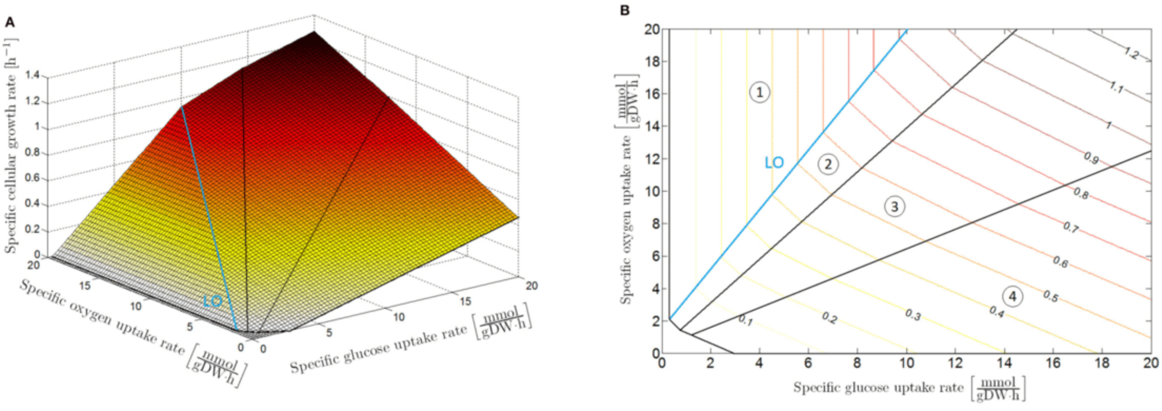


Figure 3.4. An example of a phenotypic phase plane for the specific growth rate as a function of oxygen and glucose uptake rates as (A) 3D plot and (B) contour plot [106].

3.7. Flux Variability Analysis

Flux variability analysis (FVA) finds the minimum and maximum available fluxes for each reaction while obeying the provided constraints (for example fixed glucose uptake or growth rate). FVA is mainly used 1) to evaluate the robustness of the model [107], 2) to find alternative optimum states [108], 3) to check flux distributions when growth is not at optimum level [109]. Many other applications of FVA can be found in the literature [110].

Flux variability analysis solves two optimization problems for each reaction

$$\begin{aligned} \max_v/\min_v \quad & v_i \\ \text{subject to} \quad & S_{m\times n}\cdot v = 0 \\ & w^T \cdot v \geq \gamma \cdot Z_0 \\ & v_{lb} \leq v \leq v_{ub} \end{aligned} \tag{3.19}$$

where w is the objective function equals to c in the problem, Equation (3.3), $Z_0 = w^T \cdot v_0$ describes an optimal solution to the problem, Equation (3.3), γ is an indicator to check whether the FVA is done at the optimal state (where objective flux is the same and $\gamma = 1$) or any other state (where $0 \leq \gamma < 1$).

3.8. Random Sampling of Solution Space

Constraints applied to a model define a solution space, a convex polytope, where every flux distribution is accessible. Random sampling of the solution space is an unbiased tool to explore metabolic models. Mainly, Markov Chain Monte Carlo methods are used to sample this space using algorithms such as (Artificially Centered) Hit-and-Run (HRB) [111, 112] algorithm, and this method has proven to be helpful in the analysis of genome-scale metabolic models [113]. The COBRA Toolbox [114] provides a workflow to generate random points by using the Hit-and-Run algorithm. Briefly, the random sampling method collects points that are uniformly distributed in the solution space and calculates the most probable flux value for each reaction.

Since the computational burden of the loopless sampling is high, generated random points in the solution space of GECKO integrated Yeast8 model include thermodynamically unfeasible states. To this sense, instead of using Hit-and-Run algorithm to sample inside the region of allowed solutions, corners of the solution space is sampled by using the workflow provided by the RAVEN Toolbox [115].

The workflow from the RAVEN Toolbox uses the simplex method, one of the main algorithms in mathematical optimization, to calculate corners of the solution space with a random set of objective functions. Maximization of these random functions (reactions in our case) brings out the corners of the allowed solution space. This workflow also includes flux variability analysis from the COBRA Toolbox in order to prevent internal loops.

3.9. Minimization of Metabolic Adjustment

An approach for predicting mutant behavior similar to FBA is based on the Minimization of Metabolic Adjustment (MOMA). The idea behind this methodology arises because a mutant is more likely to display a sub-optimal flux distribution since the selection through evolutionary adaptation has altered the genetic background of the evolved strains. The application of FBA on both wild-type and mutant strains is considered insubstantial. While the FBA method uses the objective function to present an optimal solution, MOMA choses the objective function as the minimization of the Euclidean distance between the optimal flux distribution calculated by FBA on the reference strain and the sub-optimal flux distribution of the mutant where it remains initially as close as possible to the optimal flux distribution of the wild-type.

Distance in the flux space is a quadratic function, thus, quadratic programming (QP) algorithm is used in MOMA instead of linear programming (LP). The COBRA

Toolbox provides an approach in the form of a function called *MOMA()* which can easily be implemented with command line in MATLAB. Within the function, the following FBA problem is solved first

$$\begin{aligned} \max \quad & c_{wt}^T \cdot v_{wt} \\ \text{subject to} \quad & S_{wt} \cdot v_{wt} = 0 \\ & lb_{wt} \leq v_{wt} \leq ub_{wt} \end{aligned} \tag{3.20}$$

and then the second problem is solved

$$\begin{aligned} \min \quad & \sum (v_{wt} - v_{mut})^2 \\ \text{subject to} \quad & S_{wt} \cdot v_{wt} = 0 \\ & S_{mut} \cdot v_{mut} = 0 \\ & lb_{wt} \leq v_{wt} \leq ub_{wt} \\ & lb_{mut} \leq v_{mut} \leq ub_{mut} \\ & c_{wt}^T \cdot v_{wt} = f_{wt} \end{aligned} \tag{3.21}$$

where f_{wt} is the optimal objective value found by FBA in the first problem.

4. RESULTS

4.1. Model Constraints and Curation

To make sure the *in-silico* growth rate predictions are in agreement with the physiological kinetic parameters obtained from the laboratory experiments, fine adjustment on the chosen metabolic model is a requirement. Since the growth-associated maintenance (GAM) and non-growth associated maintenance (NGAM) reactions in a model play a determinant role in the simulation results, fluxes through these reactions must be constrained to a fixed value. Flux of the NGAM reaction is constrained to 0.7 mmol/gDWh^{-1} for aerobic model, and 0 mmol/gDWh^{-1} for anaerobic model simulations as calculated in the previous studies [100]. Since the GAM depends on the biomass composition defined, results of a chemostat experiment [116] are used as a guide to fit simulation predictions to. The model is simulated iteratively with a range of values for the GAM and the best fit is found at the level of 31.4 mmol/gDWh^{-1} (Figure 4.1).

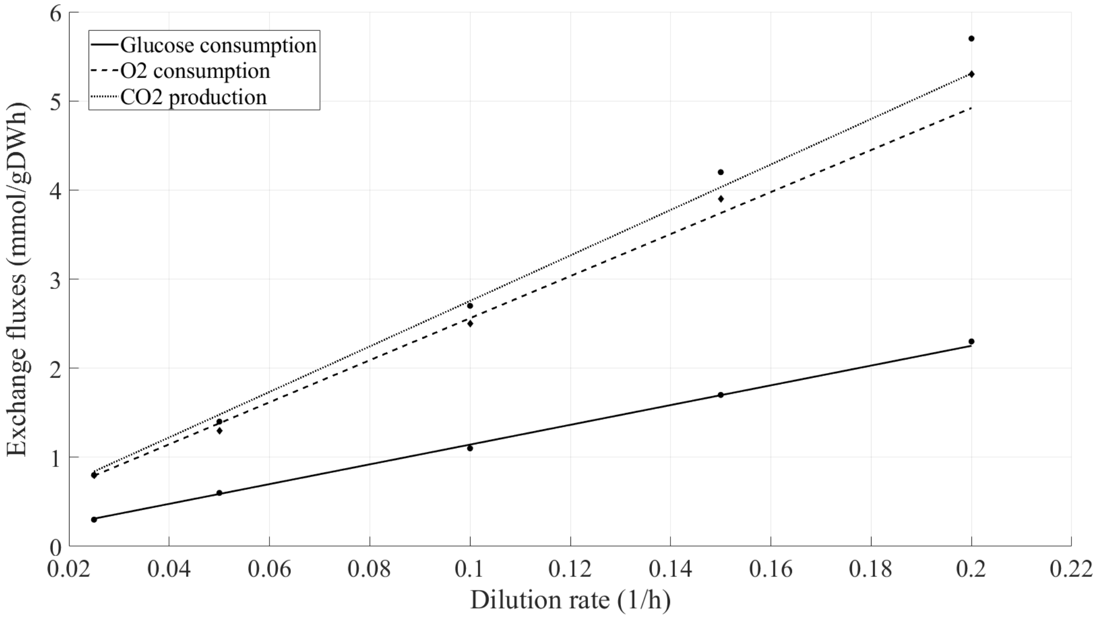


Figure 4.1. The required flux for the growth associated maintenance reaction is determined by the fitting simulation results to the experimental data.

Average enzyme saturation value is set to 50% where the lowest error rate is obtained as a fraction to biomass (Figure 4.2). By constraining the total enzyme mass, the metabolic model became able to show overflow metabolism without any other constraint applied.

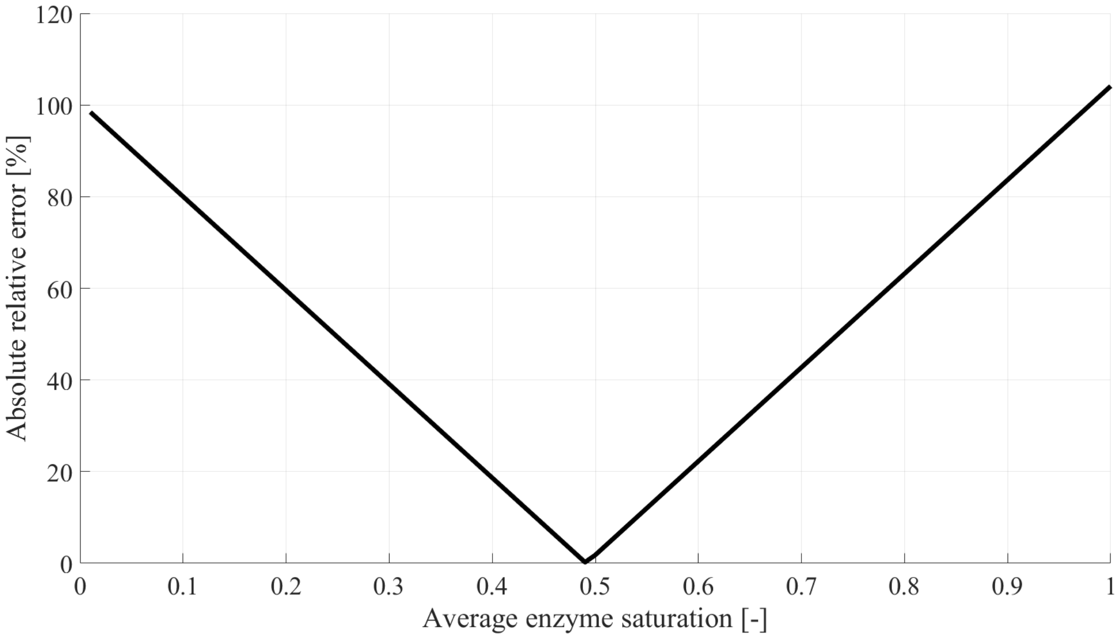


Figure 4.2. Average enzyme saturation factor for the GECKO method is determined by the iterative simulations for the lowest absolute relative error to the experimental data. The value is set to 50% where the lowest error rate is obtained.

To prevent over-constraining the model with the enzyme kinetics, the most limiting proteins in the model are found based on their sensitivity on the objective function, i.e., growth, iteratively. In each iteration, the k_{cat} values for the most limiting enzymes are updated with the maximum available values in BRENDA [117] database. Iterations are repeated until the simulation results for the growth rate agrees with the experimental growth rate with the error less than 10%. These alterations on the k_{cat} values for enzymes with their objective control coefficients can be found in the Table 4.1.

Table 4.1. Iterations on the network to find the top limiting proteins by their objective control coefficients (cc), i.e. sensitivities on the growth, with previous and updated k_{cat} values and error improvements.

#	**Protein**	**Reaction Name**	**prev k_{cat}**	**new k_{cat}**	**CC**	**Err%**
1	P38604	lanosterol synthase	0.002	4.076	0.937	-42.6
2	P38972	5'-phosphoribosylformyl glycinamidine synthetase	0.05	5.07	0.244	-28.7
3	P00931	tryptophan synthase (indoleglycerol phosphate)	0.022	775.75	0.129	-19.4
4	P48445	acetyl-CoA carboxylase	1.23	450000	0.065	-14.2
5	P05694	methionine synthase	0.33	3.5	0.050	-10.3
6	P39006	PS decarboxylase (1-16:1, 2-16:1)	0.053	366.667	0.043	-6.4

4.2. Differential Expression Analysis and Integration

The obtained expression datasets were quantile-normalized for their own experiments. Three replicates for both reference and the evolved strains for each experiment (exception with the ethanol-b8 strain which has two replicates) is used for the differential expression analysis. The distribution across normalized gene expression levels across each experiment can be seen in the Figure 4.3. No further normalization was applied to prevent introducing errors, and datasets are used as obtained from the database.

A statistical threshold ($p < 0.05$) for the differential analysis of the normalized gene expression levels detected a total of 1606 genes in ethanol-b2, 2947 genes in ethanol-b8, 4743 genes in caffeine, 2267 genes in coniferyl aldehyde, 3448 genes in iron, 152 genes in nickel, 4796 genes in phenylethanol, and 4796 genes in silver-resistant strains differentially expressed.

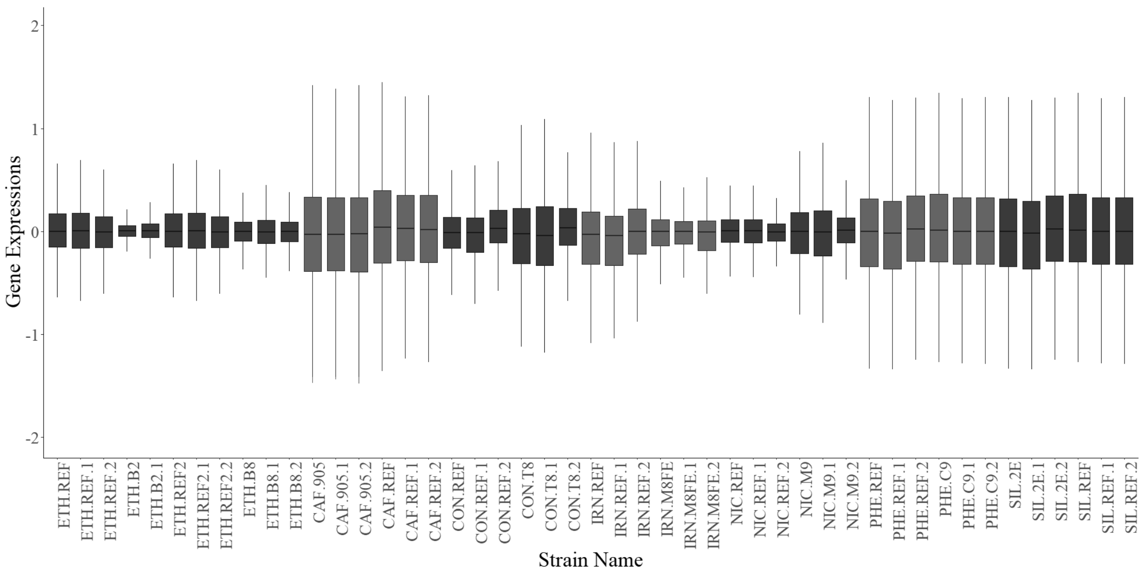


Figure 4.3. Boxplots of the gene expression levels obtained from Gene Expression Omnibus. Quantile normalization performance is indicated by the medians (black horizontal lines) that are almost at the same level with their reference reads.

To find the commonalities between evolved strains in terms of differentially expressed genes, the overlapping genes (that are differentially expressed significantly, $p < 0.05$) were grouped together in Figure 4.4, and differentially expressed genes are plotted separately for upregulated and downregulated genes.

Despite the high number of differentially expressed genes in the silver-resistant strain, the intersection with the greatest number of strains discluded it. There were not any common differentially expressed genes considering all strains. The most common genes (as shown in the first bars in Figure 4.4) are reported in Table 4.2.

The log2(fold-change) values of differentially expressed genes and their frequencies are plotted for each experiment in Figure 4.5. Although the high number of differentially expressed genes are collected from the analysis (1606 genes for ethanol-b2, 2947 genes for ethanol-b8, 4743 genes for caffeine, 2267 genes for coniferyl aldehyde, 152 genes for nickel, 4796 genes for phenylethanol and 4796 genes for silver-resistant strains), due to the limited reaction number in the genome-scale metabolic models, only a part of obtained fold-change expression values were able to be integrated.

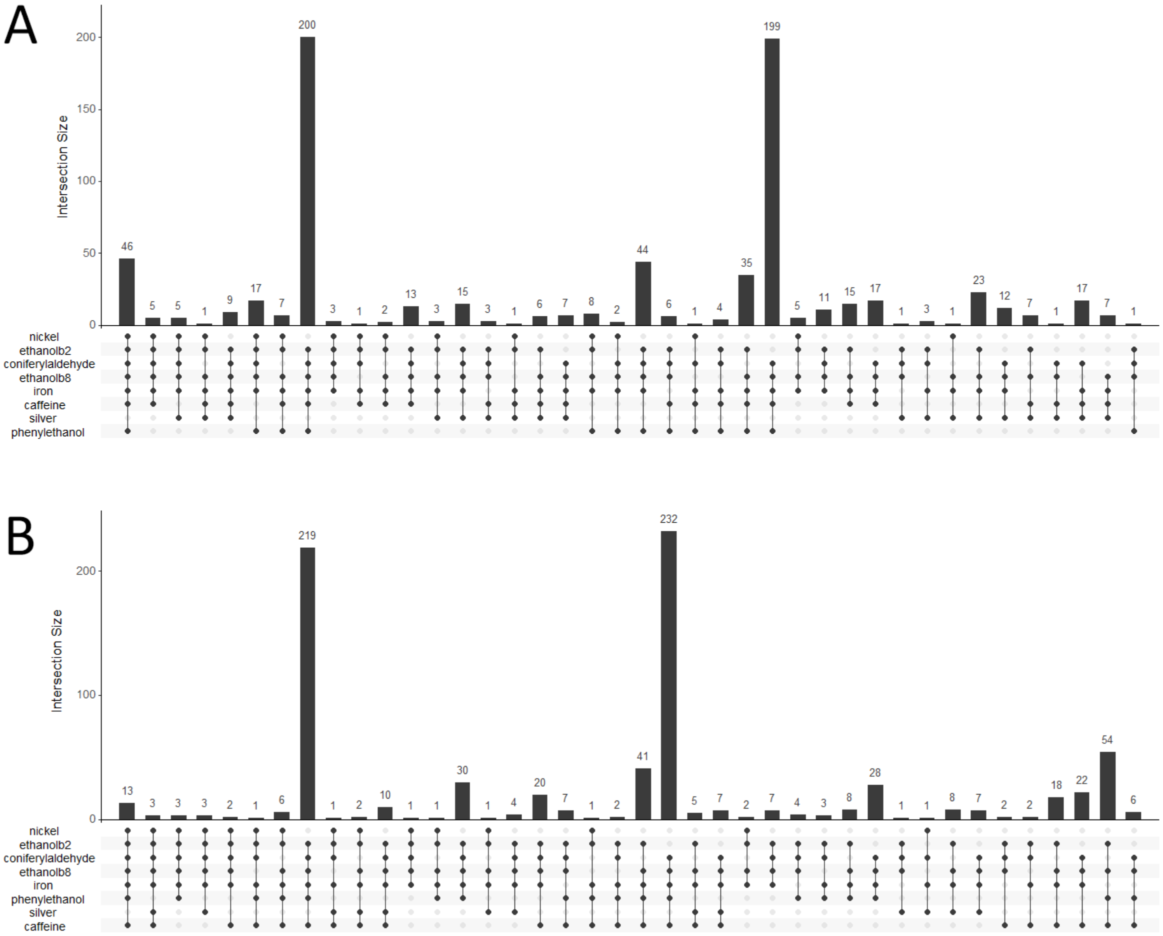


Figure 4.4. UpSet plots showing the number of common genes that are differentially expressed significantly for (A) upregulated genes, and (B) downregulated genes. Connected circles in the below represent intersections between experiment strains, and the number of genes are shown in the top bar graphs.

Total of 277 genes in ethanol-b2, 464 genes in ethanol-b8, 783 genes in caffeine, 374 genes in coniferyl aldehyde, 552 genes in iron, 28 genes in nickel, 759 genes in phenylethanol and 759 genes in silver-resistant strains were integrated into the metabolic model to generate strain-specific models (Figure 4.6). In comparison, 17.24% of genes in ethanol-b2, 15.74% of genes in ethanol-b8, 16.51% of genes in caffeine 16.50% of genes in coniferyl aldehyde, 16.01% of genes in iron, 18.42% of genes in nickel, 15.82% of genes in phenylethanol and 15.82% of genes in silver-resistant strains were found as annotated genes in the metabolic model, and they were integrated (Table 4.3).

Table 4.2. The common upregulated and downregulated genes within ethanol, caffeine, coniferyl aldehyde, iron, nickel and phenylethanol-resistant strains.

Upregulated Genes	**Downregulated Genes**
PPH3 IZH1 PEX21 RHO3 CDC123 TRM12 PGA3 EFM6 KAR1 AGA1 IZH2 RGS2 NHP6A	ACH1, AGP2, YBR285W, MAL32, ATG15, GUD1, PST1, ZRG8, RRT6, RMR1, MTL1, GPC1, RTS3, MAL13, MAL11, MAL12, FYV10, MET28, MHP1, PEP8, IML2, YJL132W, SUL2, RFX1, VPS34, YLR446W, YMR052C-A, STB2, ISF1, YMR103C, YPK2, PGM2, CVM1, YMR194C-A, ROY1, TDA1, YPT53, MET2, CIT1, RRI2, STI1, VPS17, GAC1, WTM1, MRL1, PIN3

Table 4.3. Numerical values of differentially expressed (DE) genes ($p < 0.05$) for mutant strains, and percentage integration into the reference metabolic model.

Strain Name	**# of DE genes**			**# of integrated DE genes**			**Integration Percentage**
	Up.	**Down.**	**Total**	**Up.**	**Down.**	**Total**	
b2-ethanol	764	842	1606	183	94	277	17.24
b8-ethanol	1439	1508	2947	303	161	464	15.74
caffeine	2174	2569	4743	349	434	783	16.51
coniferyl aldehyde	1146	1121	2267	237	137	374	16.50
iron	1855	1593	3448	377	175	552	16.01
nickel	109	43	152	24	4	28	18.42
phenylethanol	2426	2370	4796	364	395	759	15.82
silver	2370	2426	4796	395	364	759	15.82

In order to validate the expression data integration method, fluxes for each enzyme draw reaction (i.e., protein flux from protein pool to reaction for each protein) is predicted by flux balance analysis for all the strain models and for the wild-type model (wild-type here is referred as the metabolic model without expression data integrated). Flux ratios between wild-type and evolved models are calculated, and linear regression parameters are fitted after elimination of the outliers by the generalized extreme studentized deviate test to compare the flux-changes with the protein fold-changes from the expression data. Results confirm a good correlation between simulations and protein fold-change values (Figure 4.7).

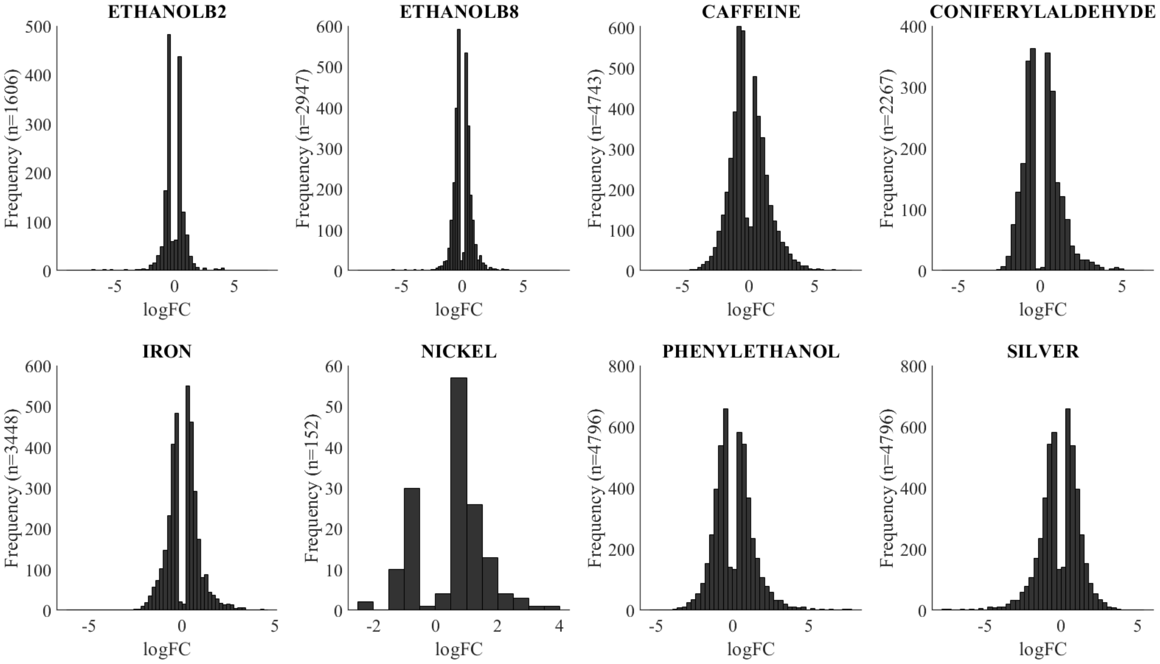


Figure 4.5. Frequencies of the differentially expressed genes after the gene expression analysis.

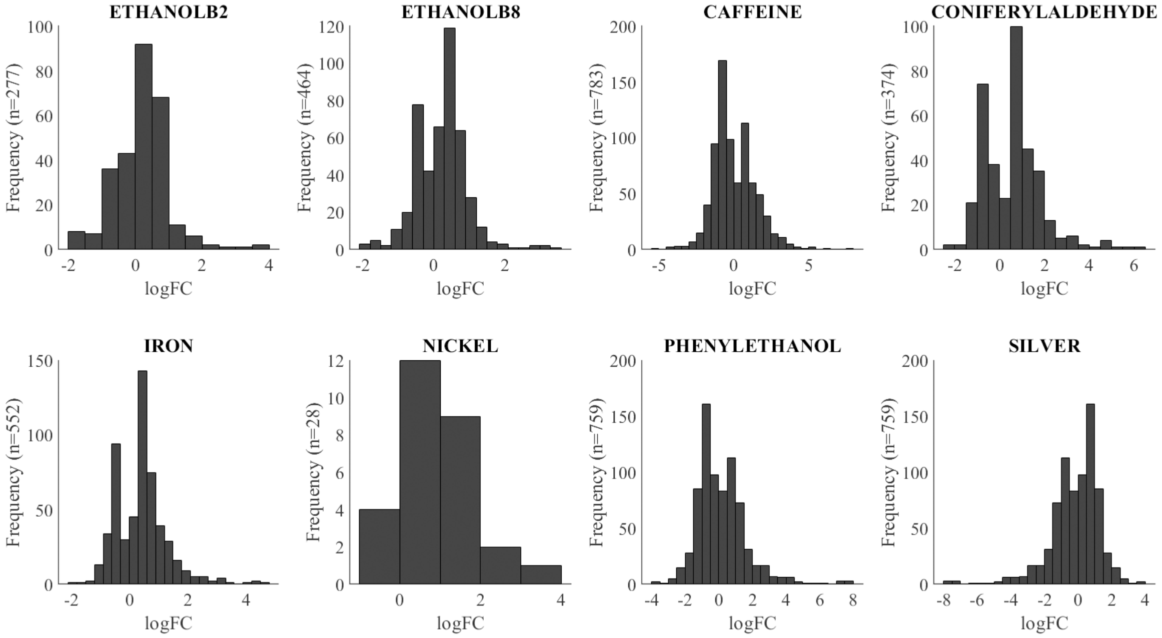


Figure 4.6. Frequencies of the differentially expressed genes that are integrated into the metabolic model.

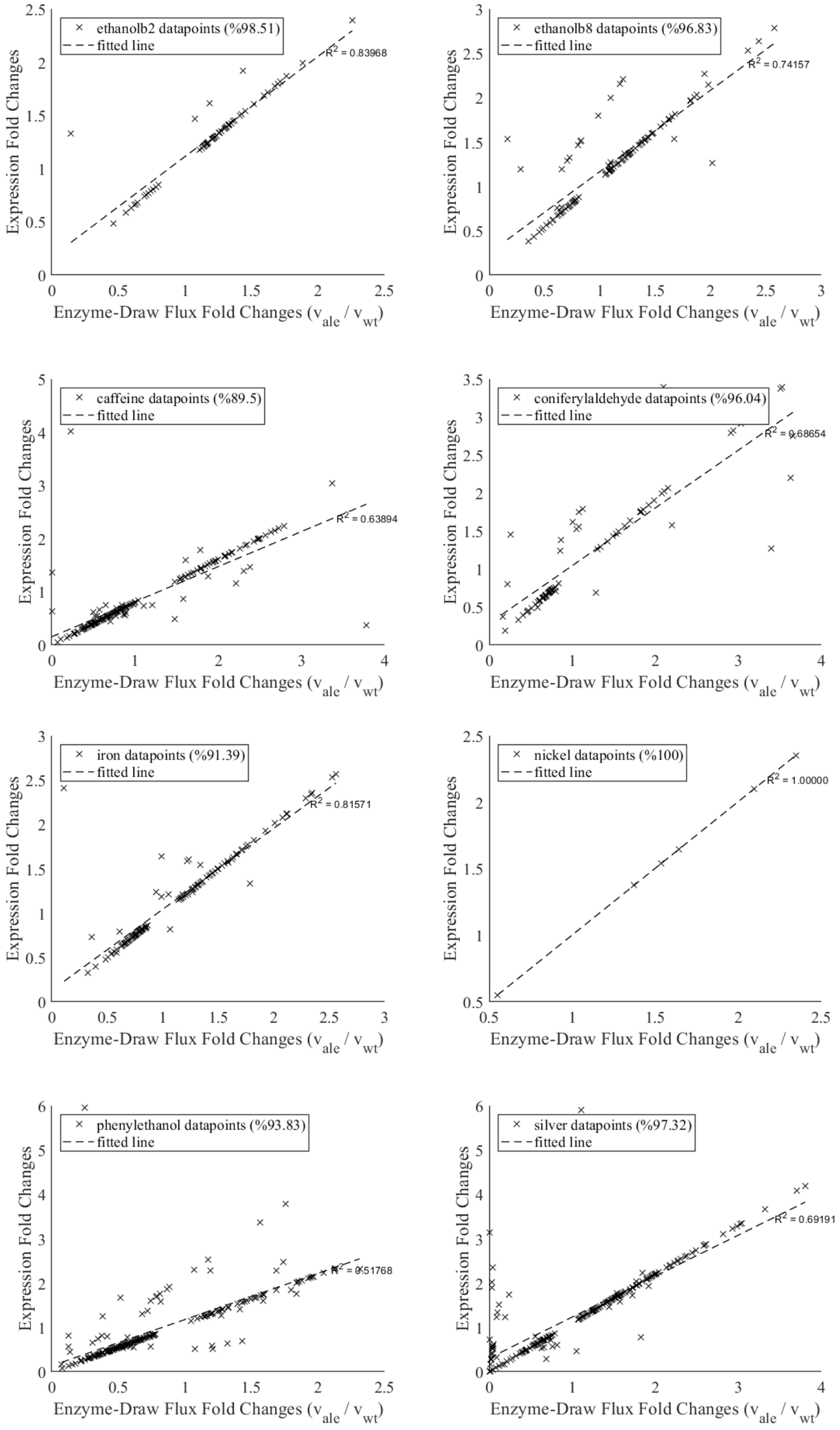


Figure 4.7. Linear regression analyses of the differential expression fold-changes obtained from microarray studies and the enzyme-draw flux fold-changes obtained from flux balance analysis simulations of corresponding proteins.

4.3. Flux Balance Analysis Simulations

The metabolic model is first simulated without the integration of the expression data (this model will be referred as wild-type model). Since the model is already enzymatically constrained, the only additional constraint required is the total amount of enzymes available in the enzyme pool. All the lower bounds for reactions were set to 0, and the upper bounds were set to 1000, except for the enzyme pool reaction. By doing that, the model was set free to be able to allocate enzymes as the reactions require. Flux rates of several exchange reactions as a function of growth rate are plotted in Figure 4.8.

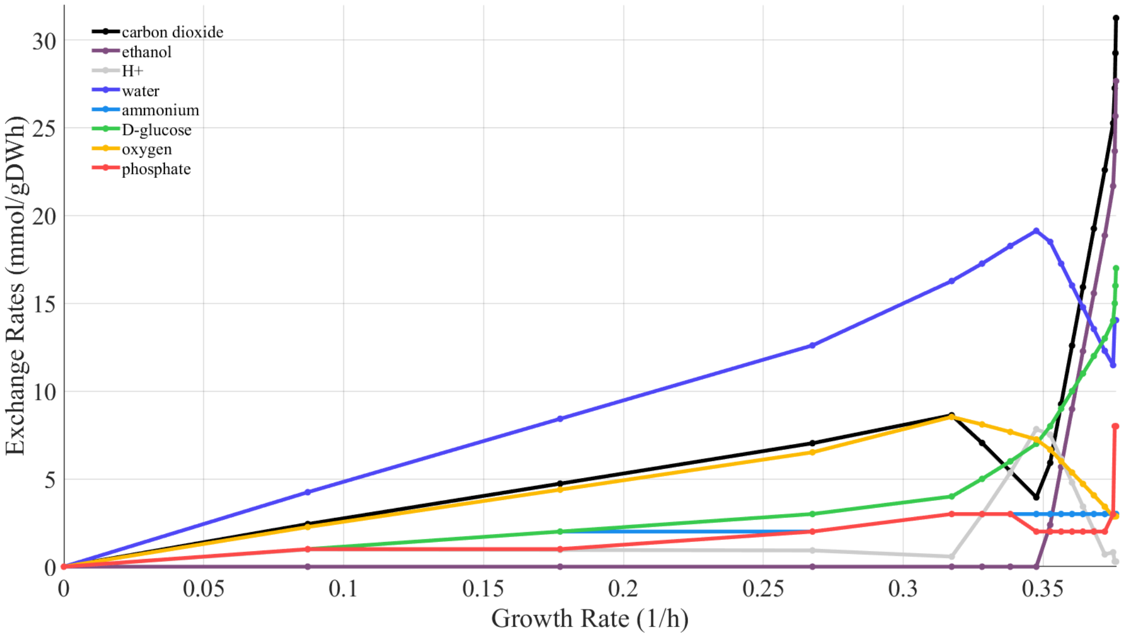


Figure 4.8. Metabolic model shows the overflow metabolism (crabtree effect) under fully aerobic conditions.

A behavioral change is observed around the time when growth rate exceeds 0.3 h^{-1}. Since the model is simulated under fully aerobic conditions (no constraints applied) and produces ethanol, this behavior can be explained by the Crabtree effect, where the yeast metabolism switches to perform both respiration and fermentation at the same time at a critical specific growth rate.

Robustness analysis is performed on the wild-type model for the growth rate with the varying levels of glucose uptake, oxygen uptake, acetate secretion and ethanol secretion rates (Figure 4.9). Unlike the traditional metabolic models, enzymatically constrained model showed non-linear graphs.

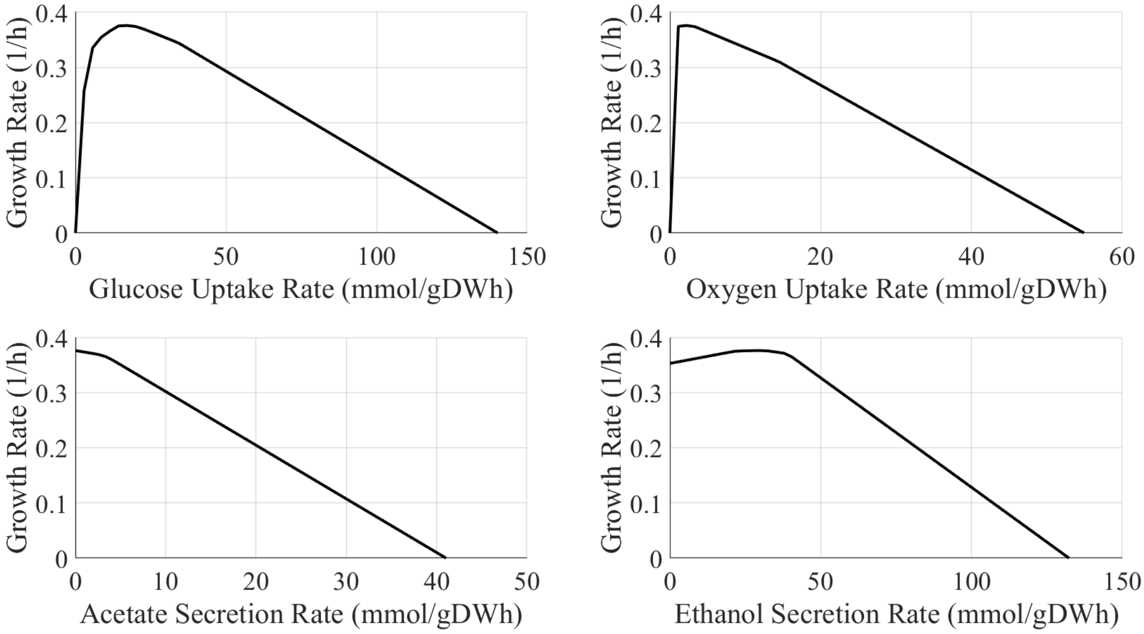


Figure 4.9. Robustness analyses on the glucose uptake, oxygen uptake, acetate secretion and ethanol secretion rates for the reference model.

After the validation of the wild-type model, all models for strains are simulated for the growth rates as a function of iteratively increasing glucose uptake rates (Figure 4.10). Models were able to grow up to the point where the protein accessibility becomes a limiting factor. It has been found that the evolution models except for the caffeine-resistant strain model show similar patterns on the points where the enzymes become limiting (break points in the lines).

Caffeine-resistant strain model shows different pattern in a way that it can consume more glucose to reach higher growth rates compared to others. Additionally, it must be noted that while only caffeine and coniferyl aldehyde-resistant strain models can grow at higher rates than wild-type model, ethanol, iron, nickel, phenylethanol and silver models can grow at lower maximum rates. The maximum growth rates and corresponding glucose uptake rates for each model are summarized in the Table 4.4.

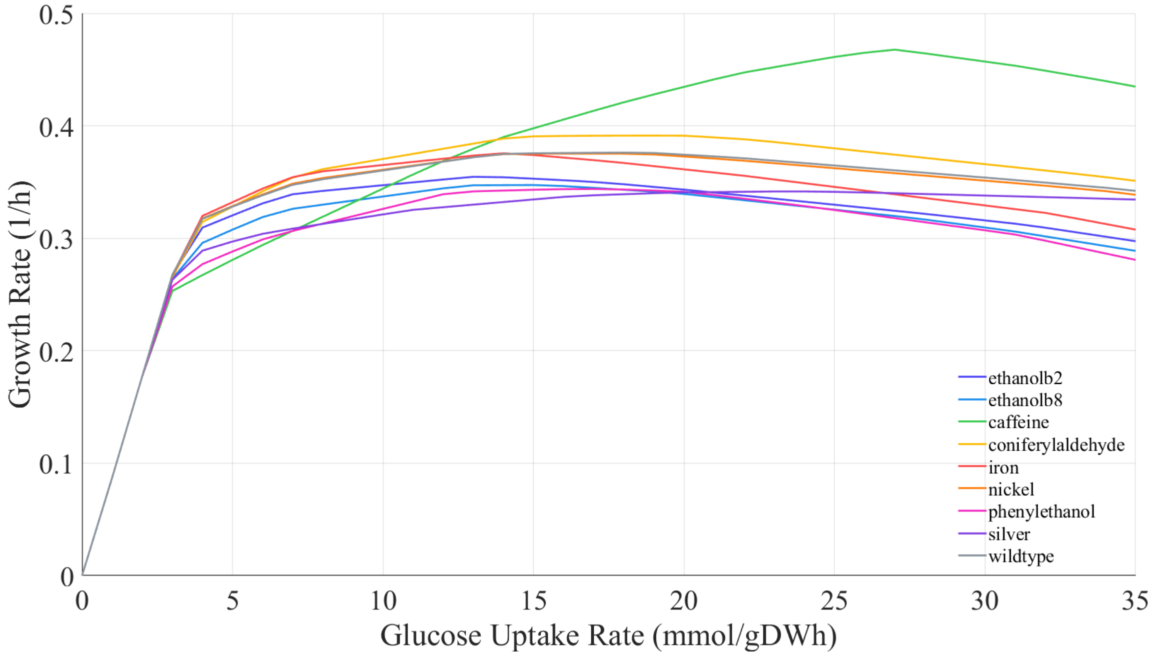


Figure 4.10. Growth rates as a function of iteratively increasing glucose uptake rates.

When the analysis results of the evolved strains were investigated within themselves, several novel behaviors are observed. Reactions that are only used in evolved strains individually and that are deactivated compared to reference strain is collected in Table 4.5. One reaction is found to be activated in all evolved strains while it is inactive in the reference strain was acetaldehyde uptake reaction. Since it is an exchange reaction which transports acetaldehyde between extracellular matrix and the cell, it does not have any gene annotations. Contrarily, the NDE2 isozyme draw reaction (to form NADH dehydrogenase) did not carry any flux in all evolved strains. Compared to the differential expression analysis, NDE2 was surprisingly upregulated and NDE1 was downregulated for almost all evolved strain, except the silver-resistant strain.

Table 4.4. Maximum growth rates (μ) from FBA for each strain and their maximum glucose uptake rates (GUR) required for that growth; and possible minimum and maximum glucose and oxygen uptake rates in mmol (gDW)$^{-1}$ h^{-1} from FVA for at least minimum of 90% of growth rate.

Strain Name	**Maximum μ**	**GUR (at maxμ)**	**GUR Range**	**O2 Range**
b2-ethanol	0.35466	14	5 - 34	2 - 12
b8-ethanol	0.34736	16	6 - 32	2 - 10
caffeine	0.46769	28	8 - 50	2 - 11
coniferyl aldehyde	0.39132	20	5 - 47	2 - 15
iron	0.37554	15	5 - 36	2 - 13
nickel	0.37523	18	5 - 46	2 - 16
phenylethanol	0.34382	18	8 - 31	2 - 7
silver	0.34163	25	7 - 50	2 - 16
reference	0.37618	19	5 - 47	2 - 16

Table 4.5. Number of reactions that are either activated or deactivated (carried flux or no flux) in the evolved strains compared to wildtype strain.

	# of Activated Rxns		**# of Dectivated Rxns**	
Strain	**Enzyme Draws**	**Total Number**	**Enzyme Draws**	**Total Number**
b2-ethanol	16	64	12	55
b8-ethanol	25	89	25	89
caffeine	30	131	70	172
coniferyl aldehyde	23	87	19	80
iron	27	102	26	107
nickel	6	17	6	14
phenylethanol	30	113	26	98
silver	31	114	33	116

4.3.1. Comparison of Exchange Fluxes Across ALE

Exchange reaction fluxes for uptake and secretion of the metabolites are compared across all evolved models. Objective function was the growth rate in all flux balance analysis simulations. Rates for the exchange reactions are collected while the glucose uptake rate is increased iteratively (Figure 4.11 and 4.12).

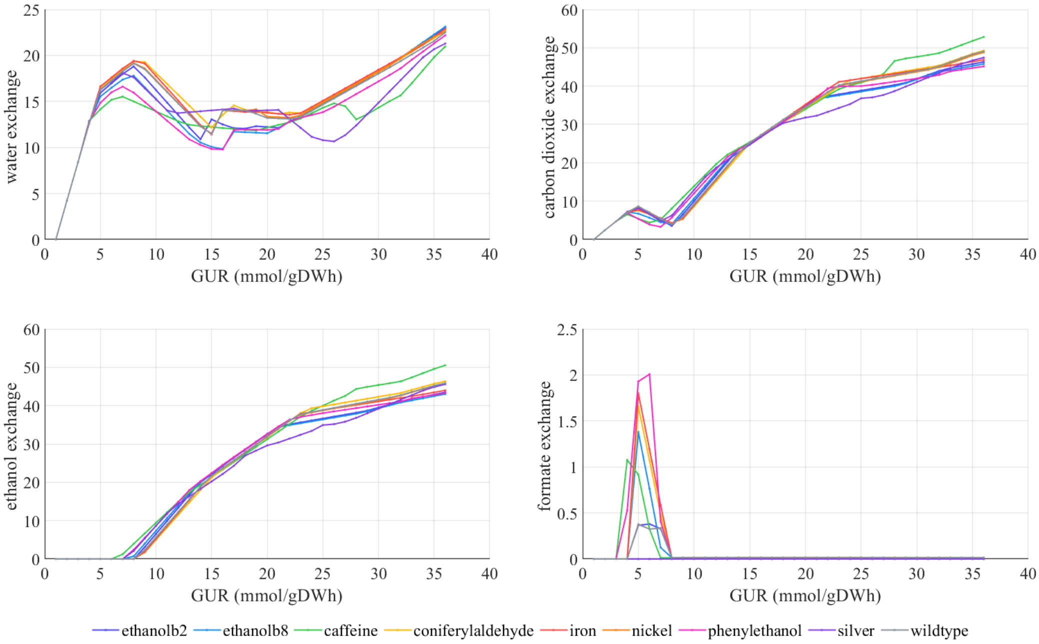


Figure 4.11. Exchange fluxes of water uptake, carbon dioxide, ethanol production and formate secretion as a function of iteratively increasing glucose uptake rates.

When the glucose uptake rate exceeds 20 mmol/gDWh^{-1}, models start to respirate as it can be understood from the water production plot. However, silver and caffeine-resistant models show a different behavioral pattern as either they start to respirate at higher glucose uptake rates (around 30 mmol/gDWh^{-1}) or they undergo a metabolic change that causes them to decrease water production. This behavior can be further investigated with the oxidative phosphorylation fluxes.

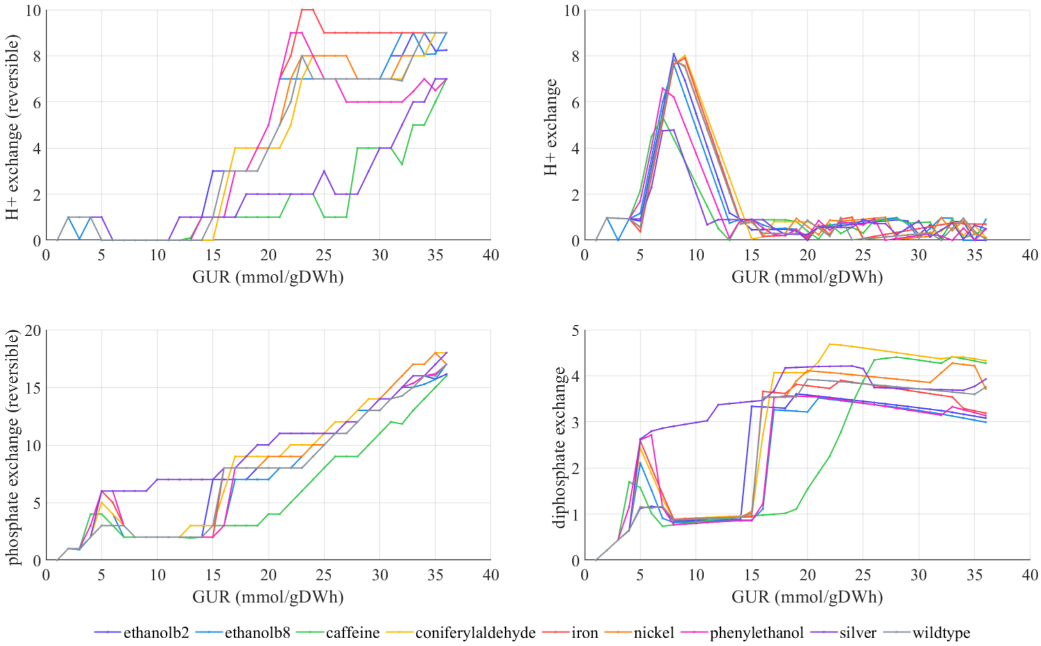


Figure 4.12. Exchange fluxes of hydrogen and phosphate uptakes (on left) and secretions (on right) as a function of iteratively increasing glucose uptake rates.

All exchange reactions that carry flux in all models are reported in Figure 4.13. One should remember that too small numeric values in metabolic fluxes (for example, L-lysine secretion reaction flux is calculated as 10^{-17} mmol/gDWh^{-1}) are biologically meaningless. Therefore, such reactions should not be taken into account when comparing evolved strains. Other insignificant results are the reactions that have the same flux value for both secretion and uptake fluxes, with net flux of 0 mmol/gDWh^{-1}. These results arise from the loops that occur in linear programming (see Methods section for details). Flux values of exchange reactions will be discussed in the next chapter.

Similar to the previous results of the iterative simulations on increasing glucose uptake rates (Figure 4.10), the most distinguished results were collected from the caffeine and silver-resistant strains' simulations. While silver-resistant strain is found to be able to secrete higher amounts of acetate and ethanol, its growth rate shorts fall compared to other strains. caffeine-resistant strain on the other hand, uses its resources to grow at higher rates and, therefore, its secretion rates are almost equal to average of other strains including the reference strain.

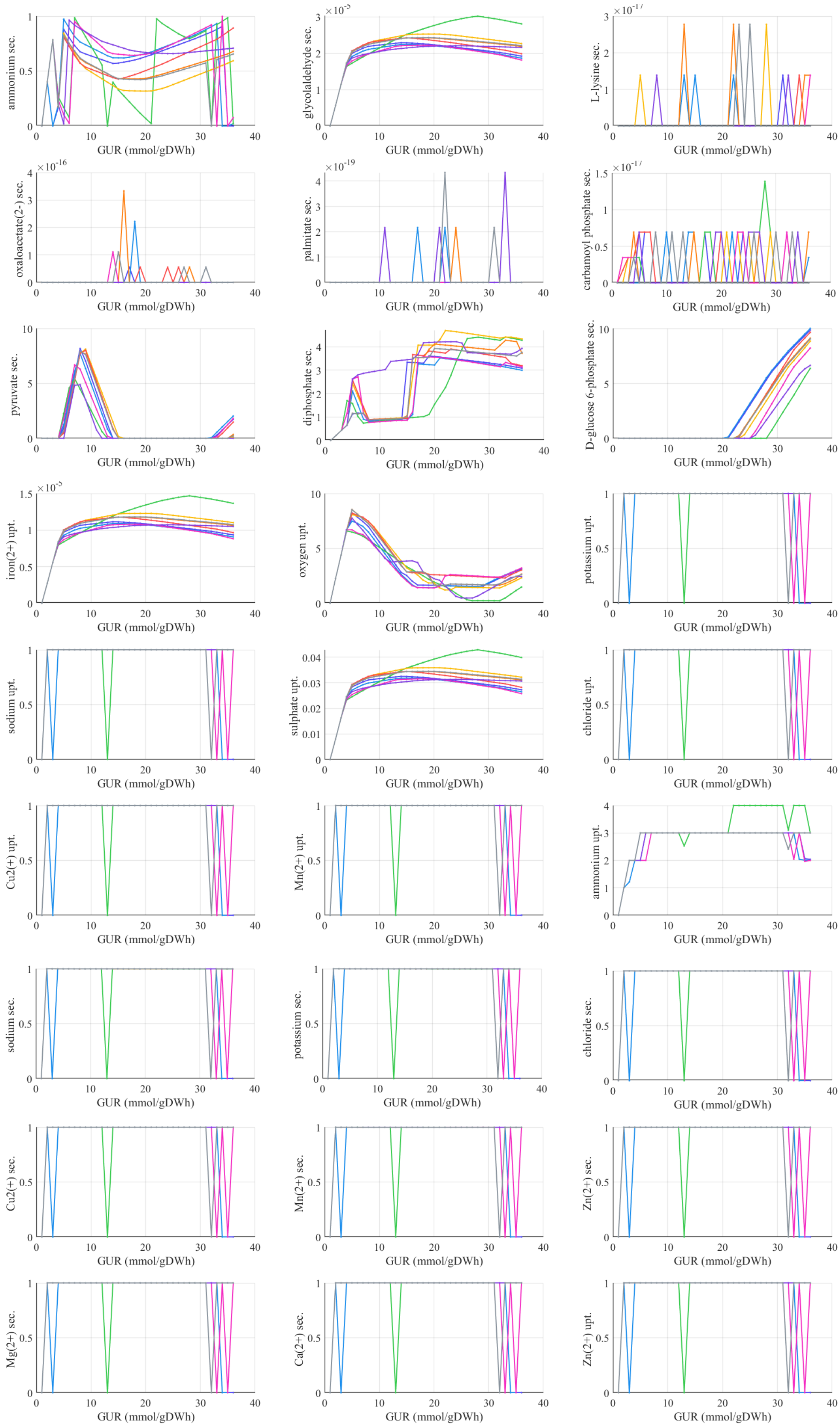


Figure 4.13. Exchange values in mmol/gDWh^{-1} for all exchange reactions that carry flux in ALE simulations.

4.3.2. Robustness Analysis

Robustness analysis is performed on all models for the growth rate with the varying levels of glucose uptake, oxygen uptake, acetate secretion and ethanol secretion rates. Each robustness analysis in Figure 4.14 has two plots with zoomed sections shown in gray areas.

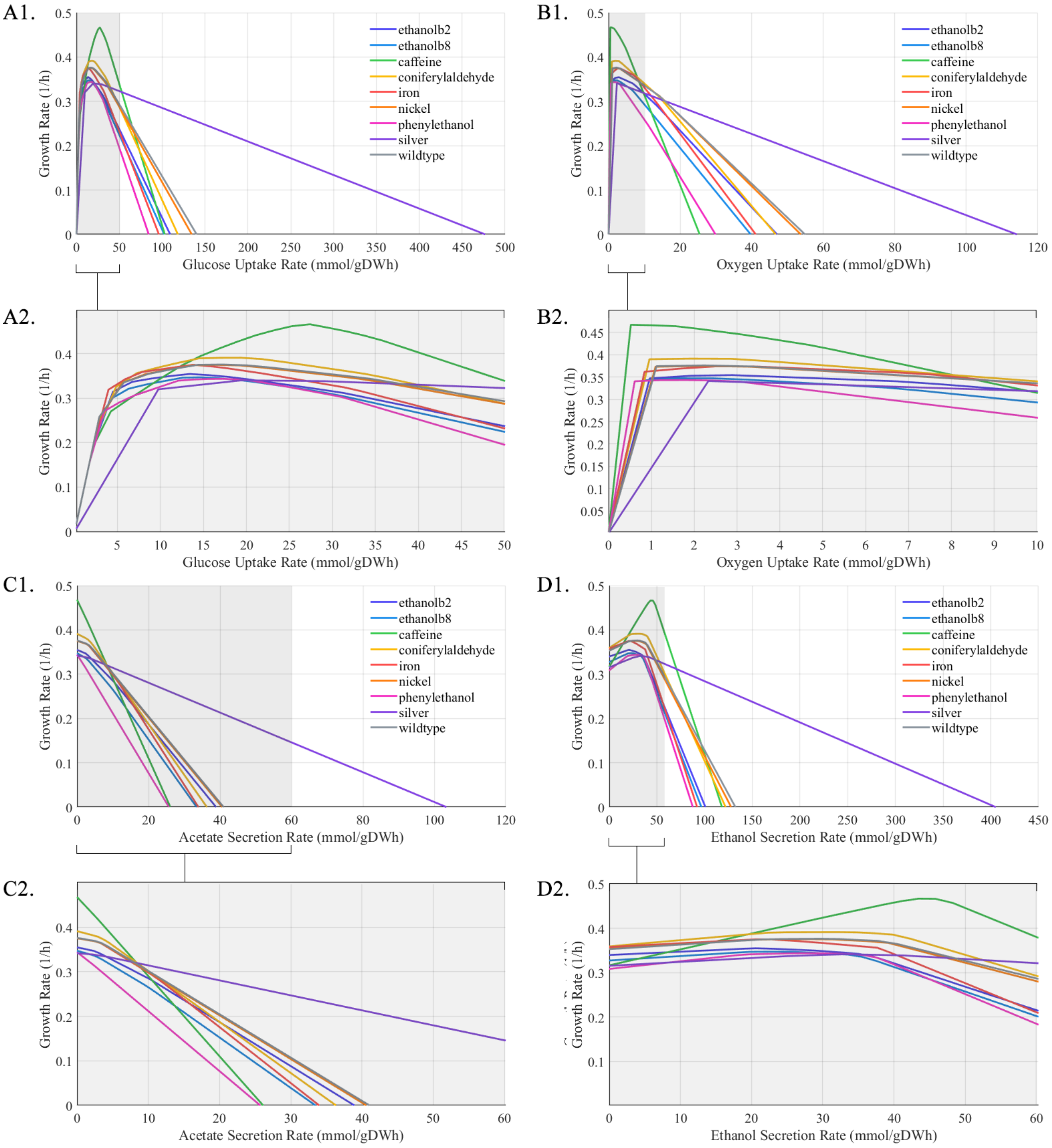


Figure 4.14. Robustness analyses on the A) glucose uptake, B) oxygen uptake, C) acetate secretion, and D) ethanol secretion rates for all models.

4.3.3. Phenotype Phase Planes

Phenotype phase planes are obtained through double robustness analysis to see the effect of oxygen on growth while the systems are fed with the gradually increasing glucose levels. Contour plots showing different regions can be seen in the Figure 4.15. Minimum and maximum values of oxygen uptake rates (y-axis) and glucose uptake rates (x-axis) are kept same in all 8 plots to be able to observe growth capacity of different resistant models clearly. Since the total area of the plots are the same (x-axis and y-axis ranges), available growth of models can be compared. On the low glucose uptake rates, growth rate contours have steep slopes for all models. They reach the growth rate of 0.3 h^{-1} when the glucose uptake rates are varying between 5-10 mmol/gDWh. Despite having the highest possible growth rate at 0.46497 h^{-1}, caffeine-resistant model has the smallest area to produce biomass. It cannot grow if the oxygen uptake rate is higher than 25 mmol/gDWh, making it sensitive hyperoxic conditions. silver-resistant model on the other hand is able to grow at high rates of both oxygen and glucose uptake rates. Ranges for the glucose and oxygen uptake rates can be found in the Table 4.6

Table 4.6. Glucose and oxygen uptake rate ranges (mmol/gDWh) at the point where growth rate is 0.3 h^{-1} for all strains.

ALE Experiment	**GUR range**	**GUR difference**	**O2 range**	**O2 difference**
b2-ethanol	5 - 34	29	2 - 12	10
b8-ethanol	6 - 32	26	2 - 10	8
caffeine	8 - 50	42	2 - 11	9
coniferyl aldehyde	5 - 47	42	2 - 15	13
iron	5 - 36	31	2 - 13	11
nickel	5 - 46	41	2 - 16	14
phenylethanol	8 - 31	23	2 - 7	5
silver	7 - 50	43	2 - 16	14
reference	5 - 47	42	2 - 16	14

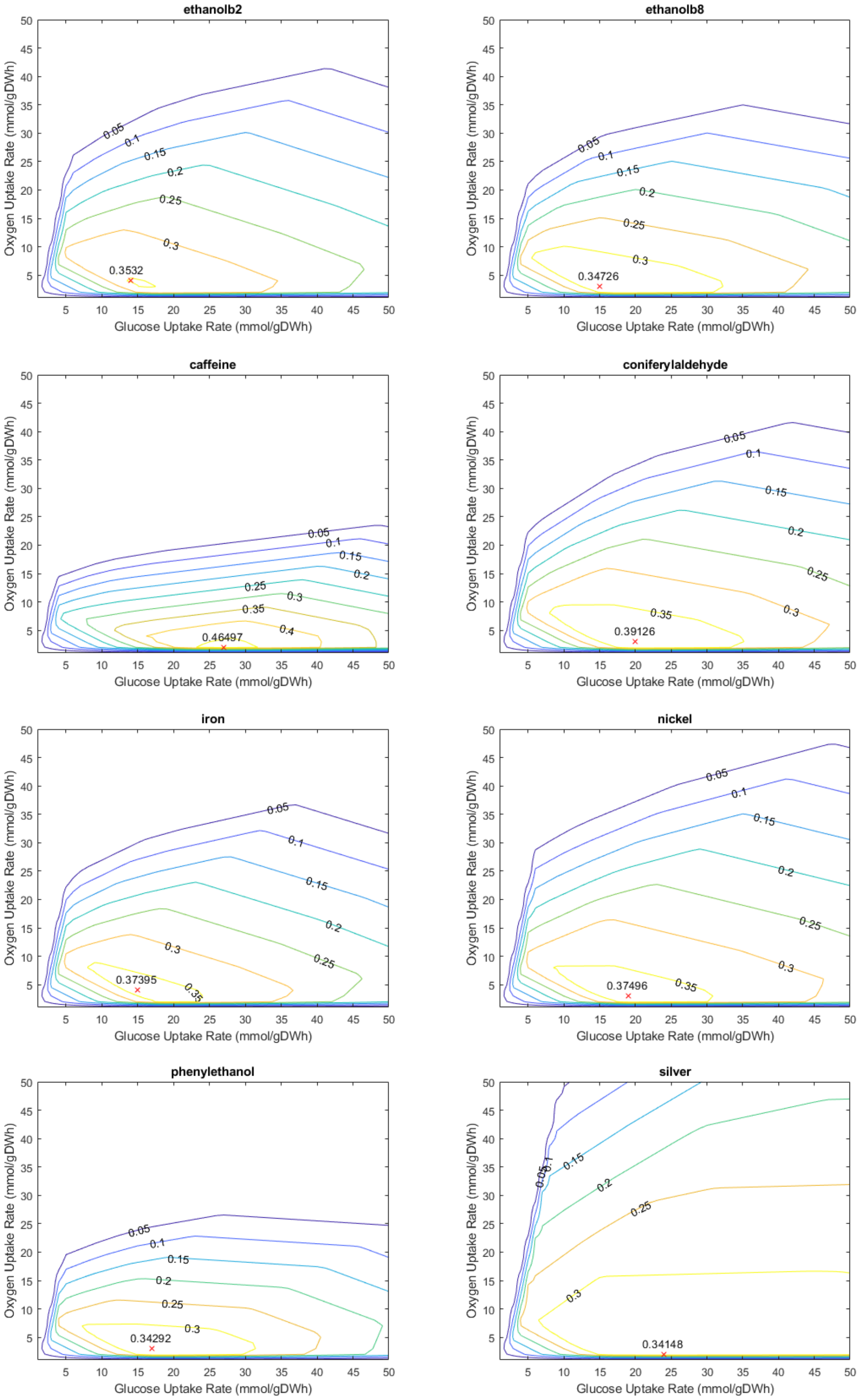


Figure 4.15. The glucose uptake rate (mmol/gDWh) versus oxygen uptake rate (mmol/gDWh) phenotype phase planes show the cellular growth rate as colored contours for each adapted model simulations. Maximum points are shown as red cross.

4.3.4. Cumulative Fluxes

In order to characterize the differentiating reactions in the flux balance analysis results, cumulative flux vectors of each experiment are plotted (Figure 4.16). It must be remembered that at this point, the only constraint applied to the models is the level of enzymes available in the enzyme pool. In other words, the models are able to allocate enzymes as the reactions require while all other sources (carbon, nitrogen, etc.) are unlimited. Standard deviations for each reaction between strains are calculated to find most diverging points and plotted as a bar graph.

Reactions with the standard deviation values more than 10 mmol/gDWh are listed in Table 4.7. The same reactions with different numbers in their names use different enzymes. For example, glyceraldehyde-3-phosphate dehydrogenase (GADPH) reaction can be carried out in the presence of YGR192C or YJL052W or YJR009C genes' proteins, in other words, TDH1 or TDH2 or TDH3 isozymes. Therefore, there are three versions of the same reaction (specified with the numbers in pharantheses), using different enzymes. It must be noted that these multiple reactions carried out with isozymes in the metabolic model may not mean a significant difference if the enzymes do not catalyze any other reaction. This means that the chosen isozyme may not matter mathematically if the flux value across the isozyme reactions are the same.

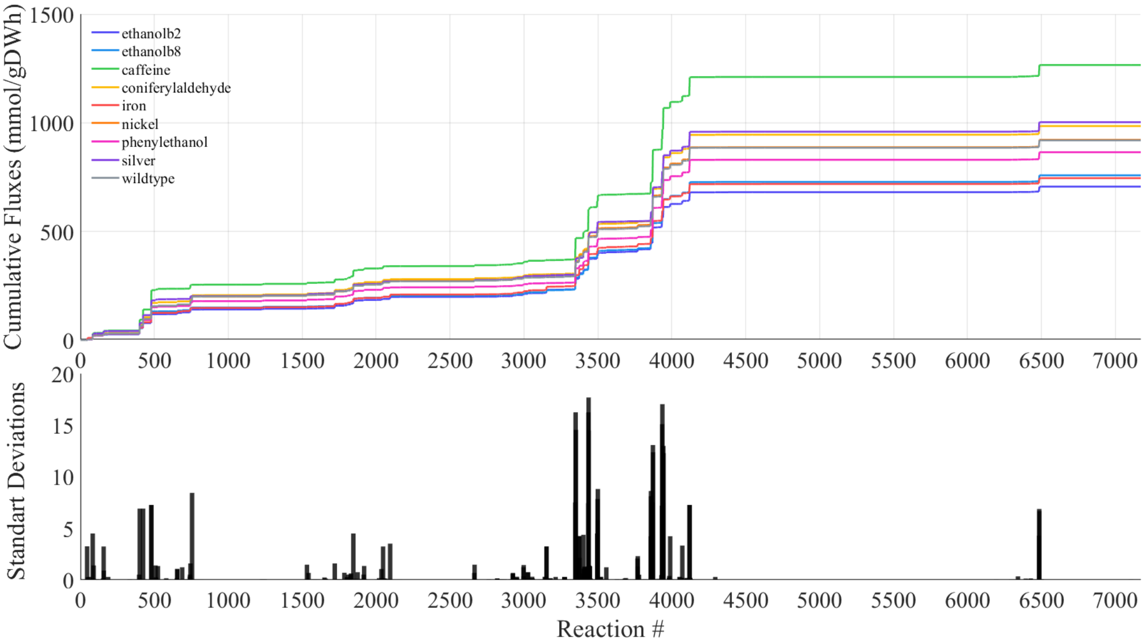


Figure 4.16. Cumulative flux vectors of each model simulations under unlimited glucose uptake constraints and the standard deviations of reactions across each adaptation.

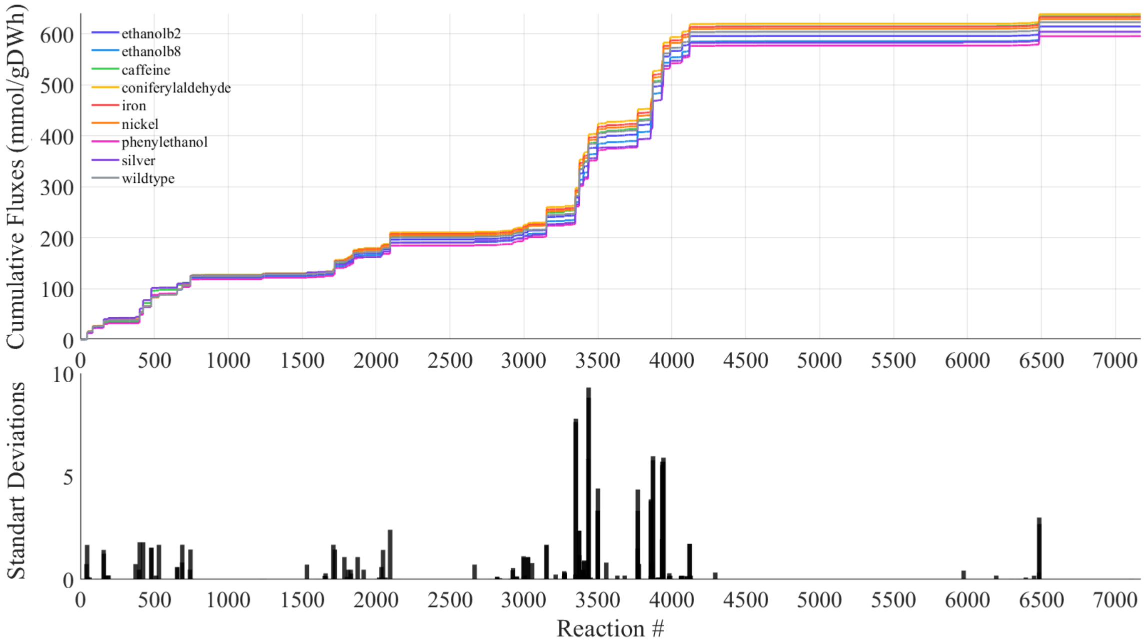


Figure 4.17. Cumulative flux vectors of each simulation and standard deviations. Glucose uptake rate is constrained to 10 mmol/gDWh for all models.

Because of the reason behind the differentiation, i.e., it is mostly caused by different uptake values of glucose, cumulative flux vectors of each experiment are plotted once more when the boundaries of the glucose uptake reaction is constrained to 10 $gDWh^{-1}$ for each model (Figure 4.17). This has done to eliminate the differences caused by the free glucose uptake rates (as shown, each model can uptake different maximum amount of glucose) and therefore affecting the downstream reaction fluxes. Reactions with the standard deviation values more than 5 mmol/gDWh are listed in Table 4.7. Despite the decrease in the flux values (caused by the lower uptake rates of glucose), the top divergent reactions obtained are in agreement with the no-constraint FBA results.

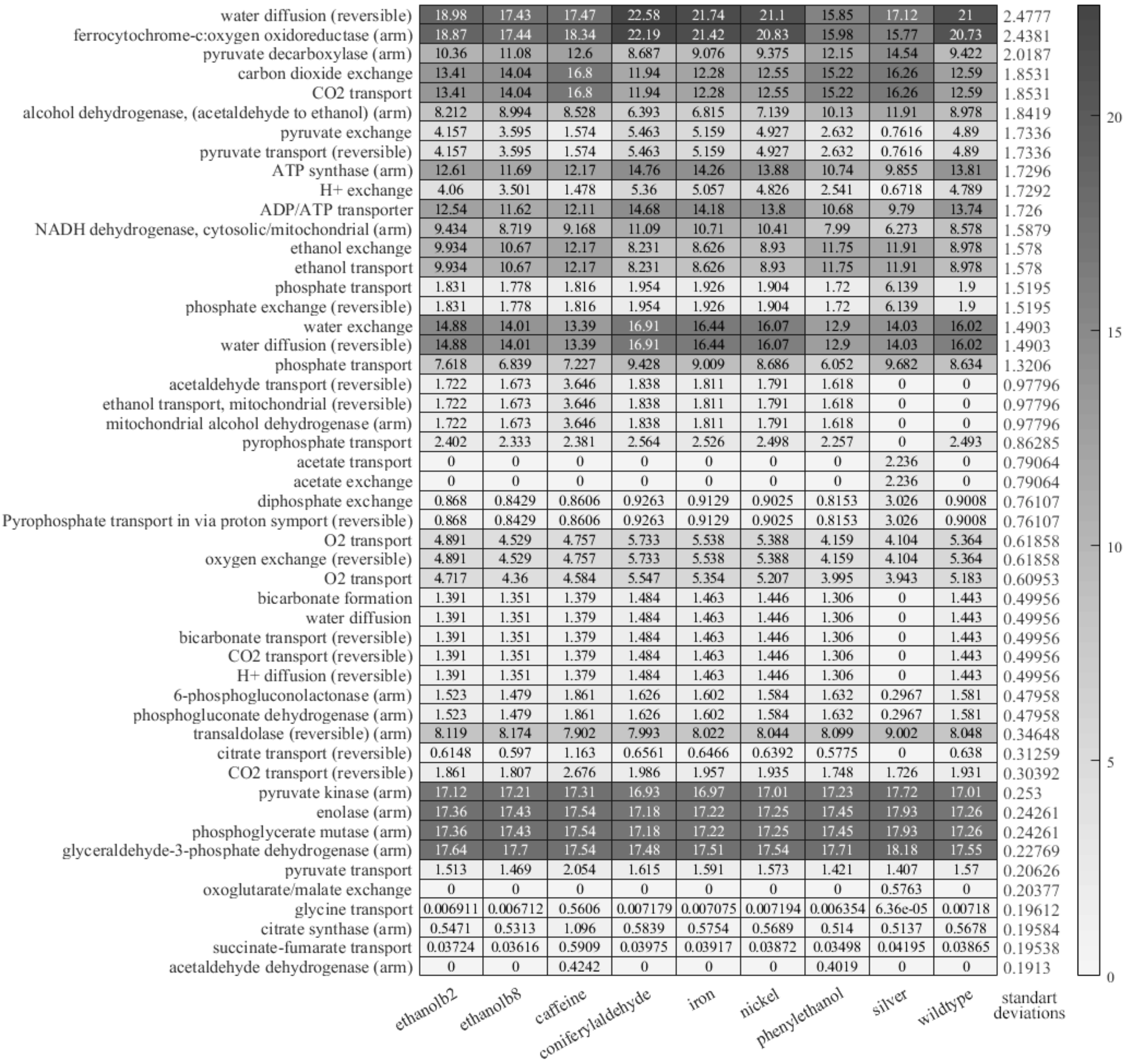
Table 4.7. The most divergent reactions according to their standard deviations across all strains and their flux values in mmol/gDWh.

#	**Reaction Name**	**Gene**	**eth-b2**	**eth-b8**	**caffeine**	**con. ald.**	**iron**	**nickel**	**phen.**	**silver**	**ref.**	**std**
1	glyceraldehyde-3-phosphate dehydrogenase (2)	TDH1	24.13	0	0	36.82	0	32.79	0	37.4	32.87	17.69
2	pyruvate decarboxylase (3)	PDC5	0	0	44.31	0	22.11	0	30.08	0	0	17.04
3	glyceraldehyde-3-phosphate dehydrogenase (1)	TDH3	0	0	48.77	0	0	0	0	0	0	16.26
4	enolase (2)	ENO1	23.84	0	48.77	36.5	25.04	32.49	32.46	0	32.56	16.25
5	pyruvate decarboxylase (2)	PDC1	20.94	25.19	0	33.69	0	29.42	0	35.33	29.49	15.1
6	enolase (4)	ERR3	0	28.03	0	0	0	0	0	37.13	0	14.55
7	glyceraldehyde-3-phosphate dehydrogenase (3)	TDH2	0	28.31	0	0	25.34	0	32.74	0	0	14.52
8	phosphoglycerate mutase (1)	YOR283W	23.84	28.03	48.77	36.5	25.04	32.49	32.46	0	32.56	13.08
9	pyruvate kinase (1)	CDC19	23.6	27.79	48.45	36.24	24.79	32.23	32.23	0	32.31	12.99
10	phosphoglycerate mutase (2)	GPM1	0	0	0	0	0	0	0	37.13	0	12.38
11	pyruvate kinase (2)	PYK2	0	0	0	0	0	0	0	36.9	0	12.3

Table 4.8. The most divergent reactions across all strains and their flux values in mmol/gDWh when the glucose uptake rate is constrained to 10 mmol/gDWh.

#	**Reaction Name**	Gene	**eth-b2**	**eth-b8**	**caffeine**	**con. ald.**	**iron**	**nickel**	**phen.**	**silver**	**ref.**	**std**
1	glyceraldehyde-3-phosphate dehydrogenase (2)	TDH1	17.64	0	0	17.48	0	17.54	0	18.18	17.55	9.32
2	glyceraldehyde-3-phosphate dehydrogenase (3)	TDH2	0	17.7	0	0	17.51	0	17.71	0	0	8.82
4	enolase (4)	ERR3	0	17.43	0	0	0	0	0	17.93	0	7.8
5	enolase (2)	ENO1	17.36	0	17.54	17.18	17.22	17.25	17.45	0	17.26	7.64
6	phosphoglycerate mutase (2)	GPM1	0	0	0	0	0	0	0	17.93	0	5.98
8	pyruvate kinase (2)	PYK2	0	0	0	0	0	0	0	17.72	0	5.91
9	glyceraldehyde-3-phosphate dehydrogenase (1)	TDH3	0	0	17.54	0	0	0	0	0	0	5.85
10	phosphoglycerate mutase (1)	YOR283W	17.36	17.43	17.54	17.18	17.22	17.25	17.45	0	17.26	5.78
11	pyruvate decarboxylase (3)	PDC5	0	0	12.6	0	9.08	0	12.15	0	0	5.72
12	pyruvate kinase (1)	CDC19	17.12	17.21	17.31	16.93	16.97	17.01	17.23	0	17.01	5.7
13	pyruvate decarboxylase (2)	PDC1	10.36	11.08	0	8.69	0	9.37	0	14.54	9.42	5.55

Finally, only the arm reactions (where the isozyme reactions were separated from, see Methods) where the glucose uptake rate is constrained to 10 mmol/gDWh are plotted as heatmap in Figure 4.18. Arm reactions in this context refer to the combined reactions that normally have multiple sub-reactions because of isozymes catalyzing the same reaction. As it can be seen from the previous tables, most of the differences were caused by the isozymes with the flux values close to each other. By considering only arm reactions and a fixed glucose uptake rate, the reactions that actually differ because of the expression integration are found.



	ethanolb2	ethanolb8	caffeine	coniferylaldehyde	iron	nickel	phenylethanol	silver	wildtype	standart deviations
water diffusion (reversible)	18.98	17.43	17.47	22.58	21.74	21.1	15.85	17.12	21	2.4777
ferrocytochrome-c:oxygen oxidoreductase (arm)	18.87	17.44	18.34	22.19	21.42	20.83	15.98	15.77	20.73	2.4381
pyruvate decarboxylase (arm)	10.36	11.08	12.6	8.687	9.076	9.375	12.15	14.54	9.422	2.0187
carbon dioxide exchange	13.41	14.04	16.8	11.94	12.28	12.55	15.22	16.26	12.59	1.8531
CO2 transport	13.41	14.04	16.8	11.94	12.28	12.55	15.22	16.26	12.59	1.8531
alcohol dehydrogenase, (acetaldehyde to ethanol) (arm)	8.212	8.994	8.528	6.393	6.815	7.139	10.13	11.91	8.978	1.8419
pyruvate exchange	4.157	3.595	1.574	5.463	5.159	4.927	2.632	0.7616	4.89	1.7336
pyruvate transport (reversible)	4.157	3.595	1.574	5.463	5.159	4.927	2.632	0.7616	4.89	1.7336
ATP synthase (arm)	12.61	11.69	12.17	14.76	14.26	13.88	10.74	9.855	13.81	1.7296
H+ exchange	4.06	3.501	1.478	5.36	5.057	4.826	2.541	0.6718	4.789	1.7292
ADP/ATP transporter	12.54	11.62	12.11	14.68	14.18	13.8	10.68	9.79	13.74	1.726
NADH dehydrogenase, cytosolic/mitochondrial (arm)	9.434	8.719	9.168	11.09	10.71	10.41	7.99	6.273	8.578	1.5879
ethanol exchange	9.934	10.67	12.17	8.231	8.626	8.93	11.75	11.91	8.978	1.578
ethanol transport	9.934	10.67	12.17	8.231	8.626	8.93	11.75	11.91	8.978	1.578
phosphate transport	1.831	1.778	1.816	1.954	1.926	1.904	1.72	6.139	1.9	1.5195
phosphate exchange (reversible)	1.831	1.778	1.816	1.954	1.926	1.904	1.72	6.139	1.9	1.5195
water exchange	14.88	14.01	13.39	16.91	16.44	16.07	12.9	14.03	16.02	1.4903
water diffusion (reversible)	14.88	14.01	13.39	16.91	16.44	16.07	12.9	14.03	16.02	1.4903
phosphate transport	7.618	6.839	7.227	9.428	9.009	8.686	6.052	9.682	8.634	1.3206
acetaldehyde transport (reversible)	1.722	1.673	3.646	1.838	1.811	1.791	1.618	0	0	0.97796
ethanol transport, mitochondrial (reversible)	1.722	1.673	3.646	1.838	1.811	1.791	1.618	0	0	0.97796
mitochondrial alcohol dehydrogenase (arm)	1.722	1.673	3.646	1.838	1.811	1.791	1.618	0	0	0.97796
pyrophosphate transport	2.402	2.333	2.381	2.564	2.526	2.498	2.257	0	2.493	0.86285
acetate transport	0	0	0	0	0	0	0	2.236	0	0.79064
acetate exchange	0	0	0	0	0	0	0	2.236	0	0.79064
diphosphate exchange	0.868	0.8429	0.8606	0.9263	0.9129	0.9025	0.8153	3.026	0.9008	0.76107
Pyrophosphate transport in via proton symport (reversible)	0.868	0.8429	0.8606	0.9263	0.9129	0.9025	0.8153	3.026	0.9008	0.76107
O2 transport	4.891	4.529	4.757	5.733	5.538	5.388	4.159	4.104	5.364	0.61858
oxygen exchange (reversible)	4.891	4.529	4.757	5.733	5.538	5.388	4.159	4.104	5.364	0.61858
O2 transport	4.717	4.36	4.584	5.547	5.354	5.207	3.995	3.943	5.183	0.60953
bicarbonate formation	1.391	1.351	1.379	1.484	1.463	1.446	1.306	0	1.443	0.49956
water diffusion	1.391	1.351	1.379	1.484	1.463	1.446	1.306	0	1.443	0.49956
bicarbonate transport (reversible)	1.391	1.351	1.379	1.484	1.463	1.446	1.306	0	1.443	0.49956
CO2 transport (reversible)	1.391	1.351	1.379	1.484	1.463	1.446	1.306	0	1.443	0.49956
H+ diffusion (reversible)	1.391	1.351	1.379	1.484	1.463	1.446	1.306	0	1.443	0.49956
6-phosphogluconolactonase (arm)	1.523	1.479	1.861	1.626	1.602	1.584	1.632	0.2967	1.581	0.47958
phosphogluconate dehydrogenase (arm)	1.523	1.479	1.861	1.626	1.602	1.584	1.632	0.2967	1.581	0.47958
transaldolase (reversible) (arm)	8.119	8.174	7.902	7.993	8.022	8.044	8.099	9.002	8.048	0.34648
citrate transport (reversible)	0.6148	0.597	1.163	0.6561	0.6466	0.6392	0.5775	0	0.638	0.31259
CO2 transport (reversible)	1.861	1.807	2.676	1.986	1.957	1.935	1.748	1.726	1.931	0.30392
pyruvate kinase (arm)	17.12	17.21	17.31	16.93	16.97	17.01	17.23	17.72	17.01	0.253
enolase (arm)	17.36	17.43	17.54	17.18	17.22	17.25	17.45	17.93	17.26	0.24261
phosphoglycerate mutase (arm)	17.36	17.43	17.54	17.18	17.22	17.25	17.45	17.93	17.26	0.24261
glyceraldehyde-3-phosphate dehydrogenase (arm)	17.64	17.7	17.54	17.48	17.51	17.54	17.71	18.18	17.55	0.22769
pyruvate transport	1.513	1.469	2.054	1.615	1.591	1.573	1.421	1.407	1.57	0.20626
oxoglutarate/malate exchange	0	0	0	0	0	0	0	0.5763	0	0.20377
glycine transport	0.006911	0.006712	0.5606	0.007179	0.007075	0.007194	0.006354	6.36e-05	0.00718	0.19612
citrate synthase (arm)	0.5471	0.5313	1.096	0.5839	0.5754	0.5689	0.514	0.5137	0.5678	0.19584
succinate-fumarate transport	0.03724	0.03616	0.5909	0.03975	0.03917	0.03872	0.03498	0.04195	0.03865	0.19538
acetaldehyde dehydrogenase (arm)	0	0	0.4242	0	0	0	0.4019	0	0	0.1913

Figure 4.18. Heatmap of the reaction fluxes where the standard deviation between models are the highest (Top 50).

Overall, the most divergent strain can be seen as silver-resistant strain while the nickel strain has the closest values to the reference model as expected. The highest deviation caused in the water diffusion from mitochondria to cytoplasm indicates that even at the same glucose feeding rates, phenylethanol-resistant strain requires higher (89.5%) water in the mitochondria. Followed with the ferrocytochrome-c:oxygen oxidoreductase, pyruvate decarboxylase reactions and different exchange reactions (such as CO2, H+), the results conclude that the models of evolved strains follow different paths to produce biomass with the same environmental conditions provided.

4.3.5. Hierarchical Relationships

As these different models belong to evolved strains, hierarchical relationship between models are drawn as dendrogram in Figure 4.19. Euclidean pairwise distance is calculated in MATLAB between flux balance analysis solution vectors (fluxes) for each model where the glucose uptake rate is constrained to 10 mmol/gDWh first and then clustered by the unweighted average distance (UPGMA) method to draw a tree. Cophenetic correlation coefficient is found as 0.9744 while the Spearman's rank correlation between the dissimilarities and the cophenetic distances is calculated as 0.9096, indicating a high-quality solution.

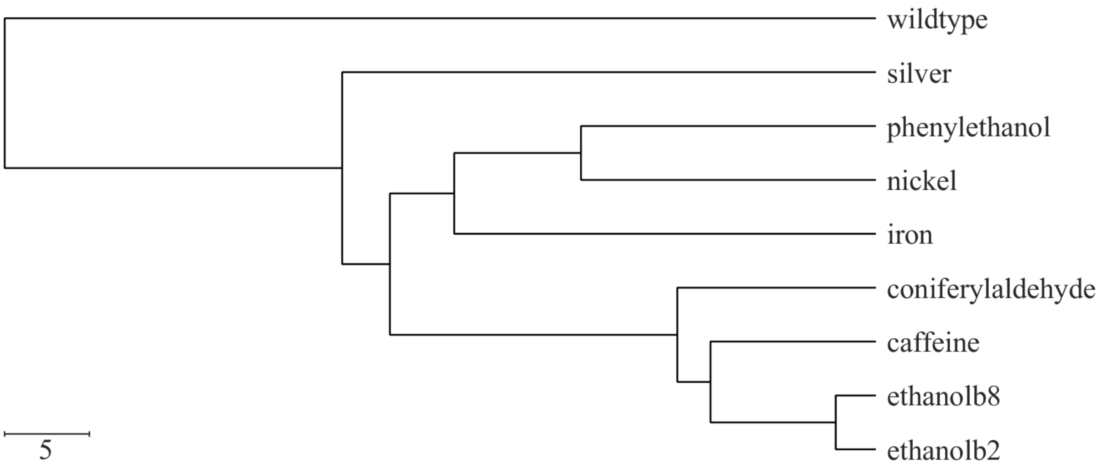


Figure 4.19. Hierarchical relationships between the strains drawn as a dendrogram plot. Euclidian distances are calculated from the reaction fluxes (flux balance analysis) where the glucose uptake rate is constrained to 10 mmol/gDWh.

4.3.6. Gene Regulation

Despite the high correlation between the fold changes obtained from differential expression analysis and the flux balance analyses for particular enzymes, outliers must be investigated for the indirect regulation of enzymes. Three groups can be formed for these: Firstly, enzymes may show transcriptional regulation behaviors (TR) if their flux fold changes (between evolved and reference strain models) correlate with the differential gene expression results. Secondly, enzymes may show metabolic regulation behaviors (MR) if the flux change is observed even in the absences of the expressional changes. And lastly, enzymes may show post-transcriptional regulation behaviors (PR) if their expression levels change without showing any flux changes (Table 4.9). One should remember that only 965 enzymes exist in a metabolic model and the values for the non-differentially expressed enzymes are assigned as 1 to show no change. Enzymes that have all the required values for the regulation decision calculation are collected in the Figure 4.20.

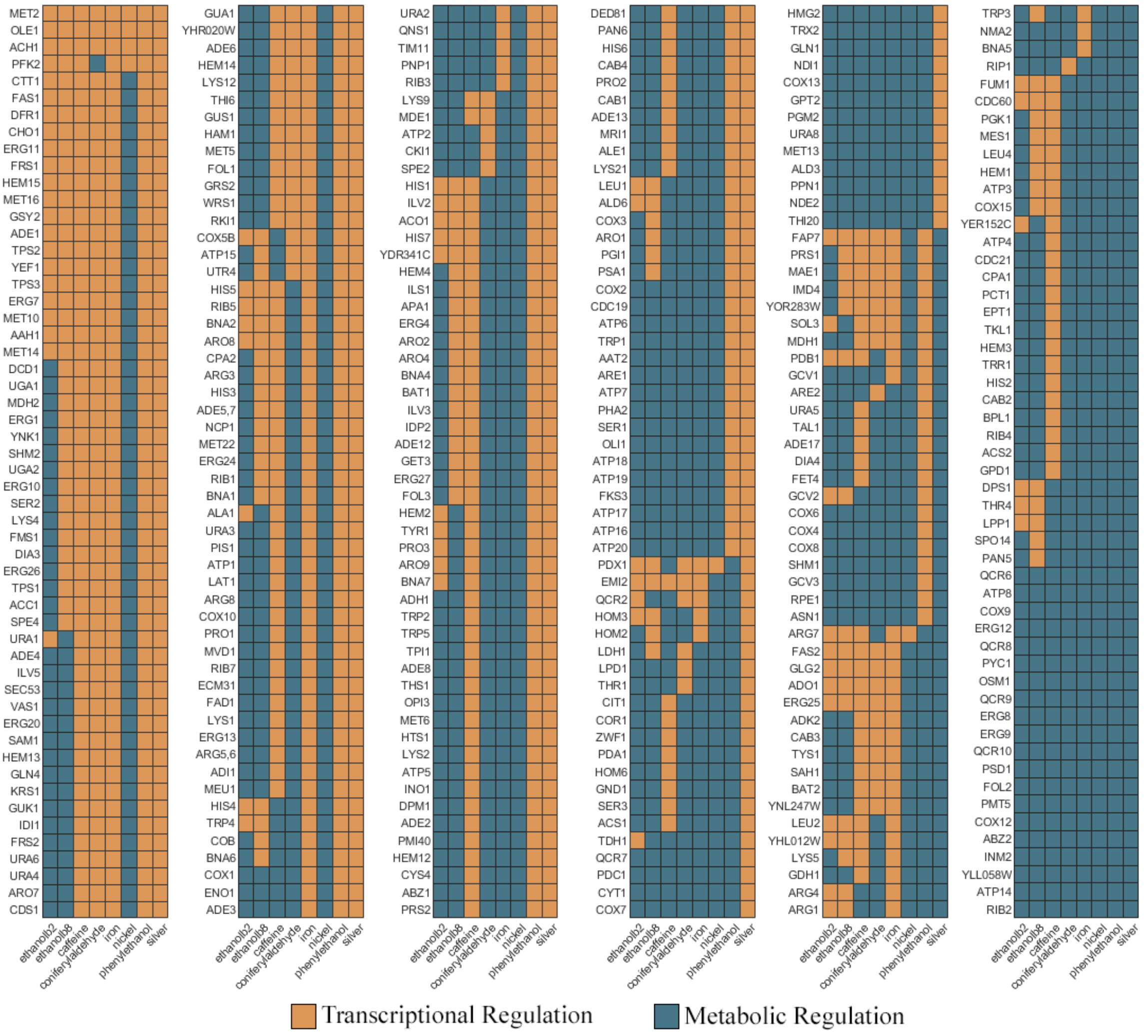


Figure 4.20. Regulation of enzymes according to the table 4.9 shown as heatmap.

Table 4.9. Enzymatic regulation decision table [118] of expressional fold change (from differential analysis) and flux fold change (from flux balance analysis) comparison.

		Flux FC		
		Positive	**Negative**	**No Change**
ExprFC	**Positive**	Transcriptional	Metabolic	Post-transcriptional
	Negative	Metabolic	Transcriptional	Post-transcriptional
	No Change	Metabolic	Metabolic	-

Unfortunately, only the enzymes that carry a flux value in the FBA simulations (under no constraint on the exchange metabolites) of the reference strain were able to provide a fold change value, therefore, the enzymes with zero fluxes were discluded from the comparison. It is found that only three enzymes, MET2 (homoserine O-acetyltransferase), OLE1 (stearoyl-CoA 9-desaturase), and ACH1 (acetyl-CoA hydrolase), were transcriptionally regulated in all evolved strains. This means at least one of the strains' simulations did not correlate with the expressional results. In nickel-resistant strain, almost all the enzymes were regulated within metabolic states as expected considering the lack of high numbers of the differential expression input. Total of 20 enzymes show metabolic regulation in all strains: QCR6 (ubiquinol-cytochrome-c reductase subunit 6), ATP8 (F1F0 ATP synthase subunit 8), COX9 (cytochrome c oxidase subunit VIIa), ERG12 (mevalonate kinase), QCR8 (ubiquinol-cytochrome-c reductase subunit 8), PYC1 (pyruvate carboxylase 1), OSM1 (fumarate reductase), QCR9 (ubiquinol-cytochrome-c reductase subunit 9), ERG8 (phosphomevalonate kinase), ERG9 (bifunctional farnesyl-diphosphate farnesyltransferase/squalene synthase), QCR10 (ubiquinol–cytochrome-c reductase subunit 10), PSD1 (phosphatidylserine decarboxylase 1), FOL2 (GTP cyclohydrolase I), PMT5 (putative dolichyl-phosphate-mannose-protein mannosyltransferase), COX12 (cytochrome c oxidase subunit VIb), ABZ2 (aminodeoxychorismate lyase ABZ2), INM2 (inositol monophosphate 1-phosphatase), YLL058W (cystathionine gamma-synthase), ATP14 (F1F0 ATP synthase subunit h), RIB2 (bifunctional DRAP deaminase/tRNA pseudouridine synthase).

4.3.7. Shadow Prices

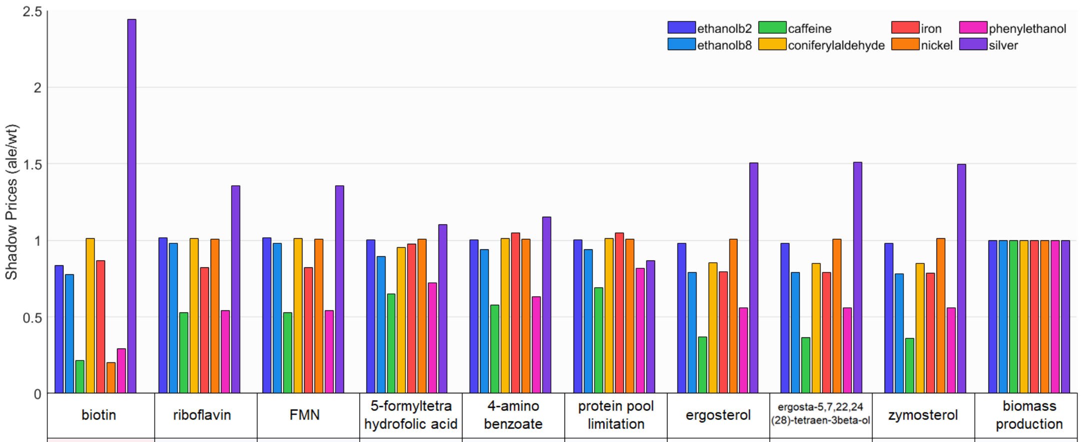
Shadow prices are calculated by solving the optimization problem for the objective function (growth rate) by increasing the upper constraints by one unit for all reactions. This way, the sensitivity of the reactions to the growth rate is calculated for all models.

In Table 4.10, the shadow prices of the reactions that limit the growth rate and their values can be seen. Only the reactions that has a higher negative value compared to growth reaction are included. The values for each reaction are divided to the wild-type model's shadow price values to uncover the adaptive evolution effects and plotted in Figure 4.21.

Table 4.10. Shadow prices of the exchange metabolites' uptake reactions that has a value higher than the growth itself.

exchange metabolites	**b2-eth**	**b8-eth**	**caff.**	**con. ald.**	**iron**	**nickel**	**phen.**	**silver**	**ref.**
biotin	-78.78	-76.60	-38.46	-104.17	-82.63	-99.87	-32.60	-255.57	-100.10
riboflavin	-44.24	-44.65	-43.65	-47.91	-36.10	-46.06	-27.84	-65.51	-46.16
flavin mononucleotide (FMN)	-44.24	-44.65	-43.65	-47.91	-36.10	-46.06	-27.84	-65.51	-46.16
5-formyltetrahydrofolic acid	-13.92	-12.99	-16.91	-14.40	-13.70	-14.69	-11.86	-16.92	-14.72
4-aminobenzoate	-6.81	-6.68	-7.41	-7.50	-7.21	-7.19	-5.09	-8.69	-7.20
ergosterol	-3.47	-2.93	-2.41	-3.30	-2.85	-3.75	-2.42	-5.77	-3.76
ergosta-5,7,22,24-(28)-tetraen-3beta-ol	-3.38	-2.86	-2.35	-3.21	-2.76	-3.67	-2.37	-5.65	-3.67
zymosterol	-3.19	-2.67	-2.21	-3.04	-2.59	-3.46	-2.24	-5.28	-3.46
14-demethyllanosterol	-2.95	-2.47	-2.02	-2.79	-2.32	-3.14	-1.99	-4.94	-3.14
hexanoate	-2.69	-3.43	-6.28	-5.48	-4.91	-3.56	-5.11	-2.07	-3.58
lanosterol	-2.13	-1.84	-1.60	-1.88	-1.67	-2.26	-1.57	-3.31	-2.26
growth	-1	-1	-1	-1	-1	-1	-1	-1	-1

Shadow prices of metabolites show us that the metabolites zymosterol, ergosta-5,7,22,24(28)-tetraen-3beta-ol, ergosterol, lanosterol, 14-demethyllanosterol, biotin, riboflavin, flavin mononucleotide (FMN), 5-formyltetrahydrofolic acid, 4-aminobenzoate and hexanoate are the most required metabolites in the media to increase growth rate in all models, in the order of their mean values across models. Although all the models would increase their biomass production with these metabolites supplied, increment amounts would differ for each model. While the silver-resistant model would increase its growth rate by 250% with biotin supplement, phenylethanol-resistant model could only achieve 33% increase.



	biotin	riboflavin	FMN	5-formyltetra hydrofolic acid	4-amino benzoate	protein pool limitation	ergosterol	ergosta-5,7,22,24 (28)-tetraen-3beta-ol	zymosterol	biomass production
ethanolb2	-0.26	0.02	0.02	0.00	0.00	0.00	-0.03	-0.03	-0.03	0
ethanolb8	-0.37	-0.03	-0.03	-0.16	-0.09	-0.09	-0.34	-0.34	-0.36	0
caffeine	-2.22	-0.92	-0.92	-0.63	-0.79	-0.53	-1.44	-1.45	-1.48	0
coniferylaldehyde	0.02	0.01	0.01	-0.07	0.02	0.02	-0.23	-0.24	-0.23	0
iron	-0.21	-0.28	-0.28	-0.03	0.07	0.07	-0.33	-0.34	-0.35	0
nickel	-2.33	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0
phenylethanol	-1.78	-0.89	-0.89	-0.47	-0.66	-0.29	-0.84	-0.84	-0.84	0
silver	1.29	0.44	0.44	0.14	0.21	-0.21	0.59	0.60	0.58	0

Figure 4.21. Barplots of the changes in the shadow prices of the metabolites that have contributions higher than the biomass (growth) itself. Shadow price of each reaction for evolved models is divided by the shadow price of that reaction in the reference model.

Sensitivity of silver-resistant strain to biotin showed an increase almost 2.5 fold whereas caffeine, nickel and phenylethanol-resistant strains decreased their sensitivities down 4 folds. FMN and riboflavin sensitivities resulted almost the same in all strains because their reactions are defined as a chain reaction in the model, therefore they carry the same flux values and shadow prices. The same can also applied to the reactions of ergosterol, and ergosta-5,7,22,24(28)-tetraen-3beta-ol which is the product of ergosterol oxidation. Zymosterol of cholesterol biosynthesis pathway, 4-aminobenzoate in folate biosynthesis, and 5-formyltetrahydrofolic acid which is folate coenzyme, are also showed high shadow prices.

Reported metabolites with increased or decreased sensitivities direct us to investigate metabolic changes in the energy maintenance reactions (especially coenzyme-A and FAD utilization), cell division pathway, and cell wall formation reactions (sterol and cholesterol biosynthesis).

4.3.8. Co-factor Production

When we look at the main ATP producing reactions in the metabolic models, we see the pyruvate kinase,

$$ADP + H^{+} + phosphoenolpyruvate \xrightarrow{PK} ATP + pyruvate \tag{4.1}$$

phosphoglycerate kinase,

$$\begin{aligned} 1,3 - bisphospho - D - glycerate + ADP \xrightarrow{PGK1} \\ 3 - phosphonato - D - glycerate^{3-} + ATP \end{aligned} \tag{4.2}$$

and, ATP synthase for the oxidative phosphorylation,

$$ADP + 3H^{+} + phosphate \xrightarrow{ATPsynthase} ATP + 2H^{+} + H2O \tag{4.3}$$

are the reactions that carry flux and actively used. Flux comparisons under both unconstrained and constrained system (where glucose uptake rate is constrained to maximum 10 mmol/gDWh) can be found in Table 4.11.

Total values of ATP production do not differ when the carbon supply is almost the same for all models. On the other hand, if we compare the simulation under free carbon source (NC columns), differences arise. Interestingly, caffeine and phenylethanol-resistant models do not prefer to produce ATP through the oxidative phosphorylation (ATP synthase) under free carbon source condition. We should consider maximum growth rate and glucose uptake rate for these models, therefore Figure 4.22 is created with the values from Table 4.4.

Table 4.11. Carried flux values (mmol/gDWh) through ATP producing reactions (A: the only constraint applied to the models is the maximum protein availability; B: additional constraint is applied to the glucose uptake rate as 10 mmol/gDW).

	eth-b2		**eth-b8**		**caffeine**		**con. ald.**		**iron**		**nickel**		**phen.**		**silver**		**ref.**	
	A	**B**	**A**	**B**	**A**	**B**	**A**	**B**	**A**	**B**	**A**	**B**	**A**	**B**	**A**	**B**	**A**	**B**
Pyruvate kinase	23.6	17.1	27.8	17.2	48.5	17.3	36.2	16.9	24.8	17	32.2	17	32.2	17.2	36.9	17.7	32.3	17
Phosphoglycerate kinase	24.1	17.6	28.3	17.7	48.8	17.5	36.8	17.5	25.3	17.5	32.8	17.5	32.7	17.7	37.4	18.2	32.9	17.6
ATP Synthase	7	12.5	4.2	11.6	0	12.1	0.1	14.7	7.7	14.2	0.1	13.8	0	10.7	0	9.8	0.1	13.7
Total	54.8	47.3	60.3	46.5	97.2	47	73.2	49.1	57.8	48.7	65.2	48.4	65	45.6	74.3	45.7	65.3	48.3

As expected for the caffeine-resistant model which has the highest glucose uptake rate and the total ATP production flux, it has the highest growth rate among all models. silver-resistant model, however, has a lower growth rate despite having the second highest glucose uptake rate and total ATP production flux. Despite the low fluxes on the glucose uptake and ATP production, the iron-resistant model shows high growth rate. The similar values from the reference strain and nickel-resistant models were also expected since the lowest amount of gene expression information was integrated for the nickel-resistant model, making it the least differentiated model from the reference strain.

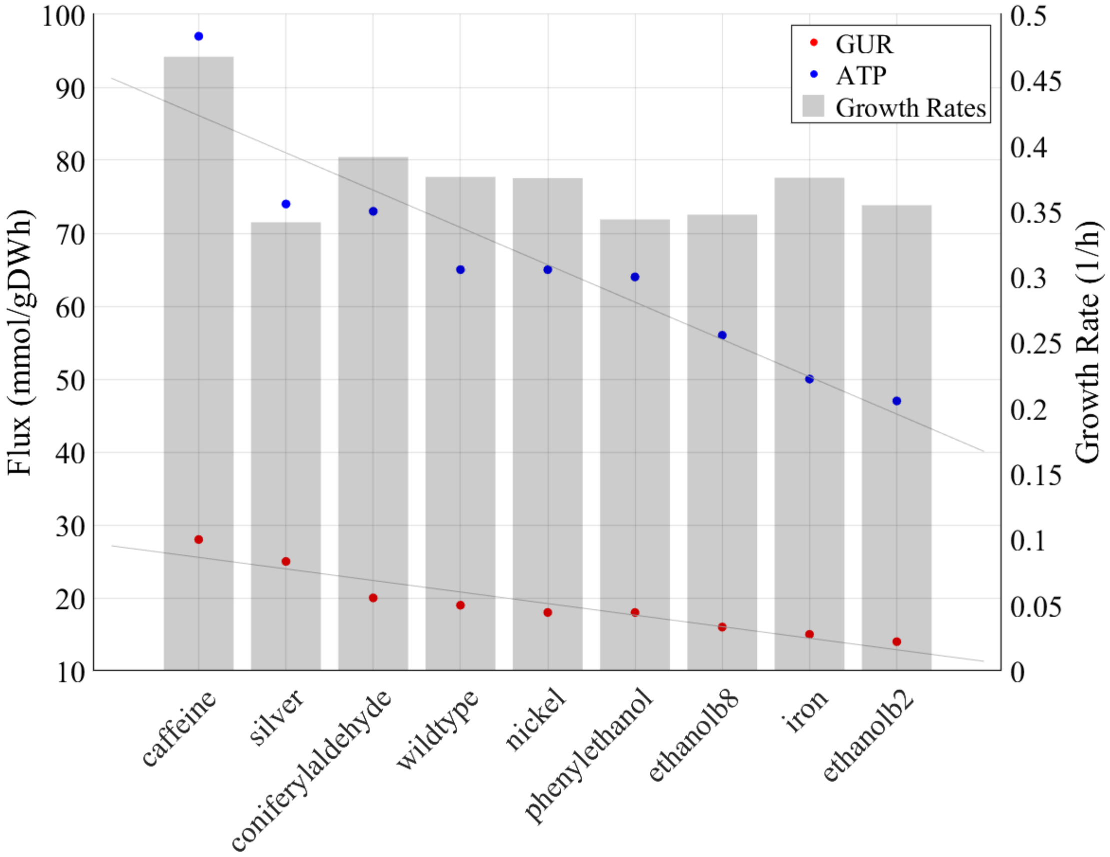
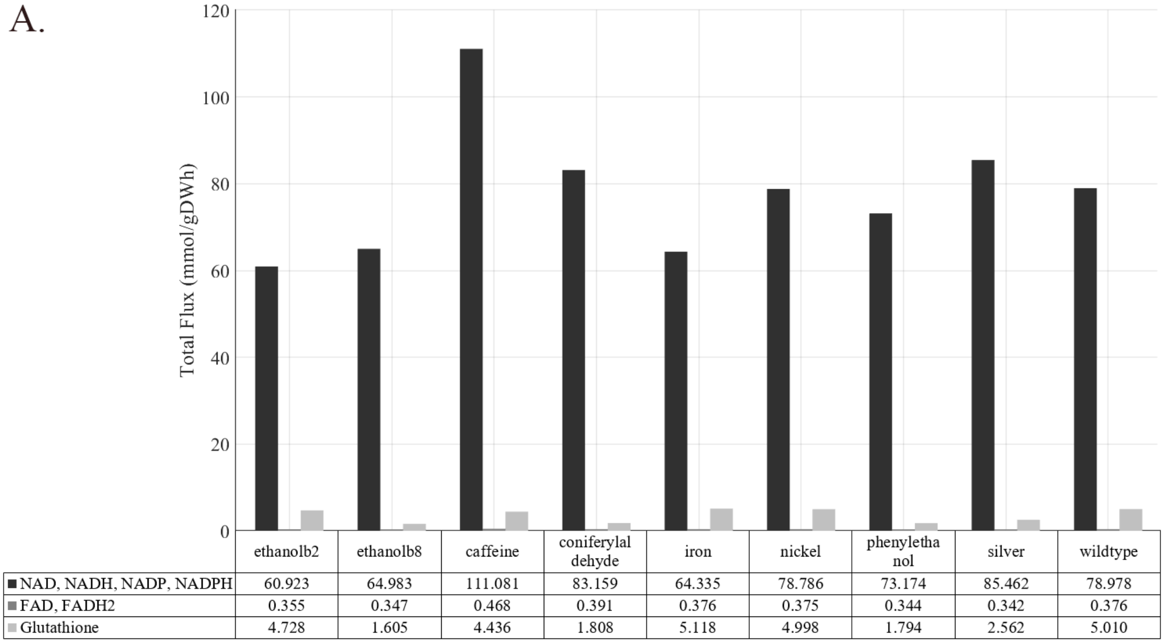


Figure 4.22. Glucose uptake rates and total fluxes through ATP producing reactions are plotted with least-squares lines. Growth rates are shown as bar graphs on the secondary y-axis on the right.

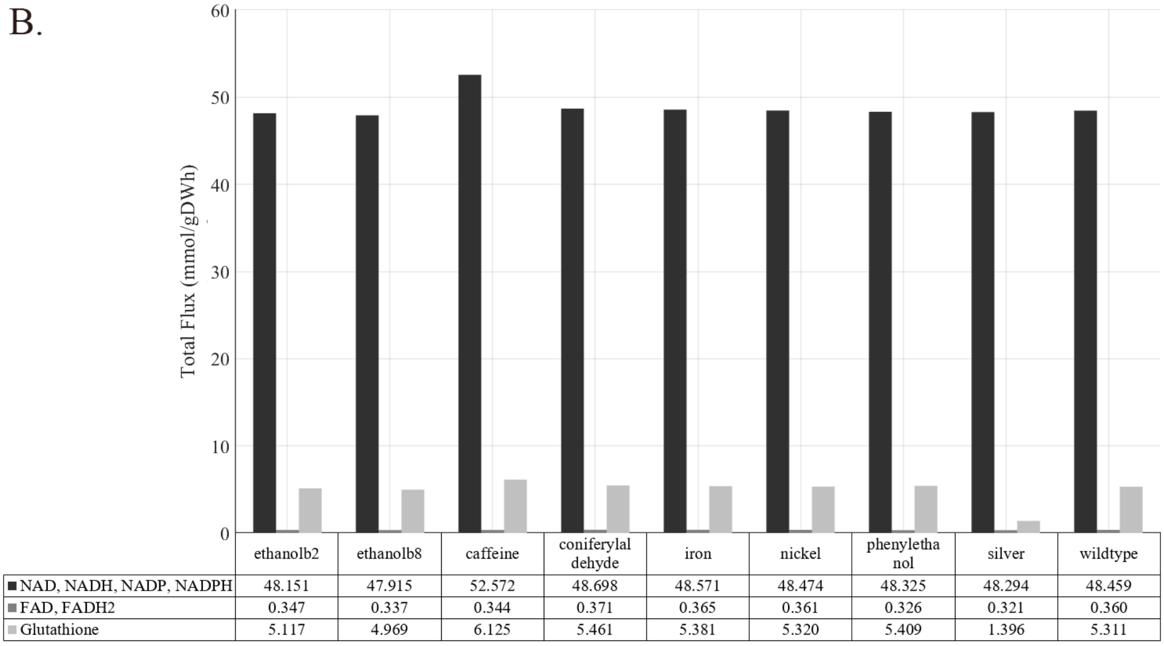
When we look at the total coenzymes NAD, NADP, FAD consumptions and the glutathione metabolism fluxes, similar results to the ATP regulations are observed with the exception of a significant difference on the glutathione metabolism in the silver-resistant model. When the glucose uptake rate is constrained same for all models, we have collected similar results from all models. Surprisingly, the silver-resistant model was not able to carry high fluxes through glutathione metabolism under the same glucose uptake rate conditions. This could explain why the silver-resistant model was unable to grow at high rates despite carrying high fluxes on the glucose uptake and ATP producing reactions (Figure 4.22).

A.



	ethanolb2	ethanolb8	caffeine	coniferylal dehyde	iron	nickel	phenyletha nol	silver	wildtype
■ NAD, NADH, NADP, NADPH	60.923	64.983	111.081	83.159	64.335	78.786	73.174	85.462	78.978
■ FAD, FADH2	0.355	0.347	0.468	0.391	0.376	0.375	0.344	0.342	0.376
■ Glutathione	4.728	1.605	4.436	1.808	5.118	4.998	1.794	2.562	5.010

B.



	ethanolb2	ethanolb8	caffeine	coniferylal dehyde	iron	nickel	phenyletha nol	silver	wildtype
■ NAD, NADH, NADP, NADPH	48.151	47.915	52.572	48.698	48.571	48.474	48.325	48.294	48.459
■ FAD, FADH2	0.347	0.337	0.344	0.371	0.365	0.361	0.326	0.321	0.360
■ Glutathione	5.117	4.969	6.125	5.461	5.381	5.320	5.409	1.396	5.311

Figure 4.23. Total flux values of NAD/NADP, FAD consumption and glutathione metabolism reactions of the simulations where no additional constraint to uptake reactions is applied (A), and the glucose uptake rate is constrained to 10 mmol/gDWh (B), for each model.

4.3.9. Amino Acid Biosynthesis

Flux distribution within the amino acid biosynthesis subsystem in the models are collected in the Table 4.12. The reactions enolase, phosphoglycerate mutase, pyruvate kinase, glyceraldehyde-3-phosphate dehydrogenase are the most active reactions in the amino acid biosynthesis. Although most of the flux values are in parallel with the corresponding glucose uptake rates, some differences are observed. 2-isopropylmalate synthase enzyme is only used by coniferyl aldehyde andiron-resistant models. While caffeine-resistant model chooses to use 3-deoxy-D-arabino-heptulosonate 7-phosphate synthetase enzyme, all other models carry their fluxes through 2-deoxy-D-arabino-heptulosonate 7-phosphate synthetase. The caffeine-resistant model also does not carry a flux through phosphoglycerate dehydrogenase. And most interestingly, the silver-resistant model carries flux with lower values in the reverse direction through transketolase 1 and transketolase 2 reactions,

$$\begin{aligned} D-xylulose-5-phosphate+ribose-5-phosphate \xrightarrow{transketolase1} \\ glyceraldehyde-3-phosphate+sedoheptulose-7-phosphate \\ D-erythrose-4-phosphate+D-xylulose-5-phosphate \xrightarrow{transketolase2} \\ D-fructose-6-phosphate+glyceraldehyde-3-phosphate \end{aligned} \tag{4.4}$$

as the forward direction is shown.

Table 4.12. Fluxes through the reactions in the amino acids biosynthesis subsystem.

Reactions	**b2-eth.**	**b8-eth.**	**caff.**	**con. ald.**	**iron**	**nickel**	**phen.**	**silver**	**ref.**
2-aceto-2-hydroxy-butanoate synthase	0.085	0.084	0.113	0.094	0.090	0.090	0.083	0.082	0.091
2-aminoadipate transaminase	0.127	0.124	0.167	0.140	0.134	0.134	0.123	0.122	0.134
2-deoxy-D-arabino-heptulosonate-7-phosphate synthetase	0.119	0.117	0	0.132	0.126	0.126	0.116	0.115	0.126
2-isopropylmalate synthase	0	0	0	0.145	0.139	0	0	0	0
3-deoxy-D-arabino-heptulosonate-7-phosphate synthetase	0	0	0.157	0	0	0	0	0	0
acetolactate synthase	0.249	0.243	0.328	0.274	0.263	0.263	0.241	0.239	0.264
citrate synthase	0.559	0.547	0.595	0.617	0.592	0.591	0.437	0.539	0.593
enolase	23.842	28.028	48.769	36.501	25.038	32.485	32.463	37.133	32.564
glyceraldehyde-3-phosphate dehydrogenase	24.128	28.308	48.769	36.816	25.341	32.787	32.740	37.397	32.867
homoacontinate hydratase	0.127	0.124	0.167	0.140	0.134	0.134	0.123	0.122	0.134
methionine adenosyltransferase	0.046	0.045	0.061	0.051	0.049	0.049	0.045	0.045	0.049
O-succinylhomoserine lyase (reverse)	0.003	0.003	0.004	0.003	0.003	0.003	0.003	0.003	0.003
phosphofructokinase	11.896	14.158	25.467	18.412	12.486	16.216	16.356	20.641	16.256
phosphoglycerate dehydrogenase	0.286	0.280	0	0.315	0.302	0.302	0.277	0.264	0.303
phosphoglycerate mutase	23.842	28.028	48.769	36.501	25.038	32.485	32.463	37.133	32.564
phosphoribosylpyro-phosphate synthetase	0.121	0.118	0.159	0.133	0.128	0.128	0.117	0.116	0.128
pyruvate carboxylase	1.264	1.237	2.450	1.394	1.121	1.337	0.899	0.902	1.340
pyruvate kinase	23.603	27.795	48.454	36.238	24.786	32.233	32.232	36.903	32.311
transaldolase (rev.)	11.434	14.041	25.041	18.281	11.984	15.728	16.206	20.642	15.766
transketolase 1	0.462	0.117	0.426	0.132	0.502	0.489	0.150	0	0.490
transketolase 1 (rev.)	0	0	0	0	0	0	0	0.001	0
transketolase 2	0.343	0	0.268	0	0.376	0.363	0.034	0	0.364
transketolase 2 (rev.)	0	0	0	0	0	0	0	0.116	0

4.3.10. Precursor Metabolites for Biomass Production

In the metabolic model, there are 12 branch-point metabolites that are considered as the precursor metabolites for the central carbon metabolism in the network. Producing and consuming reactions of these 12 precursor metabolites (Table 4.13) are investigated individually for all models.

Table 4.13. The 12 precursor (branchpoint) metabolites for biomass production.

#	**Metabolite**	**Abbreviation**	**Building Blocks Produced**
1	D-glucose-6-phosphate	G6P	glycogen, LPS
2	D-fructose-6-phosphate	F6P	cell wall
3	D-ribose-5-phosphate	R5P	His, Phe, Trp, nucleotides
4	D-erythrose-4-phosphate	E4P	Phe, Trp, Tyr
5	D-glyceraldehyde-3- phosphate	GAP	lipids
6	glycerate-3-phosphate	3PG	Cys, Gly, Ser
7	phosphoenolpyruvate	PEP	Tyr, Trp
8	pyruvate	PYR	Ala, Ile, Lys, Leu, Va
9	acetyl-CoA	ACA	Leu, lipids
10	2-ketoglutarate	2KG	Glu, Gln, Arg, Pro
11	succinyl-CoA	SCA	Met, Lys, tetrapyrroles (e.g., heme)
12	oxaloacetate	OXA	Asn, Asp, Ile, Lys, Met, Thr, nucleotides

For the first precursor metabolite, D-glucose-6-phosphate, only hexokinase reaction produces the metabolite,

$$ATP + \text{D-glucose} \xrightarrow{hexokinase} ADP + \text{D-glucose 6-phosphate} + \text{H+} \tag{4.5}$$

and the flux values for all models is the same at the provided glucose uptake rate (in this case, 10 mmol/gDWh). The main difference is the activated gene for the reaction: While the silver-resistant model prefers to use the enzyme HXK1 (YFR053C), phenylethanol-resistant model uses the enzyme HXK2 (YGL253W), while all the other models use EMI2 (YDR516C) for the same reaction activity.

Consuming reactions for the G6P are the glucose-6-phosphate isomerase, glucose 6-phosphate dehydrogenase, phosphoglucomutase, alpha,alpha-trehalose-phosphate synthase (UDP-forming), and myo-inositol-1-phosphate synthase catalyzed reactions. Sum of all the consuming reactions is the same as the production rate, 10 mmol/gDWh, as expected. The only differentiating result is observed for the silver-resistant model, all models carry 78.5% flux through the isomerase catalyzed reaction and 16.2% of total G6P consumption fluxes through dehydrogenase catalyzed reaction, the silver model carries 92.2% of flux through the isomerase catalyzed reaction and 2.9% through the dehydrogenase catalyzed reaction.

Producing and consuming reaction fluxes for the metabolite D-fructose-6-phosphate (F6P) can be seen in the Table 4.14. Similar to the G6P results, the only noticeable difference arises from the silver-resistant model.

Table 4.14. Producing (P) and consuming (C) reaction fluxes for the metabolite D-fructose-6-phosphate (F6P).

F6P	**Reaction Name**	**Genes**	**b2-eth**	**b8-eth**	**caffeine**	**con.ald.**	**iron**	**nickel**	**phen.**	**silver**	**ref.**
P	glucose-6-phosphate isomerase	PGI1	7.95	8.01	7.62	7.82	7.85	7.87	7.88	9.22	7.88
	transketolase 2	TKL1	0.39	0.38	0.5	0.42	0.41	0.41	0.43	0	0.41
C	transaldolase	TAL1, NQM1	8.12	8.17	7.9	7.99	8.02	8.04	8.1	9	8.05
	mannose-6-phosphate isomerase	PMI40	0.23	0.22	0.22	0.24	0.24	0.23	0.21	0.21	0.23
	transketolase 2	TKL2, TKL1	0	0	0	0	0	0	0	0.01	0

The metabolite d-ribulose-5-phosphate (R5P) is produced by the phosphogluconate dehydrogenase reaction that has two isozymes acting on it, GND1 (YHR183W) and GND2 (YGR256W). From the evolved model simulations, caffeine, coniferyl aldehyde, iron, nickel and phenylethanol-resistant models use the GND1 isozyme while b2-ethanol, b8-ethanol and silver-resistant models use the GND2 isozyme.

Ribulose 5-phosphate 3-epimerase and ribose-5-phosphate isomerase reactions are the only active reactions for the consumption of the metabolite R5P. Flux values of both production and consumption rates can be found in the Table 4.15.

Table 4.15. Producing (P) and consuming (C) reaction fluxes for the metabolite d-ribulose-5-phosphate (R5P).

R5P	**Reaction Name**	**Genes**	**b2-eth**	**b8-eth**	**caffeine**	**con.ald.**	**iron**	**nickel**	**phen.**	**silver**	**ref.**
P	phosphogluconate dehydrogenase	GND1	0	0	1.86	1.63	1.6	1.58	1.63	0	1.58
	phosphogluconate dehydrogenase	GND2	1.52	1.48	0	0	0	0	0	0.3	0
C	ribulose 5-phosphate 3-epimerase	YJL121C	0.9	0.87	1.12	0.96	0.94	0.93	0.98	0.09	0.93
	ribose-5-phosphate isomerase	RKI1	0.63	0.61	0.74	0.67	0.66	0.65	0.65	0.21	0.65
	3,4-dihydroxy-2-butanone-4-phosphate synthase	RIB3	0	0	0	0	0	0	0	0	0

D-erythrose-4-phosphate producing and consuming reactions are collected in Table 4.16. Interestingly, the reversible transketolase 2 reaction shows activity in the opposite direction with the flux value that is close to 0 for the silver-resistant model only.

Table 4.16. Producing (P) and consuming (C) reaction fluxes for the metabolite d-erythrose-4-phosphate (E4P).

E4P	**Reaction Name**	**Genes**	**b2-eth**	**b8-eth**	**caffeine**	**con.ald.**	**iron**	**nickel**	**phen.**	**silver**	**ref.**
P	sedoheptulose-1,7-bisphosphate-D-glyceraldehyde-3-phosphate-lyase	FBA1	8.63	8.67	8.52	8.53	8.56	8.57	8.64	9.1	8.57
	transketolase 2 (reverse)	TKL2	0	0	0	0	0	0	0	0.01	0
C	transaldolase (reverse)	TAL1, NQM1	8.12	8.17	7.9	7.99	8.02	8.04	8.1	9	8.05
	transketolase 2	TKL2, TKL1	0.39	0.38	0.5	0.42	0.41	0.41	0.43	0	0.41
	D-erythrose 4-phosphate transport	-	0.12	0.11	0.12	0.12	0.12	0.12	0.11	0.11	0.12

Similar pattern for the transketolase reaction in the silver-resistant model can also be seen for the glyceraldehyde-3-phosphate consuming reactions (Table 4.17). In all models except the silver-resistant model, transaldolase reaction uses the TAL1 (YLR354C) enzyme for its activity; however, in the silver-resistant model, the enzyme NQM1 (YGR043C) is active for the same purpose.

Table 4.17. Producing (P) and consuming (C) reaction fluxes for the metabolite glyceraldehyde-3-phosphate (GAP).

GAP	Reaction Name	Genes	b2-eth	b8-eth	caffeine	con.ald.	iron	nickel	phen.	silver	ref.
P	transaldolase (reverse)	NQM1	0	0	0	0	0	0	0	9	0
	transaldolase (reverse)	TAL1	8.12	8.17	7.9	7.99	8.02	8.04	8.1	0	8.05
	transketolase 1	TKL1	0.51	0.49	0.62	0.54	0.53	0.53	0.54	0	0.53
	triose-phosphate isomerase	TPI1	8.6	8.64	8.5	8.51	8.53	8.55	8.62	9.08	8.55
	transketolase 2	TKL1	0.39	0.38	0.5	0.42	0.41	0.41	0.43	0	0.41
	transketolase 1	TKL2	0	0	0	0	0	0	0	0.1	0
	tryptophan synthase	TRP5	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.01
C	glyceraldehyde-3-phosphate dehydrogenase	TDH3, TDH1, TDH2	17.64	17.7	17.54	17.48	17.51	17.54	17.71	18.18	17.55
	transketolase 2 (reverse)	TKL2, TKL1	0	0	0	0	0	0	0	0.01	0

Flux values for the producing and consuming reactions of the metabolite 3-phosphonato-D-glycerate (3PG) are listed in the Table 4.18. The difference is observed in the caffeine-resistant model as the inactivity of the phosphoglycerate dehydrogenase reaction while the reaction is active for other models. caffeine-resistant model consumes the metabolite with the mutase reaction activity. The phosphoenolpyruvate consuming and producing reactions are collected in the Table 4.19.

Table 4.18. Producing (P) and consuming (C) reaction fluxes for the metabolite 3-phosphonato-D-glycerate (3PG).

3PG	Reaction Name	Genes	b2-eth	b8-eth	caffeine	con.ald.	iron	nickel	phen.	silver	ref.
P	phosphoglycerate kinase	PGK1	17.64	17.7	17.54	17.48	17.51	17.54	17.71	18.18	17.55
C	phosphoglycerate mutase	GPM1, YOR283W	17.36	17.43	17.54	17.18	17.22	17.25	17.45	17.93	17.26
	phosphoglycerate dehydrogenase	SER3, SER33	0.28	0.27	0	0.3	0.29	0.29	0.26	0.25	0.29

Table 4.19. Producing (P) and consuming (C) reaction fluxes for the metabolite phosphoenolpyruvate (PEP).

PEP	**Reaction Name**	**Genes**	**b2-eth**	**b8-eth**	**caffeine**	**con.ald.**	**iron**	**nickel**	**phen.**	**silver**	**ref.**
P	phosphoglycerate mutase	GPM1, YOR283W	17.36	17.43	17.54	17.18	17.22	17.25	17.45	17.93	17.26
	phosphoglycerate dehydrogenase	SER3, SER33	0.28	0.27	0	0.3	0.29	0.29	0.26	0.25	0.29
C	pyruvate kinase	CDC19, PYK2	17.12	17.21	17.31	16.93	16.97	17.01	17.23	17.72	17.01
	citrate transport (reverse)	YBR291C	0.12	0.11	0.12	0.12	0.12	0.12	0.11	0.11	0.12
	3-phosphoshikimate-1-carboxyvinyltransferase	YDR127W	0.12	0.11	0.12	0.12	0.12	0.12	0.11	0.11	0.12

As one of the most important branchpoint metabolites, pyruvate, reactions for production and consumption of this metabolite do not differ much. As collected in the Table 4.20, it can be seen that main distinction is caused by the caffeine and silver-resistant models when they use the alanine glyoxylate aminotransferase reaction for the production of pyruvate. Although the values are small, this difference must be noted for future analyses.

Choosing a different isozyme for the pyruvate kinase in the silver-resistant model (PYK2, YOR347C; instead of PYK1 YAL038W) is observed as previous results. Flux value in the caffeine-resistant model through the phenylalanine transaminase (ARO9, YHR137W) almost triples the fluxes for the other models.

In the network, acetyl-CoA production is done by the enzyme acetyl-CoA synthetase with two isozymes, ACS1 (YAL054C) and ACS2 (YLR153C). As it can be seen from the Table 4.21, only the phenylethanol-resistant model prefers to carry flux through ACS1 isozyme and all other models uses the ACS2 isozyme.

As an intersection metabolite between central carbon metabolism and nitrogen metabolisms, producing and consuming reactions for the metabolite 2-oxoglutarate (2KG) is collected in Table 4.22. Activity of the oxoglutarate/malate exchange between mitochondria and cytoplasm is only observed in the silver-resistant model. As opposed to that, AKG transporter (mitochondrial) which exchanges citrate instead of malate is active in the reverse direction. Different preference for the isozymes of 2-aminoadipate transaminase reaction on the phenylethanol and silver-resistant models also observed.

Table 4.20. Producing (P) and consuming (C) reaction fluxes for the metabolite pyruvate (PYR).

PYR	**Reaction Name**	**Genes**	**b2-eth**	**b8-eth**	**caffeine**	**con.ald.**	**iron**	**nickel**	**phen.**	**silver**	**ref.**
P	pyruvate kinase	PYK2	0	0	0	0	0	0	0	17.72	0
	pyruvate kinase	CDC19	17.12	17.21	17.31	16.93	16.97	17.01	17.23	0	17.01
	alanine glyoxylate aminotransferase	AGX1	0	0	0.55	0	0	0	0	0.01	0
	anthranilate synthase	TRP2,TRP3	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01	0.01
C	pyruvate decarboxylase	PDC6, PDC1, PDC5	10.36	11.08	12.6	8.69	9.08	9.37	12.15	14.54	9.42
	pyruvate transport (reverse)	JEN1	4.16	3.6	1.57	5.46	5.16	4.93	2.63	0.76	4.89
	pyruvate transport	MPC1, MPC3, MPC2	1.51	1.47	2.05	1.61	1.59	1.57	1.42	1.41	1.57
	phenylalanine transaminase	ARO9	0.2	0.19	0.75	0.21	0.21	0.2	0.18	0.19	0.2
	pyruvate carboxylase	PYC2, PYC1	0.91	0.88	0.9	0.97	0.96	0.94	0.85	0.84	0.94

Table 4.21. Producing (P) and consuming (C) reaction fluxes for the metabolite acetyl-CoA (ACA).

ACA	**Reaction Name**	**Genes**	**b2-eth**	**b8-eth**	**caffeine**	**con.ald.**	**iron**	**nickel**	**phen.**	**silver**	**ref.**
P	acetyl-CoA synthetase	ACS2	0.43	0.42	0.42	0.46	0.45	0.44	0	0.4	0.44
	acetyl-CoA synthetase	ACS1	0	0	0	0	0	0	0.4	0	0
C	acetyl-CoA carboxylase	BPL1, ACC1	0.38	0.37	0.37	0.4	0.4	0.39	0.35	0.35	0.39
	fatty-acyl-CoA synthase (C16)	FAS1, FAS2	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
	fatty-acyl-CoA synthase (C18)	FAS1, FAS2	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

Table 4.22. Producing (P) and consuming (C) reaction fluxes for the metabolite 2-oxoglutarate (2KG).

2KG	Reaction Name	Genes	b2-eth	b8-eth	caffeine	con.ald.	iron	nickel	phen.	silver	ref.
P	oxoglutarate/malate exchange	YOR222W, YPL134C	0	0	0	0	0	0	0	0.58	0
	phenylalanine transaminase (r)	ARO8	0.25	0.25	0.81	0.27	0.27	0.26	0.24	0.24	0.26
	AKG transporter, mitochonrial (r)	YHM2	0.51	0.5	0.51	0.55	0.54	0.53	0.48	0	0.53
	phosphoserine transaminase	SER1	0.28	0.27	0	0.3	0.29	0.29	0.26	0.25	0.29
	2-aminoadipate transaminase	YER152C	0.12	0.12	0.12	0.13	0.13	0.13	0	0	0.13
	2-aminoadipate transaminase	ARO8	0	0	0	0	0	0	0.12	0	0
	2-aminoadipate transaminase	BNA3	0	0	0	0	0	0	0	0.11	0
	aspartate transaminase (reverse)	AAT2	0.45	0.43	0.44	0.48	0.47	0.46	0.42	0.41	0.46
	isocitrate dehydrogenase (NADP)	IDP2	0.22	0.21	0.22	0.23	0.23	0.23	0.21	0.2	0.23
	leucine transaminase (reverse)	BAT2	0.13	0.12	0.13	0.14	0.14	0.13	0.12	0.12	0.13
	saccharopine dehydrogenase	LYS1	0.12	0.12	0.12	0.13	0.13	0.13	0.12	0.11	0.13
	isoleucine transaminase (r)	BAT2	0.08	0.08	0.08	0.09	0.09	0.09	0.08	0.08	0.09
	tyrosine transaminase	ARO8	0.04	0.04	0.04	0.05	0.05	0.05	0.04	0.04	0.05
	histidinol-phosphate transaminase	HIS5	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
C	glutamate dehydrogenase (NADP)	YAL062W, YOR375C	2.02	1.96	2.28	2.15	2.12	2.1	1.89	1.86	2.09
	AKG transporter, mitochonrial	YHM2	0	0	0	0	0	0	0	0.1	0
	2-oxoadipate 2-oxoglutarate transport	ODC2, ODC1	0.12	0.12	0.12	0.13	0.13	0.13	0.12	0.11	0.13
	phosphoglycerate dehydrogenase (r)	-	0.11	0.1	0.1	0.11	0.11	0.11	0.1	0.1	0.11
	4-aminobutyrate transaminase	UGA1	0	0	0	0	0	0	0	0	0

Intermediate metabolite of the TCA cycle, oxaloacetate (OXA), is only produced from the pyruvate carboxylase reaction and consumed by the malate dehydrogenase and aspartate transaminase reactions, as collected in Table 4.23. All the flux values through these reactions are in close range with each other for all the models.

Table 4.23. Producing (P) and consuming (C) reaction fluxes for the metabolite oxaloacetate (OXA).

OXA	Reaction Name	Genes	b2-eth	b8-eth	caffeine	con.ald.	iron	nickel	phen.	silver	ref.
P	pyruvate carboxylase	PYC1	0.91	0.88	0.9	0.97	0.96	0.94	0.85	0.84	0.94
C	malate dehydrogenase, cytoplasmic (reverse)	MDH2	0.46	0.45	0.46	0.49	0.49	0.48	0.43	0.43	0.48
	aspartate transaminase (rev.)	AAT2	0.45	0.43	0.44	0.48	0.47	0.46	0.42	0.41	0.46

4.3.11. Bi-level Optimization for ATP Synthase

With the values obtained for biomass production in FBA simulations, the objective function could be set to another reaction. ATP synthase is chosen as objective function when the growt rate is constrained with the FBA results to dual optimize metabolic problem. A linearly spaced vector of from 0 to 0.5 is used to constrain lower bound of growth while the metabolic model solved for the minimization of ATP synthase. Figure 4.24 shows the minimum required ATP synthase activity to reach specific growth rate.

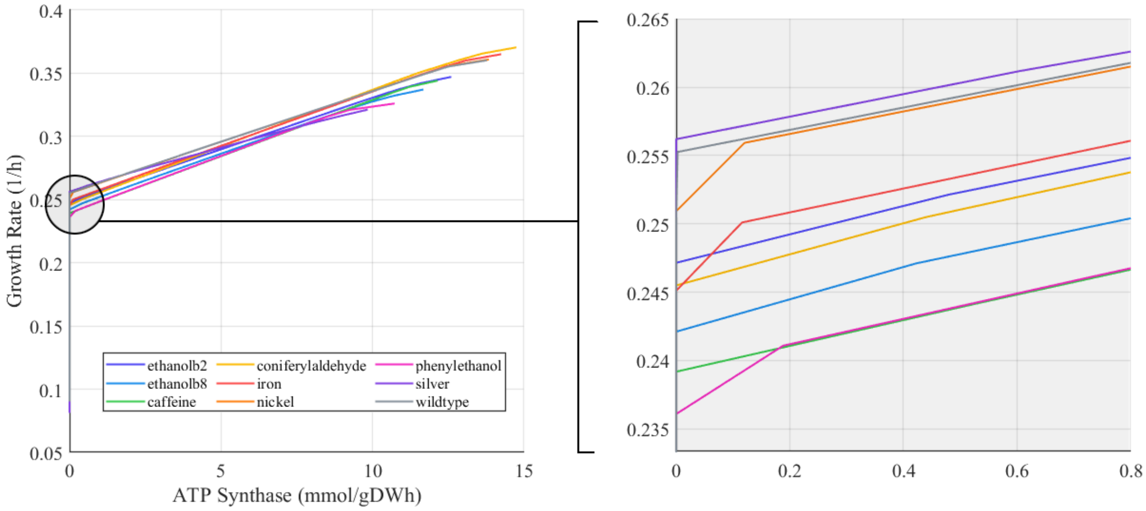


Figure 4.24. Bi-level optimization maximum growth and minimum ATP synthase activity. Lines where ATP synthase activity is close to 0 values are zoomed in.

Dual optimization showed that all the strains could grow without carrying any flux through ATP synthase at around 0.25 h^{-1}. Specifically, ethanol B8, caffeine and phenylethanol-resistant strains could grow at the rate of 0.23 h^{-1}, ethanol b2, coniferyl aldehyde and iron-resistant strains at 0.24 h^{-1}, and finally, silver-resistant and reference strains could grow at 0.26 h^{-1} while the ATP synthase reaction carries zero flux.

A linear pattern in growth rate vs. ATP synthase activity is observed until the 8.60 mmol $(gDW)^{-1}$ h^{-1}) of ATP synthase activity, at which the phenylethanol-resistant strain could not faster than 0.32 h^{-1}. Ethanol resistant strain B8 break this linearity at 0.17 mmol $(\mathrm{gDW})^{-1}$ h^{-1}) ATP synthase rate and 0.25 h^{-1} growth rate. In agreement with the previous results, the investigation around the knee point showed that the maximum flux differences were on the reactions of glyceraldehyde-3-phosphate dehydrogenase, enolase, and pyruvate decarboxylase reactions.

4.4. Minimization of Metabolic Adjustment

Minimization of metabolic adjustment (MOMA) analysis is applied to evolved models to find the closest points to wild-type model simulations (in terms of the Euclidean distances). Analysis is done with the objective of maximization of the growth rate and no constraints applied to the models except for the total protein amount availability. Results are collected for the metabolic reactions and the proteins (as enzyme-draw reactions) separately. The sum of the distances from the reference strain simulations is used as a ranking system to find most distant reactions and proteins. Individual fluxes through reactions can be seen in Figure 4.25.

The most distant reactions in the evolved models from the wild type model found as palmitoyl-CoA hydrolase, followed by the long chain fatty acid CoA ligase, deoxyuridine kinase, methylglyoxal synthase, 5' nucleosidases, glutathione metabolism reactions. On the other hand, the enzymes TDH1 (glyceraldehyde-3-phosphate dehydrogenase isozyme 1), OLI1 (F0 ATP synthase subunit c), GPT2 (glycerol-3-phosphate / dihydroxyacetone phosphate sn-1 acyltransferase), RIB7 (Diamino-hydroxy-phoshoribosylamino-pyrimidine deaminase), SCT1 (glycerol 3-phosphate / dihydroxy acetone phosphate sn-1 acyltransferase) and COB (cytochrome b) were the top distant enzymes in terms of enzyme-draw reaction fluxes.

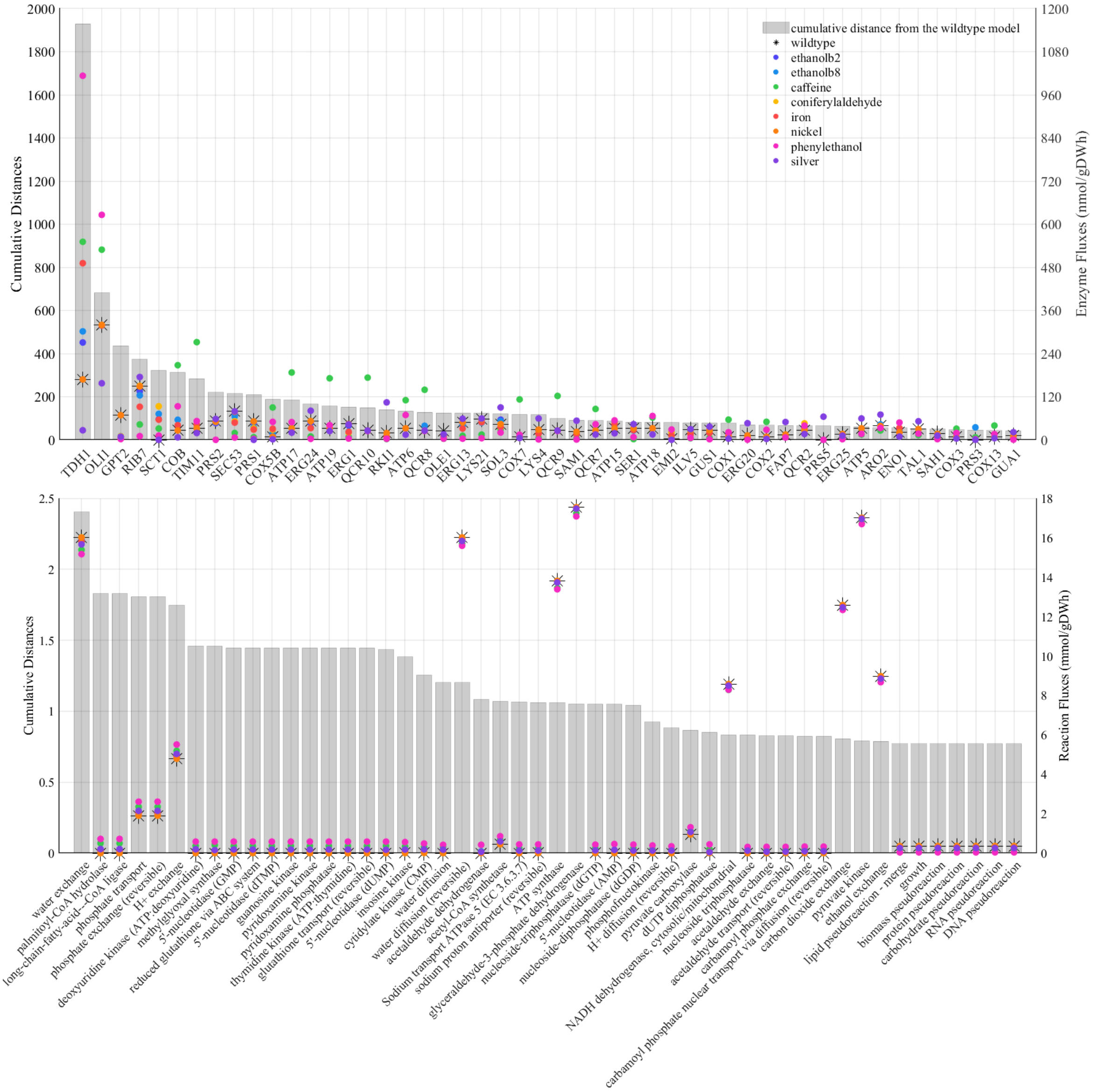


Figure 4.25. Enzymes that have MOMA flux values the most distant from the wild-type flux values in nmol/gDWh.

Draw reaction RIB7 showed 14.73-fold decrease in phenylethanol-resistant strain, while the fluxes for the others were decreased only 1.48±0.93 fold on average compared to the reference strain. phenylethanol-resistant strain decreased 12.72 fold for the enzyme SEC53.

4.5. Flux Variability Analysis

Flux ranges of each enzyme to find active enzymes and pathways on models are analyzed by the linear problem in the Flux Variability Analysis method where the objective functions to minimize and maximize all reactions iteratively with growth rate as the objective function kept at 90% of its maximum value. Minimum and maximum available fluxes are collected in the iterative process for each reaction, and results are plotted in Figure 4.5). Standard deviations on flux ranges for each reaction are also plotted to find most divergent targets.

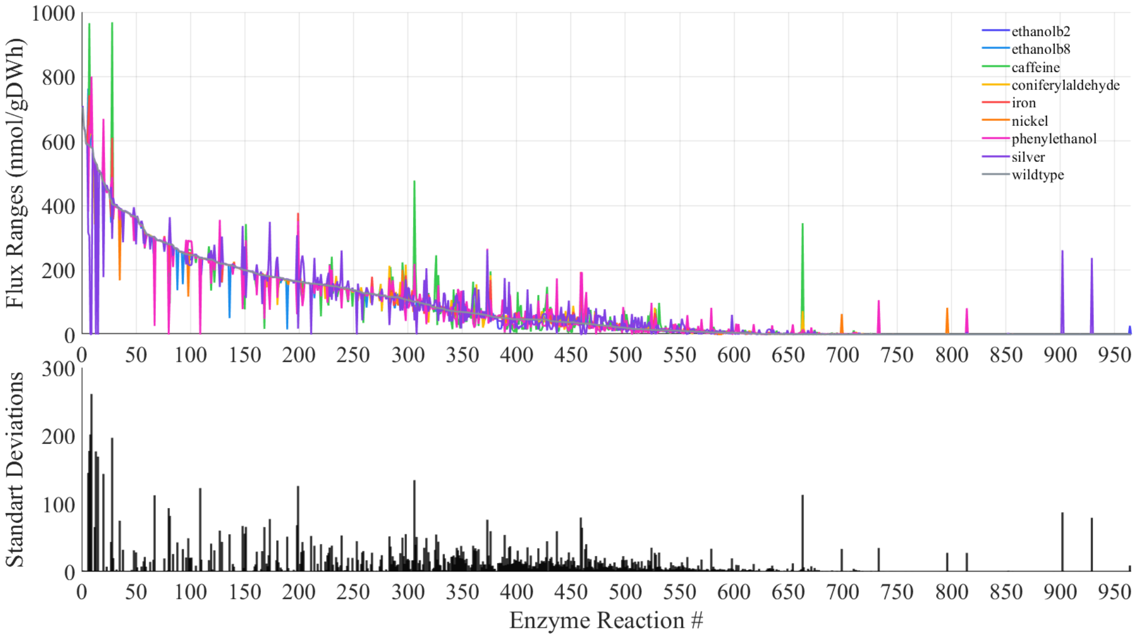


Figure 4.26. Flux variability analysis as flux ranges per enzyme usage reaction.

As it can be seen in the Table 4.24 where the most divergent proteins (std > 50) across all experiments are listed, paralog components of the glutathione system namely glutaredoxin-1 (GRX1) and glutaredoxin-2 (GRX2) shows the highest divergence across strains, followed by the glyceraldehyde-3-phosphate dehydrogenase forming isozymes triose-phosphate dehydrogenase 2 and 3 (TDH3 and TDH2).

Table 4.24. The most (std > 55 nmol/gDWh) divergent proteins across all experiments and their flux variabilities as ranges (max - min flux) in nmol/gDWh.

Enzyme	**b2-ethanol**	**b8-ethanol**	**caffeine**	**coniferyl aldehyde**	**iron**	**nickel**	**phenyl ethanol**	**silver**	**reference**	**std**
GRX1	580.52	798.71	796.91	799.92	799.02	580.52	799.86	0	580.52	261.34
GRX2	580.52	580.52	622.01	624.36	623.65	580.52	622.86	0.08	580.52	201.73
TDH3	561.29	609.88	968.54	403.74	610.11	404.14	488.51	296.71	403.66	196.83
TDH2	702.15	760.74	965.83	607.13	742.34	584.58	649.07	305.09	586.70	177.49
HYR1	530.62	530.44	529.25	531.24	530.65	531.78	529.97	0	531.14	176.88
GPX1	507.66	507.49	506.35	508.26	507.69	508.74	507.04	0	508.16	169.23
TDH1	704.05	762.80	542.42	606.64	672.41	586.16	381.54	315.49	588.29	145.14
OLI1	498.12	447.32	256.85	462.14	503.37	480.46	668.93	177.89	471.85	143.72
TRX2	83.56	76.81	477.59	85.43	91.79	101.72	218.84	43.53	102.17	134.54
SOL4	239.80	324.14	363.99	365.53	377.35	156.46	351.92	18.15	164.11	126.02
ERG10	237.06	236.98	0.01	0.02	0.02	214.75	0.01	238.31	237.27	122.94
CYC7	3.07	6.00	345.15	72.74	10.38	1.24	15.25	0.07	0.55	112.84
AYR1	274.31	301.45	301.68	63.66	301.41	301.16	26.45	302.96	301.69	112.20
PGA3	280.31	280.22	279.59	280.28	280.33	280.16	0	281.75	280.59	93.47
SNO3	0	0	0	0	0	0	0	261.32	0	87.11
DAS2	259.51	224.29	96.71	211.09	191.35	276.99	126.60	363.82	279.13	81.82
HXK1	82.23	146.33	188.50	189.24	193.89	33.59	186.56	0.01	34.44	79.54
SNO2	0	0	0	0	0	0	0	237.53	0	79.18
URA6	153.11	148.21	113.37	133.15	123.92	175.63	76.48	349.32	179.93	77.32
YDC1	265.82	137.27	265.96	265.85	265.50	265.13	265.98	261.85	58.55	76.34
YOR283W	385.25	284.27	374.59	386.44	383.88	167.70	354.85	391.22	387.67	74.91
SOL3	173.27	160.34	102.97	127.45	116.01	161.25	62.84	307.74	165.03	68.26
CDC8	180.45	172.85	222.25	182.36	168.37	198.14	78.50	336.56	201.04	67.03
ALG13	222.59	233.87	342.51	266.53	268.81	198.78	292.49	113.05	199.23	65.73
GPX2	537.40	537.23	536.02	341.26	537.43	538.58	488.86	540.21	537.93	65.52
PDC5	157.75	161.29	17.12	185.99	76.60	181.67	48.37	169.86	183.92	65.47
GLK1	27.76	50.03	133.67	135.09	92.54	33.06	193.48	0	34.35	64.44
GPP2	204.17	215.88	261.44	299.76	230.49	212.94	355.61	147.19	219.99	60.29
ALD4	44.53	37.53	140.64	138.00	78.24	40.30	174.06	1.32	39.86	59.43
PYK2	91.09	88.38	195.97	183.93	169.19	55.15	108.01	38.43	56.16	59.39
ALG14	179.10	169.99	77.84	142.43	140.31	198.78	120.83	273.04	199.23	55.73
MHT1	107.29	91.30	170.34	216.78	183.08	108.03	41.78	171.04	108.32	55.14

Pathway-wise investigation is carried out using the results of flux variability analysis results. First, the reactions that have subsystem annotations are collected separately for each subsystem.

It must be noted that some reactions can exist in multiple subsystems, for example, phosphoenolpyruvate carboxykinase reaction exists in all gluconeogenesis, glycolysis, citrate cycle (TCA cycle), pyruvate metabolism, biosynthesis of secondary metabolites, antibiotics, and carbon metabolism subsystems. Secondly, the mean flux values for each subsystem for all models in FVA are collected. These flux values for the most common subsystems are plotted as stacked bars in Figure 4.27.

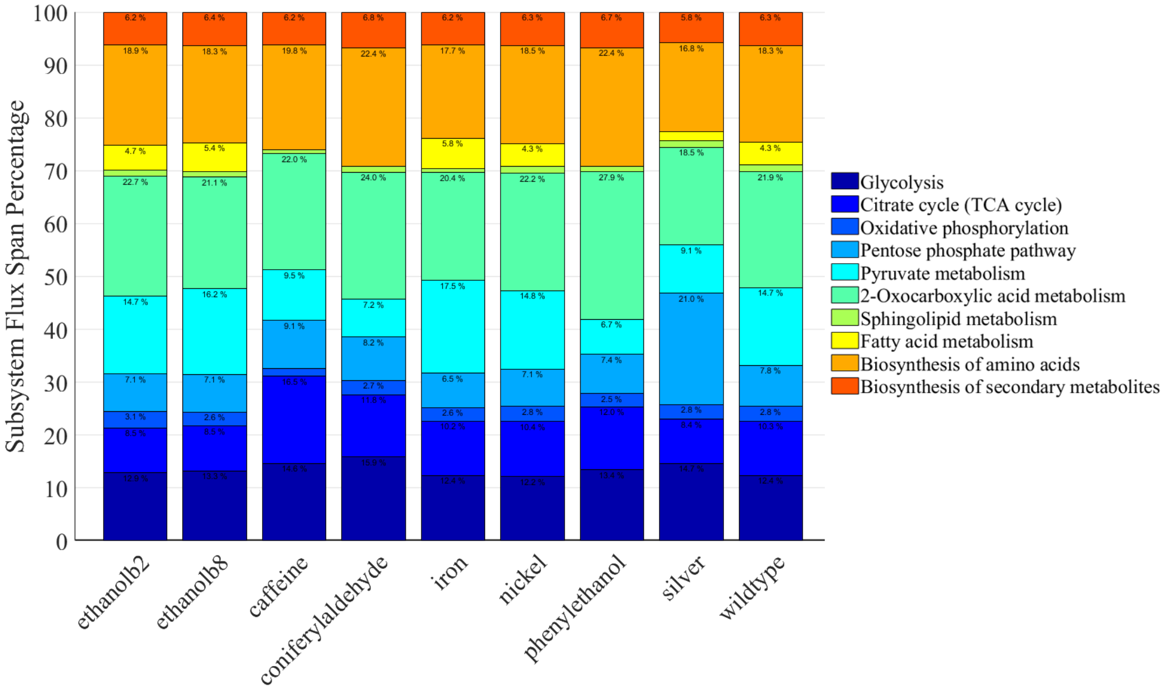


Figure 4.27. Mean carried fluxes through the subsystems are shown as stacked bars for strain-wise comparison. Sum of fluxes are normalized to a hundred percent to be able to distinguish flux distribution.

The most distinguished models are observed as caffeine, coniferyl aldehyde and phenylethanol strains as they do not carry fluxes through fatty acid metabolism reactions. Instead, the caffeine model shows higher activity in the TCA cycle reactions, compared to coniferyl aldehyde and phenylethanol strains where the activity increases in the 2-oxocarboxylic acid metabolism. Pentose phosphate pathway for the silver-resistant strain increased its flux 3.7 fold compared to reference strain. Secondly, fatty acid metabolism for the caffeine, coniferyl aldehyde and phenylethanol-resistant strains was almost deactivated.

4.6. Random Sampling Results

The random solutions for each metabolic model are obtained by maximizing for a random set of three reactions with random weights. Total of 10,000 solutions for each model are generated after the boundaries on growth rate as an objective function is set to minimum of 90% and maximum of 100% growth rate of the flux balance analysis results. The solution spaces for these reactions are plotted as histogram plots where the bin counts, N, are normalized so that the sum(N) is 1 for each analysis.

The most divergent reactions excluding diffusion, transport and exchange reactions from FBA results are plotted in Figure 4.28, the most divergent enzymes from FVA results are plotted in Figure 4.29, followed by the MOMA hits in Figure 4.33.

Ferrocytochrome-c:oxygen oxidoreductase reaction shows a great divergence across all models despite having the 0 flux on the caffeine-resistant model. These 0 flux on the caffeine model explains the higher fluxes through alcohol dehydrogenase reaction, meaning that the oxidative phosphorylation rates are lower in the caffeine-resistant model compared to ethanol producing pathways. Pyruvate decarboxylase reaction shows almost the pattern with the pyruvate kinase reaction indicating a chain reaction. Remember that although the models are fed with the same amount of carbon sources, small differences in their reaction fluxes comes from the differential expression analyses. That reminded, the small flux values on the 6-phosphogluconolactonase reaction may not mean a big difference biologically, however the different behavior of the silver-resistant model (having a wider range over other models) could mean the availability of the alternative pathways for the same biological objective.

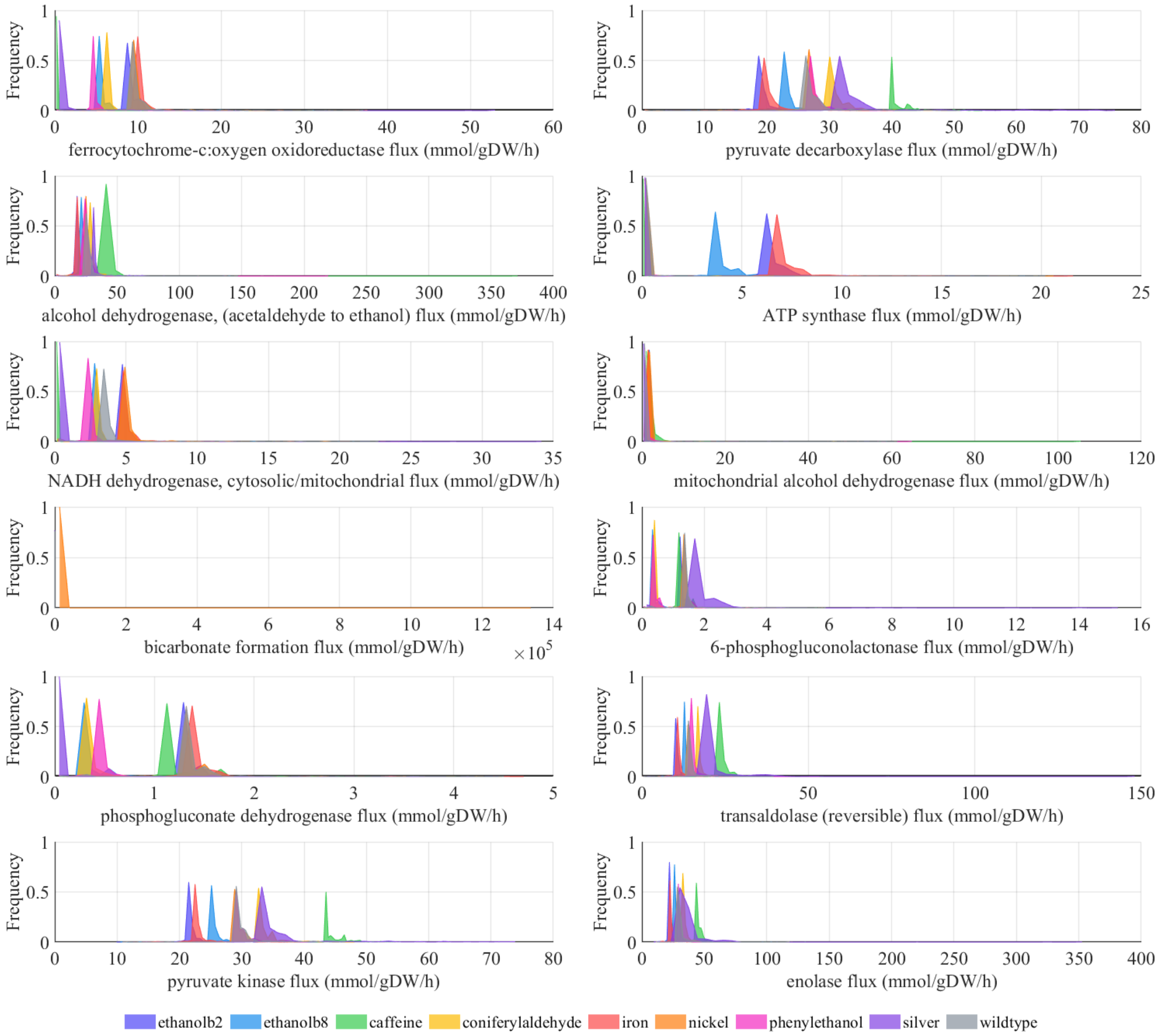


Figure 4.28. Flux distributions of the reactions that diverge most in flux balance analysis, obtained from the random sampling of solution spaces for each model shown as histogram plots.

Enzyme fluxes are numerically hard to distinguish in histogram plots because of the lower magnitude of the enzyme kinetics. The main target enzymes from the flux variability analysis are found as the glutaredoxin paralogs (GRX1 and GRX2) mostly did not show very different behaviors on the sampled solution spaces. Glyceraldehyde-3-phosphate dehydrogenase isozymes (TDH1, TDH2 and TDH3) on the other hand shows the different preferences between models.

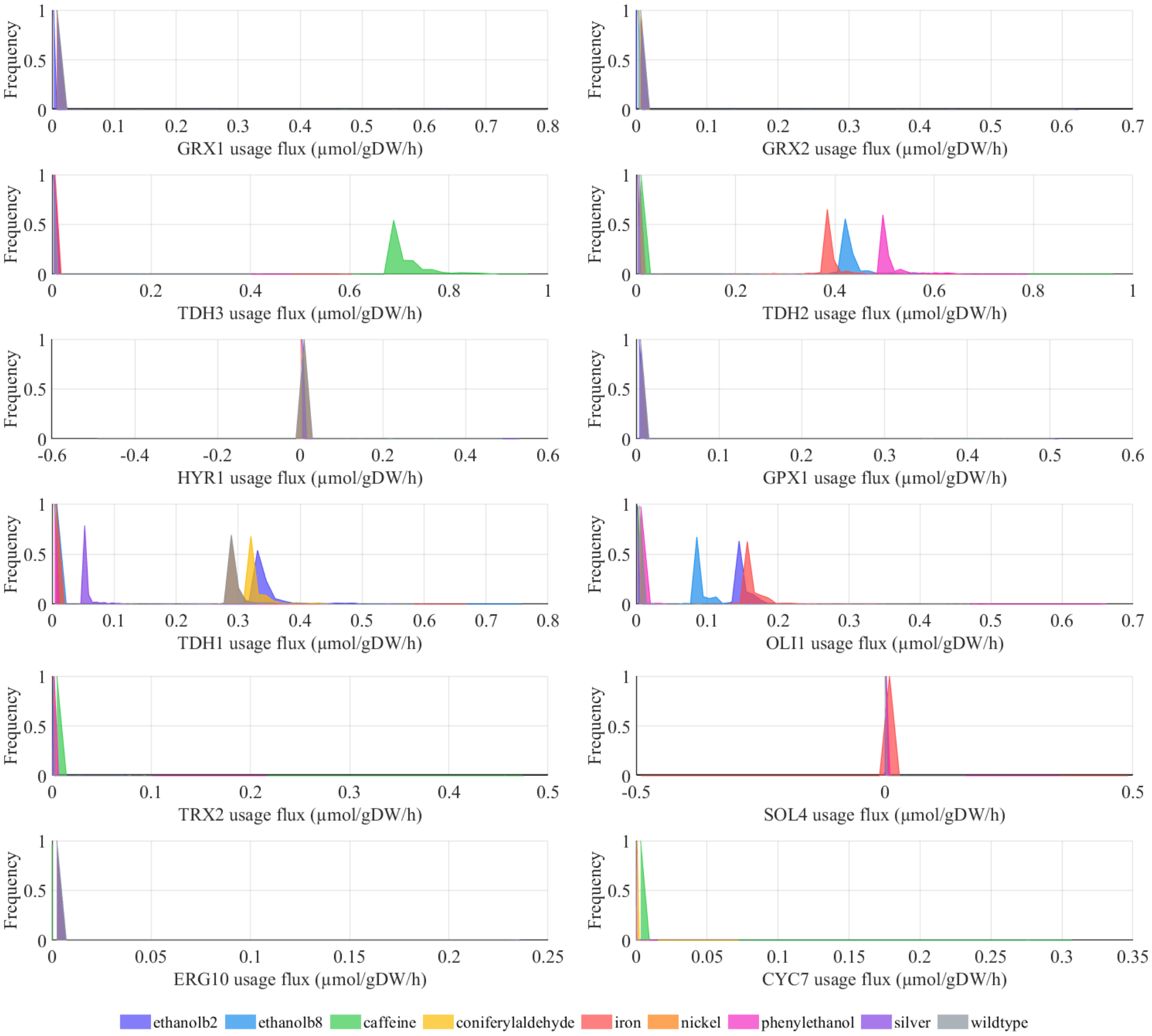


Figure 4.29. Flux distributions of the reactions that diverge most in flux variability analysis, obtained from the random sampling of solution spaces for each model shown as histogram plots.

In order to investigate enzyme relationships, correlation coefficients for each enzyme is calculated from the solution spaces for all models. Considering only the correlation values higher than R > 0.5, GRX enzymes are found to be positively correlated with the phospholipid hydroperoxide glutathione peroxidase (GPX) enzymes (Figure 4.30).

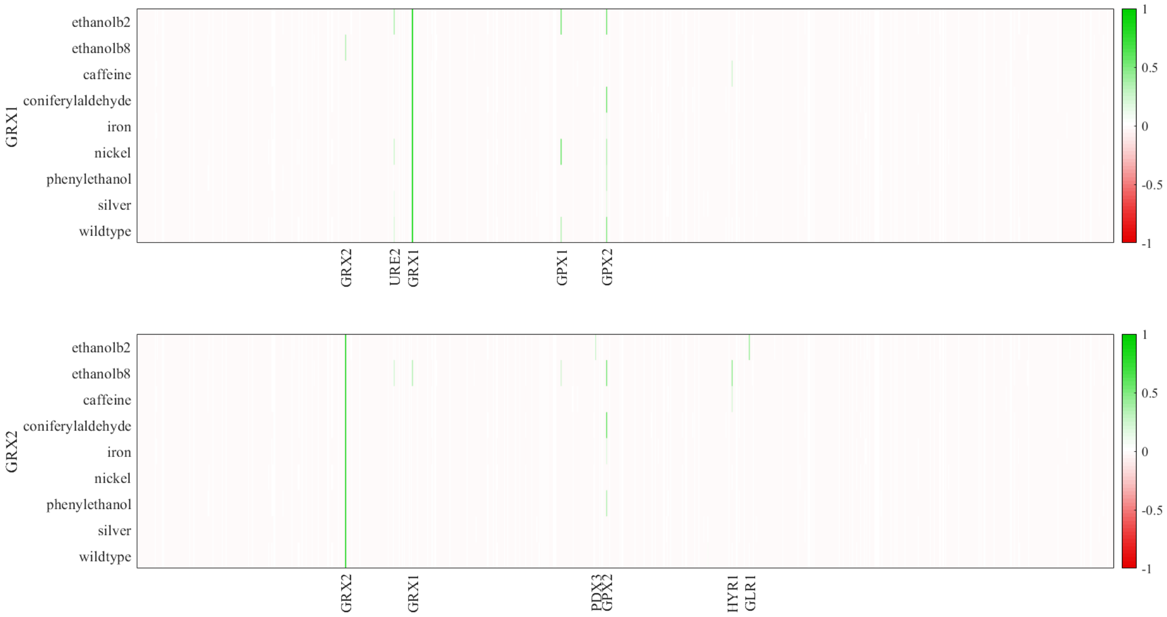


Figure 4.30. Correlations of GRX1 (top) and GRX2 (bottom) enzymes to other enzymes calculated from the random sampling of solution spaces. Only the name of enzymes with the correlation value c>0.3 are shown in the y-axes

TDH3 enzyme shows positive correlation only in the caffeine-resistant model to 4 enzymes: An enzyme responsible for the regulation of telomerase (CDC13), minor isoform of pyruvate decarboxylase (PDC5), subunit of phosphofructokinase (PFK2), and to a protein of unknown function (EMI2). TDH1 and TDH3 enzymes on the other hand show both positive and negative correlations to multiple enzymes (Figure 4.31).

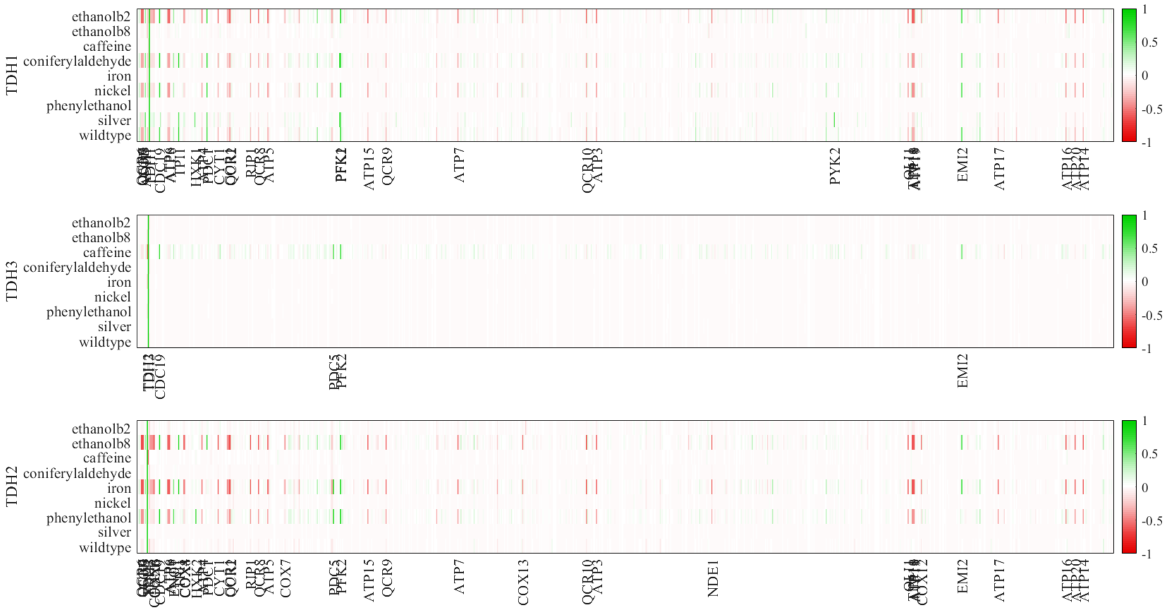


Figure 4.31. Correlations of TDH1 (top), TDH3 (middle) and TDH2 (bottom) enzymes to other enzymes calculated from the random sampling of solution spaces. Only the name of enzymes with the correlation value c>0.5 are shown in the y-axes.

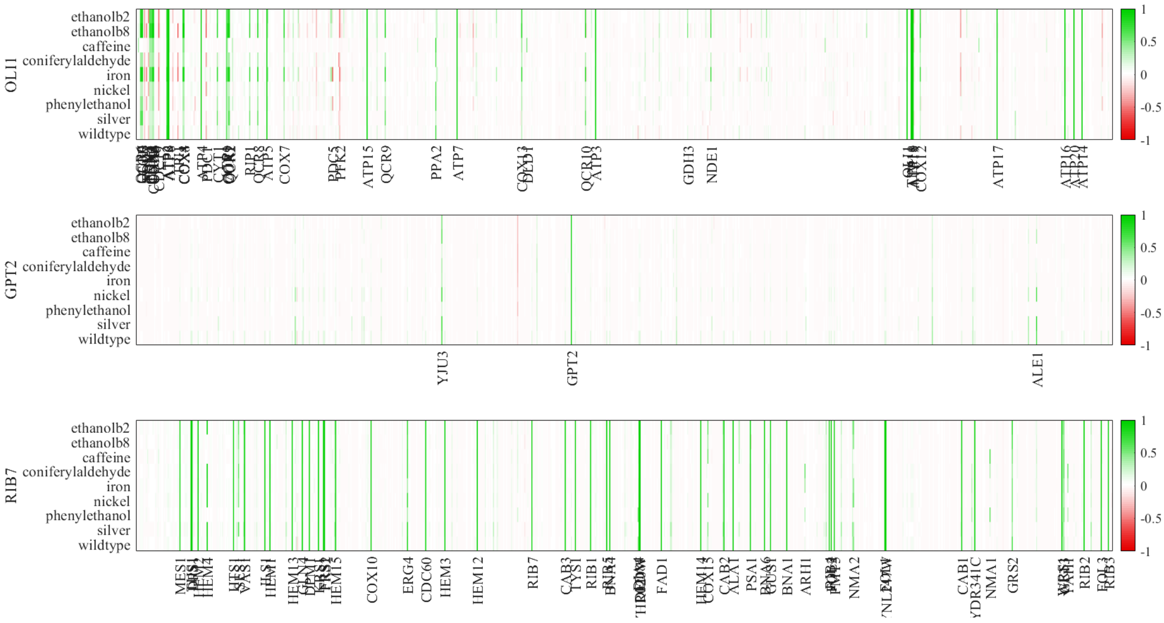


Figure 4.32. Correlations of OLI1 (top), GPT2 (middle) and RIB7 (bottom) enzymes to other enzymes calculated from the random sampling of solution spaces. Only the name of enzymes with the correlation value c>0.5 are shown in the y-axes.

When we further investigate the enzymatic correlations of OLI1 (F0 ATP synthase subunit c), GPT2 (glycerol-3-phosphate/dihydroxyacetone phosphate sn-1 acyltransferase), and RIB7 (Diamino-hydroxy-phoshoribosyl-amino-pyrimidine deaminase) to other enzymes. We see that RIB7 shows similar results across models with the exception of correlations with HEM4 enzyme (uroporphyrinogen III synthase) absent in b8-ethanol, caffeine and phenylethanol-resistant models.

Similar to HEM4 enzyme, COX15 (heme a synthase), ARH1 (mitochondrial oxidoreductase), NMA1 (nicotinic acid mononucleotide adenylyltransferase) and YAH1 (ferredoxin) enzymes must be investigated since they show different correlations across models although they are all involved in the heme biosynthesis pathways.

In the case of OLI1, both positive and negative correlations are observed, although the most distinct strain appears as caffeine-resistant model by showing less correlations. GPT2 enzyme positively correlates with YJU3 (monoglyceride lipase) and ALE1 (lysophospholipid acyltransferase) enzymes in the reference, silver-, and nickel-resistant models with lower correlation values in caffeine-resistant model.

Glyceraldehyde-3-phosphate dehydrogenase reaction ranks once again number 1 in a different analysis, MOMA. It can be seen that separation is clear in all models with a wider flux range in the silver-resistant model meaning that its flux value differs in multiple solutions. Followed by palmitoyl-CoA hydrolase, long chain fatty acid-CoA ligase and methylglyoxal synthase reactions are only in the Figure because they are only active in the silver-resistant model. However, since their flux ranges are too close to 0, they cannot be investigated as solid findings. Pyruvate kinase reaction shows the same pattern as in previous analyses. Interestingly, proton leak reaction (leakage of hydrogen atoms from cytoplasm to mitochondria) carries flux up to 20 mmol/gDWh, indicating that in the phenylethanol, coniferyl aldehyde and nickel-resistant models, more than available proton is required in the mitochondria. Lastly, 5'-nucleotidase reactions, guanosine kinase and pyridoxamine kinase reactions are only active for silver-resistant model, same as the previously mentioned reactions.

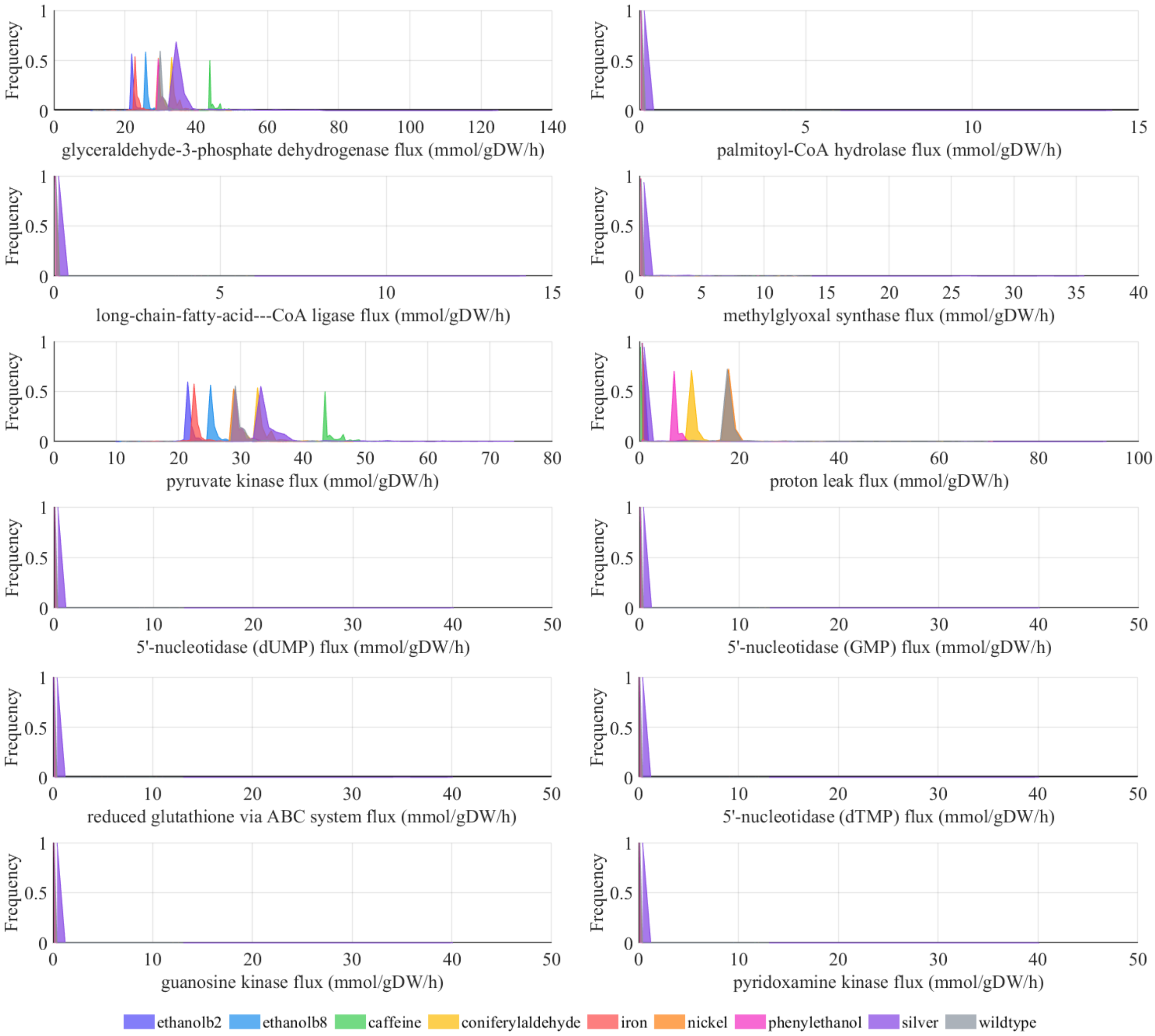


Figure 4.33. Flux distributions of the reactions that diverge the most in the minimization of metabolic adjustment analysis, obtained from the random sampling of solution spaces for each model shown as histogram plots.

Finally, mean values for each enzyme from solution spaces are calculated. Z-scores are obtained from the difference between this means in each strain model to the reference model fluxes, divided by the square root of the sum of standard deviations of this difference. Z-scores for enzymes are converted to p-values within gaussian distribution, and compared to the p-values obtained from the differential expression analysis. Discarding the non-significant enzymes in both cases, the following Figure 4.34 is constructed to compare regulations.

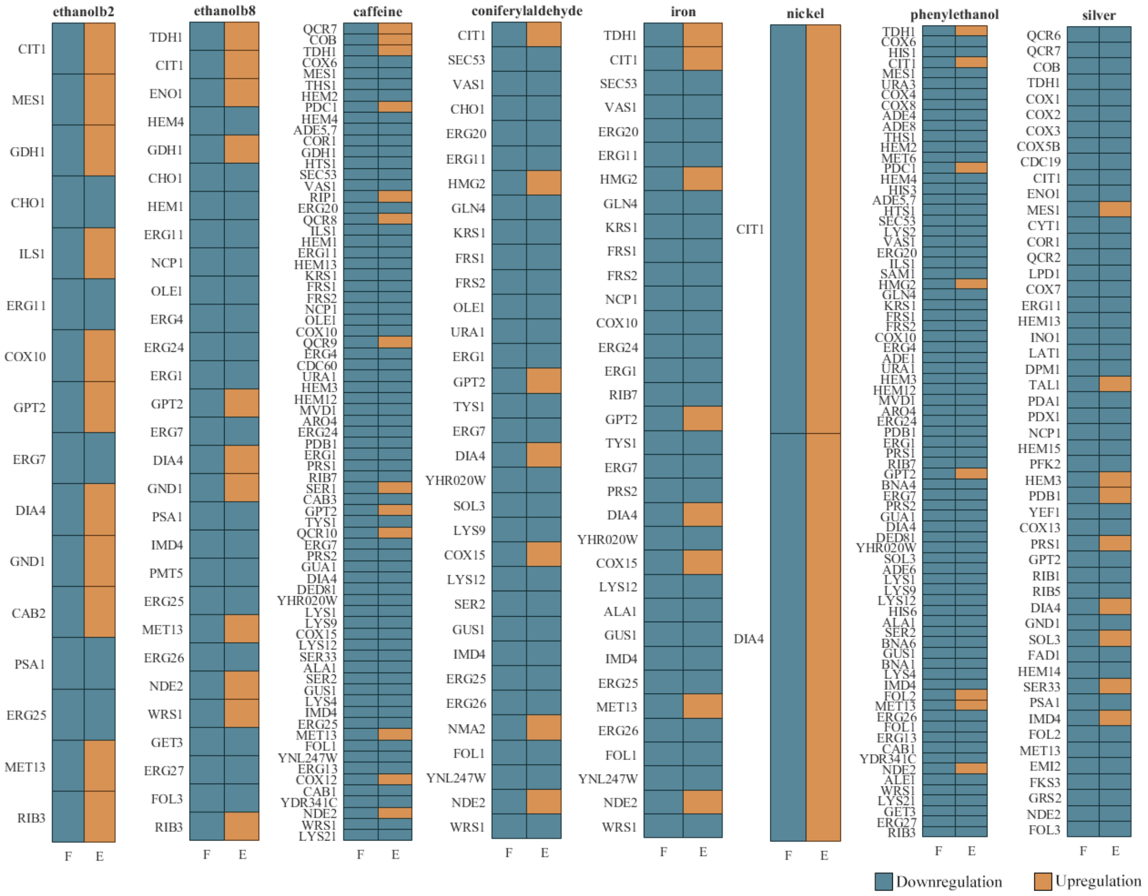


Figure 4.34. Regulation of the enzymes obtained from the random sampling of solution spaces (F) and differential expression analysis (E) in all models. Only significant ($p<0.05$) changes are plotted in both flux and expression changes.

Sampled points were also investigated for dual objective, the growth vs ATP synthase. In Figure 4.35, the edges where the points are denser show that the strains were able to grow at high rates without any ATP synthase activity (vertical line forming points), and the strains were able to grow at a specific rate while ATP synthase rate could vary (horizontal line forming points). Strains were able to reach their maximum growth rates without ATP synthase activity (vertical edges), and they were able to maintain the same growth rate for various flux values carried through ATP synthase (horizontal edges). Random sampling of the available solution space simply shows that multiple solution vectors can be calculated for a desired growth rate. The solutions scattered around the middle are more likely to reflect biological status of the evolved cells, therefore bi-level optimization could help investigate evolved strains.

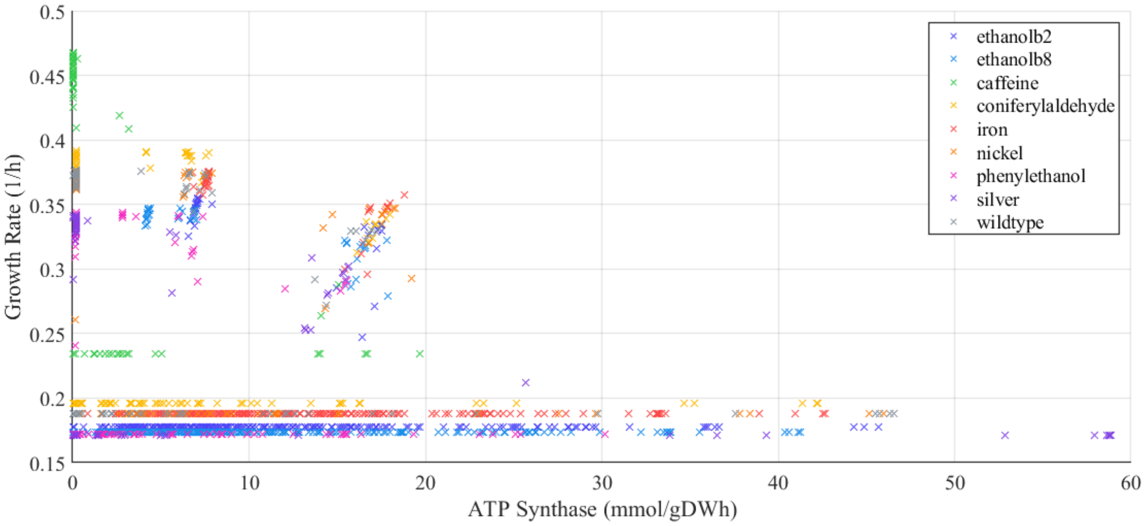


Figure 4.35. Scatter plot of the randomly sampled points for ATP synthase and the corresponding growth rates from the solution space.

4.7. Survivability Analysis: Knock-out Simulations

To calculate the essentiality of the reactions and the proteins on the growth, *in-silico* single deletion analyses are done on all the reactions for each model. Growth rate ratios between deletion applied models and the prior models are collected as percentages for each strain. The reactions that have the same value for each experiment are discarded from the results considering they do not differ after evolution. To be able to catch the most important reactions, only the reactions with the minimum of %95 growth rate decreases are reported here. One should remember that GECKO framework includes enzyme usages as separate reactions into the metabolic models, therefore the impacts of the deleted reactions are shown as heatmaps separately in Figure 4.36 for metabolic reactions and in Figure 4.37 protein reactions.

Although all the models were able to grow in single deletion of any enzymes on the reactions (with the minimum 45% lower growth rate from the non-deletion models), 2 enzymes (SAM1 and PMT5) are found essential for all models. Interestingly, MET2 (homoserine O-acetyltransferase), UGA1 (4-aminobutyrate transaminase), COX10 (protoheme IX farnesyltransferase), UGA2 (succinate-semialdehyde dehydrogenase), HEM14 (oxygen-dependent protoporphyrinogen oxidase), THI6 (bifunctional hydroxyethylthiazole kinase/thiamine-phosphate diphosphorylase) and MRI1 (S-methyl-5-thioribose-1-phosphate isomerase MRI1) enzymes were essential for all models except the caffeine-resistant model. On the other hand, the deletion of the PGK1 (phosphoglycerate kinase), PFK2 (6-phosphofructokinase subunit beta), TPI1 (triose-phosphate isomerase TPI1) and PGI1 (glucose-6-phosphate isomerase) enzymes decreases the caffeine-resistant model's growth ability higher than other models.

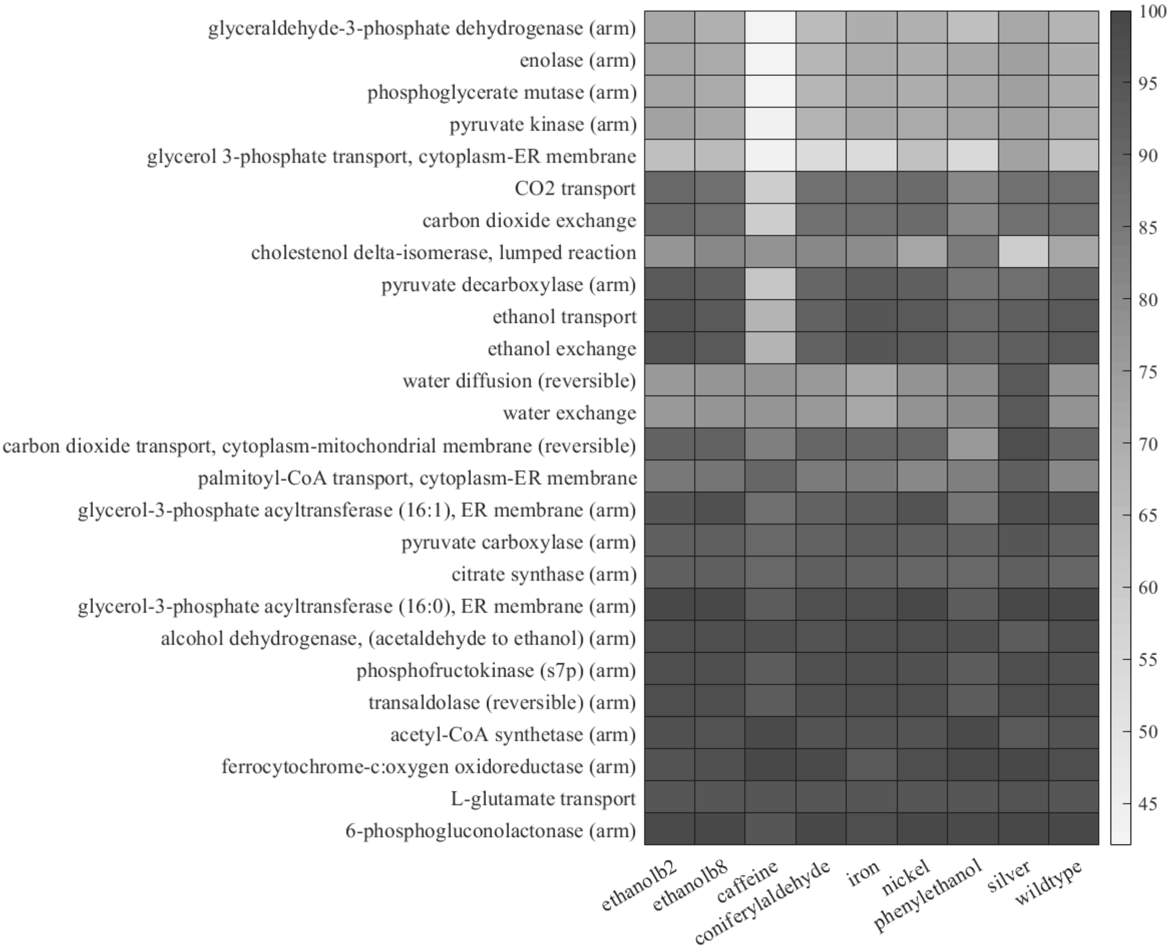


Figure 4.36. Heatmaps of the growth rate ratios between deletion strains for each experiment. The knocked-out reactions that cause a decrease higher than %95 are reported.

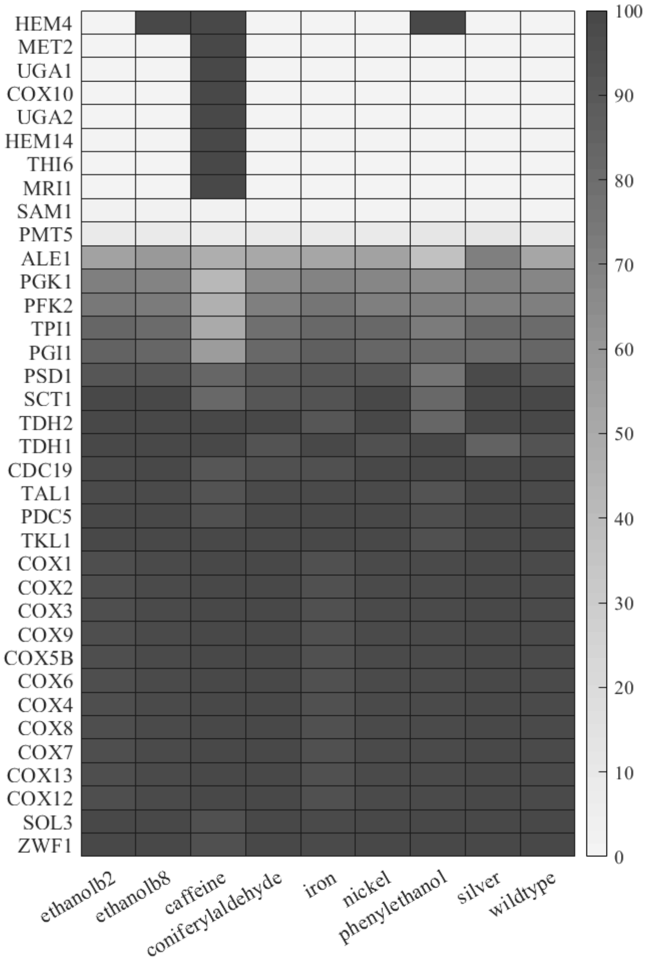


Figure 4.37. Heatmaps of the growth rate ratios between deletion strains for each experiment. The knocked-out proteins that cause a decrease higher than %95 are reported.

5. DISCUSSION

In this work, genome-scale metabolic modelling methods are used to analyze transcriptomics of evolved *S. cerevisiae* strains that have been obtained by in vivo evolutionary engineering strategies for different environmental conditions where the following substances are gradually increased in the media: Ethanol [69], caffeine [70], coniferyl aldehyde [71], iron [72], nickel [73], phenylethanol [74], and silver [75]. Overall workflow of the study is described Figure 5.1.

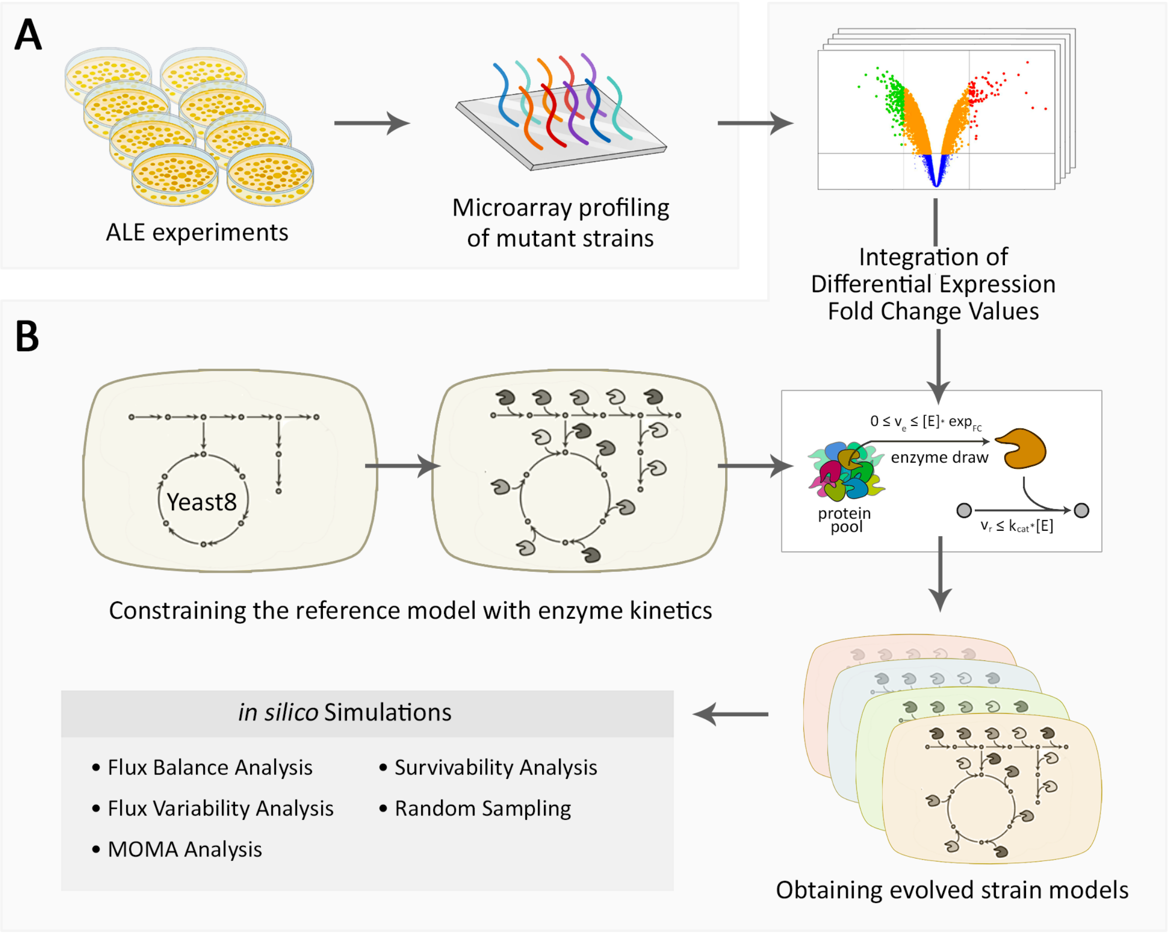


Figure 5.1. Overall workflow of the study: A) preprocessed microarray and experimental data on evolved strains are obtained from Cakar's Lab, B) model reconstruction and *in silico* analyses are carried out.

5.1. Flux Distribution Comparisons

FBA and robustness analyses where growth rates are maximized against varying levels of glucose uptake and oxygen uptake rates predicted non-linear robustness graphs. This non-linearity shows that the reconstructed metabolic models which are used in this study were able to simulate overflow metabolism without the *ad hoc* constraints.

By simulating the varying the carbon concentration in the media, a transition region that is known as Janusian region is predicted [119]. Using enzymatic constraints with GECKO methodology helps us to capture this region since it occurs when the cells are limited in nutrient and protein availability-wise. In the analyses, this region is observed for growth rates above 0.3 h^{-1} in all strains as the metabolic enzymes became saturated due to protein pool limitation.

Evolved strains produce biomass at different rates, with different exchange metabolite preferences due to their non-identical protein availability. Caffeine resistant strain reaches the highest growth rate amongst evolved strains, and the strain was able to maintain its energy without using the oxidative phosphorylation pathway and it is highly sensitive to oxygen levels in the media. Simulations show that the strain cannot grow at all if the oxygen uptake rate is higher than 25 mmol $(gDW)^{-1}$ h^{-1} and if enough carbon is available in the media to uptake, cells prefer to simulate hypoxia in order to produce biomass at higher rates. However, when the carbon source is limited, cells prefer to take more oxygen inside the cell. On the other hand, silver resistant strain which is the most distinguished evolved strain next to caffeine resistant strain, had the lowest maximum μ but it was able to produce biomass at higher rates under higher oxygen and glucose uptake rates compared to other strains.

Multiple *in silico* analyses showed that, although the predictions for the glucose uptake and ATP production rates of silver resistant strain is high, the strain does not carry fluxes through oxidative phosphorylation, in mitochondria. When total flux values through the reactions of coenzymes NAD, NADP, FAD and glutathione system are investigated, an outstanding difference on the glutathione metabolism activity is observed. It has previously shown in the literature that glutathione metabolism is highly connected with stress response mechanisms, and glutathione itself has a role in the maintenance of cellular structure integrity [120]. With that information, the change in the glutathione metabolism of silver resistant strain could be further investigated in stress response coupled with adaptive evolution. SEC53, a phosphomannomutase which localizes to cytoplasmic stress granules also a stress related response [121] and its activity is decreased notably in phenylethanol resistant strain.

Another metabolite, riboflavin, has a monofunctional role in FMN and FAD biosynthesis pathways [122]. Phenylethanol resistant strain showed decreased activity through riboflavin biosynthesis pathway by RIB7 (2,5-diamino- 6-(ribosylamino)-4(3H)-pyrimidinone 5'-phosphate reductase) despite its increased activities through TDH1 (glyceraldehyde-3-phosphate dehydrogenase) and OLI1 (F0 ATP synthase subunit c) in MOMA results, directs us to focus on energy utilization pathways.

OLI1, the main transmembrane subunit of F-type ATP synthases, and TIM11, subunit e of mitochondrial F1F0-ATPase take part in the mitochondrial enzyme F0F1-ATPase / ATP synthase synthesizes during oxidative phosphorylation [123,124]. Changes in the OLI1 and TIM11 activities in the mutant strains are important to investigate since the ATPase is an evolutionarily conserved enzyme complex [125]. COB (cytochrome b), a subunit of the mitochondrial respiratory chain complex III (transmembrane cytochrome bc1), catalyzes the quinol oxidation and cytochrome c reduction reactions, meanwhile generating a force to proton motive in ATP synthesis [126].

Caffeine, ethanol B8, and phenylethanol resistant strains have increased their activities compared to reference strains, and the silver resistant strain was the only strain which has decreased its flux for COB activity. As protein allocation decides which reactions to catalyze or which pathways to activate in cells, these differential findings are useful for analysis of energy metabolism in evolved strains.

A common change between all evolved strains of decreasing glycerol-3-phosphate or dihydroxyacetone phosphate sn-1 acyltransferase (GPT2) activity is observed, except the nickel resistant strain. However, since only a few expression data integrated in the reconstruction of nickel resistant strain, its exception could be ignored in terms of finding a commonality between all strains. GPT2 has a role in the phospholipid biosynthesis, where it catalyzes the acylation of glycerol-3-phosphate and dihydroxyacetone [127], and it is one of the main targets.

Two of five related enzymes of the 5-phospho-ribosyl-1(alpha)-pyrophosphate synthetases (PRSs), namely PRS1 and PRS2, were found as another important targets by MOMA as their flux values for each evolved strain differed. Although the flux values of the two enzymes' usages were equal for the reference strain; ethanol B8, caffeine, iron and phenylethanol resistant strains did not carry any fluxes through the enzyme PRS2 and only used the enzyme PRS1 for their activity (caffeine and phenylethanol resistant strains carried low fluxes through PRS1). On the contrary, only the silver resistant strain preferred to use PRS2, and carried zero flux through PRS1.

Phosphoribosyl pyrophosphate (PRPP), synthesized from PRPs, is a key compound in central carbon and nitrogen metabolism, as it has a role in purine, pyrimidine, and pyridine nucleotides synthesis pathways [128]. A previous study by Schneiter et. al. investigated the five genes of PRSs (PRS1–PRS5), and reported that the Δprs1 and Δprs3 mutants of *S. cerevisiae* are hypersensitive to caffeine and they could not survive when exposed to caffeine [129]. It was also found that PRPP plays a key role in cell wall integrity pathway and cell viability through stress signaling [130]. Here, the caffeine resistant strain carried zero flux through PRS2, and 7.10 fold decreased flux compared to the reference strain. Differential expression analysis also concluded that the caffeine stress causes downregulation on all PRS genes.

Acyltransferase of the glycerolipid biosynthesis pathway, namely SCT1, was inactive in the reference strain's FBA results. However, although the MOMA methodologically finds the solution vector for the evolved strain where the distance between the evolved strain to the reference strain's solution is minimum, SCT1 was activated in ethanol B2, ethanol B8, caffeine, coniferyl aldehyde, and iron resistant strains. It was previous shown in the literature that SCT1 deletion decreases the saturated fatty acid content in the cells, and its overexpression (which was the case for the ethanol B8, coniferyl aldehyde and phenylethanol resistant strains) cause a decrease in desaturation of fatty acids dramatically, and therefore affecting the phospholipid composition of the cells [131]. The followings are found as the most sensitive metabolites from the shadow prices (the derivatives of the objective function, i.e. growth, with respect to the exchange flux of the metabolite, see Table 4.10) in FBA: lanosterol, hexanoate, 14-demethyllanosterol, zymosterol, ergosta-5,7,24(28)-trien-3beta-ol, ergosterol, 4-aminobenzoate, 5-formyltetra hydrofolic acid, flavin mononucleotide (FMN), riboflavin, and biotin.

It is known that the acetyl-coenzyme A carboxylase is a biotin-dependent enzyme, and biotin is also involved in the generation of malonyl-CoA in fatty acid synthesis amongst its many other functionalities [132, 133]. Sensitivity of silver resistant strain to biotin availability has increased almost 2.5 fold whereas caffeine, nickel and phenylethanol resistant strains decreased their sensitivities down to 4 folds. Flavin mononucleotide (FMN) and its precursor riboflavin (vitamin B2) are the metabolites which evolved strains are most sensitive to their changes, and the values for both metabolites are almost the same because their reactions are formed as chain reactions in the metabolic network. FMN, produced from riboflavin, functions as a prosthetic group for several oxidoreductases, mainly NADH dehydrogenase [134]. The reactions for ergosterol, a sterol found in the cell membrane, and ergosta-5,7,22,24(28)-tetraen-3beta-ol, the product of ergosterol oxidation, are also chained reactions and therefore have same sensitivity values. Zymosterol, an intermediate in cholesterol biosynthesis, 4-aminobenzoate, an intermediate in folate biosynthesis, and 5-formyltetrahydrofolic acid, a folate coenzyme, are also found as important metabolites in sensitivity analysis.

The most divergent enzyme through all evolved model simulations, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is an enzyme involved in glycolysis and gluconeogenesis pathways and it is encoded by three genes TDH1, TDH2 and TDH3 (triosephosphate dehydrogenase; TDH) [135]. It catalyzes the conversion of glyceraldehyde-3-phosphate to 1,3-bisphospho-D-glycerate during glycolysis and the reverse reaction during gluconeogenesis,

$$\begin{aligned}\text{glyceraldehyde-3-phosphate} + NAD + P \leftrightarrow \\ \text{1,3-bisphospho-D-glycerate} + H^{+} + NADH\end{aligned} \tag{5.1}$$

where glyceraldehyde-3-phosphate is converted to 1,3-bisphospho-D-glycerate with the NAD and hydrogen.

In this study, it is found that carried flux through GAPDH is increased by 39% in caffeine, 13% in silver, and 11% in coniferyl aldehyde resistant strain. On the contrary, the carried flux is decreased by 31% in ethanolb2, 15% in ethanol B8, and 26% in iron resistant strain. It is also observed that each evolved model uses different isozyme for the catalytic activity. GAPDH differences are summarized in Figure 5.2.

For a long time, GAPDH was considered as a housekeeping gene for its constitutively expression in the cell, and commonly used in comparisons of expression data. In 2005, Barber et al. have reported different regulation mechanisms for GAPDH under specific conditions [136], and it is later found that the gene is upregulated under hypoxic stress [137]. Additional functions in other cellular processes have also been described such as being a chaperone protein in cellular iron homeostasis [138]. Interestingly, catalytically active TDH enzymes are found in both the cytoplasm and the cell wall.

GAPDH activity is controlled with oxygen availability, and evidence suggests that it may be the sensor of the cell in terms of oxidative stress [139]. In the simulations presented here, oxidative phosphorylation through ATP synthase was totally inactive (zero flux) for the caffeine resistant strain under unconstrained conditions, i.e., unlimited availability for uptake metabolites such as oxygen and glucose. Despite the inactivity, caffeine resistant strain was the highest ATP producer among evolved strains, and the ATP production was through pyruvate kinase and phosphoglycerate kinase activity (Table 4.11). Similar to caffeine resistant strain, silver resistant strain, too, showed increased activity on GAPDH and carried zero flux through ATP synthase in oxidative phosphorylation. In agreement with the results, ethanholb2, ethanol B8, and iron resistant strains showed decreased activity for GAPDH, and they were the only models that show ATP synthase activity.

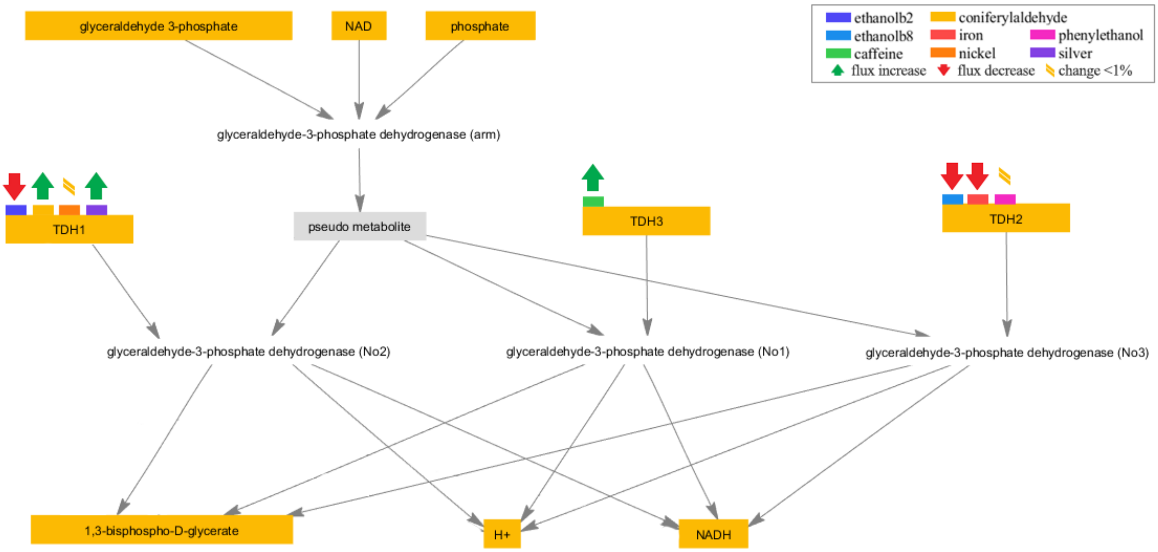


Figure 5.2. Map of glyceraldehyde-3-phosphate dehydrogenase catalyzed reaction as it used in simulations. Corresponding genes used for evolved strains are shown in colors, and flux changes compared to reference strain model simulations are indicated as increased or decreased with arrows.

The most second divergent enzyme through all evolved model simulations is found as pyruvate decarboxylase (PDC), a regulatory component of pyruvate dehydrogenase (PDH) complex that is responsible for conversion of pyruvate to acetyl coenzyme A, therefore making the complex a key metabolic point since it links glycolysis pathway to the TCA cycle. Pyruvate decarboxylase converts pyruvate into acetaldehyde with addition of one proton and release of one molecule of CO_2,

$$H^{+} + pyruvate \rightarrow \text{Acetaldehyde} + \text{carbondioxide} \tag{5.2}$$

and this reaction takes places in cytoplasm. Total of three isozymes are used for the pyruvate decarboxylase reaction in the metabolic models: The major of three pyruvate decarboxylases is PDC1 (isozyme 1, encoded by YLR044C), the second most abundant isozyme is PDC5 (isozyme 2, encoded by YLR134W), and lastly, the minor of three pyruvate decarboxylase is PDC6 (isozyme 3, encoded by YGR087C).

In simulations, models chose different isozymes to catalyze pyruvate decarboxylase reaction. While three of them (caffeine, iron, phenylethanol models) chose to use PDC5 isozyme, the remaining (ethanolb2, ethanol B8, coniferyl aldehyde, nickel, and silver models) chose PDC1; and none of the models carry fluxes through PDC6 as it can be seen from the Figure 5.3. It can be concluded that isozyme preference of models is in agreement with the abundances of isozymes. Additionally, it has been previously reported that the isozyme PDC6 is not used by yeast while the fermentation of glucose is active [140].

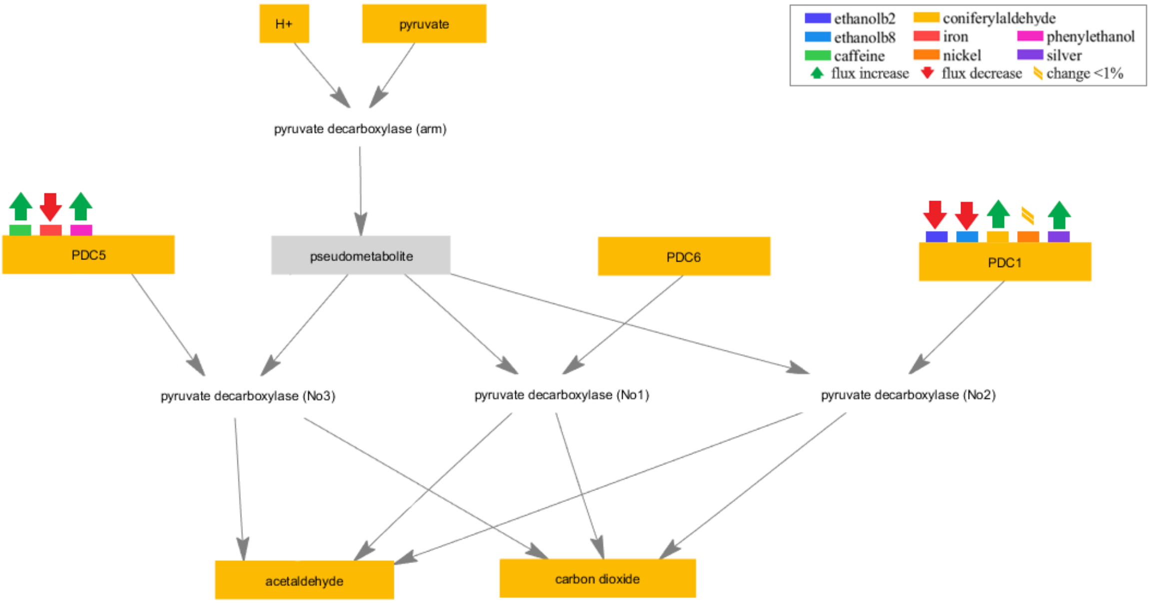


Figure 5.3. Map of pyruvate decarboxylase catalyzed reaction as it used in simulations. Corresponding genes used for evolved strains are shown in colors, and flux changes compared to reference strain model simulations are indicated as increased or decreased with arrows.

Another enzyme target found through FVA results is enolase. It is a phosphopyruvate hydratase; catalyzes conversion of 2-phosphoglycerate to phosphoenolpyruvate in forward direction during glycolysis, and the reverse direction during gluconeogenesis in yeast, the reaction

$$\text{2-phospho-D-glyceric acid} \leftrightarrow \text{H2O} + \text{phosphoenolpyruvate} \tag{5.3}$$

is catalyzed by enolase as shown [141]. This reaction in the metabolic models has five different genes annotated, defining five isozymes: ENO1 (YGR254W), ENO2 (YHR174W), ERR1 (YOR393W), ERR2 (YPL281C), and ERR3 (YMR323W). Last three of these genes are enolase-related repeats (namely ERRs) and they are identified by the sequence analysis [142]. It is known that the expression of enolase is decreased in response to glucose increase, while its protein abundance increases under replication stress

In FBA results, all models except ethanol B8 and silver resistant strains used the isozyme ENO1 to catalyze this reaction. ethanol resistant B8 and silver resistant strains, on the other hand, preferred to use ERR3 isozyme. These results were interesting because in the reconstruction steps of resistant models, the integrated fold change values of ENO1 was 0.43 (downregulation) for the silver, and 1.20 (upregulation) for the ethanol resistant B8 strains; while the values of ERR3 was 1 (no change) for the ethanol b8 model, and 0.12 (downregulation) for the silver resistant strain. In other terms, for the ethanol B8 and silver resistant strain simulations, the results were the opposite of expectations. Despite having an upregulation for the ENO1 isozyme, the ethanol resistant B8 strain preferred to use ERR3 isozyme. The similar results are also observed from the silver resistant strain, considering fold change values.

Kornblatt et al. previously showed that ERR3 region can encode a protein that has similar structure to yeast enolase, called Err3p, and it can catalyze the conversion of 2-phosphoglycerate to phosphoenolpyruvate [143]. It has also been showed that their kinetic activities are similar. Unfortunately, no further investigation on enolase-repeated regions is carried out. The preference behind the selection of enolase isozymes could arise from adaptation, therefore these regions must be investigated genetically.

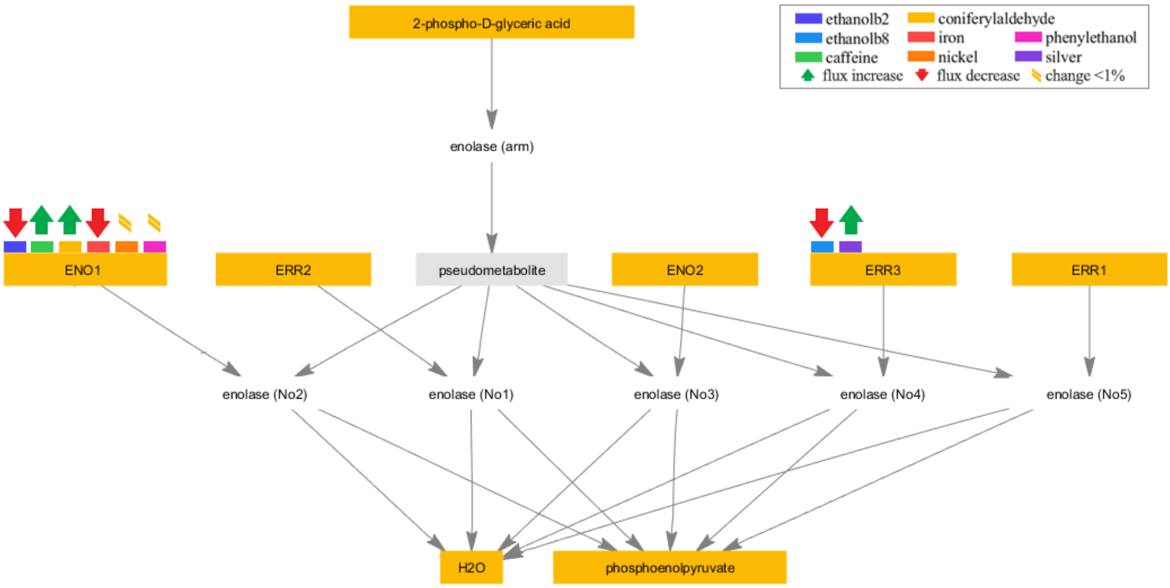


Figure 5.4. Map of enolase catalyzed reaction as it used in simulations. Corresponding genes used for evolved strains are shown in colors, and flux changes compared to reference strain model simulations are indicated as increased or decreased with arrows.

The phosphoglycerate mutase, another enzyme that caused distance between the simulations of evolved strains, is associated with the tetrameric GPM1 (YKL152C), and Q12040 (YOR283W) genes. It catalyzes the conversion of 3-phosphoglycerate to 2-phosphoglycerate in glycolysis pathway (step 8), and also the reverse reaction during gluconeogenesis. GPM1 and Q12040 have similarities in functioning as phosphoglycerate mutases, however Q12040 is a phosphatase with broader substrate specificity [144]. Studies on animals have shown that phosphoglycerate mutase activity is sensitive ionic concentration, such as high concentration of salts inhibit its mutase activity [145].

The fold change values of the gene GPM1 from differential analysis indicated an upregulation for ethanol-B8 and phenylethanol resistant, and a downregulation for silver resistant strain. No differential change is detected in the remaining resistant strains. For the enzyme Q12040, most of the strains showed downregulation, except an upregulation in silver resistant strain, and no change in ethanol-B2 and nickel resistant strains. In the FBA results, however, it is observed that all the models carry fluxes through the enzyme Q12040, while only the silver resistant strain uses the enzyme GPM1. These oppositions put forth the importance of the integrity of metabolic systems, and promotes to investigate systems as a whole instead of considering single values.

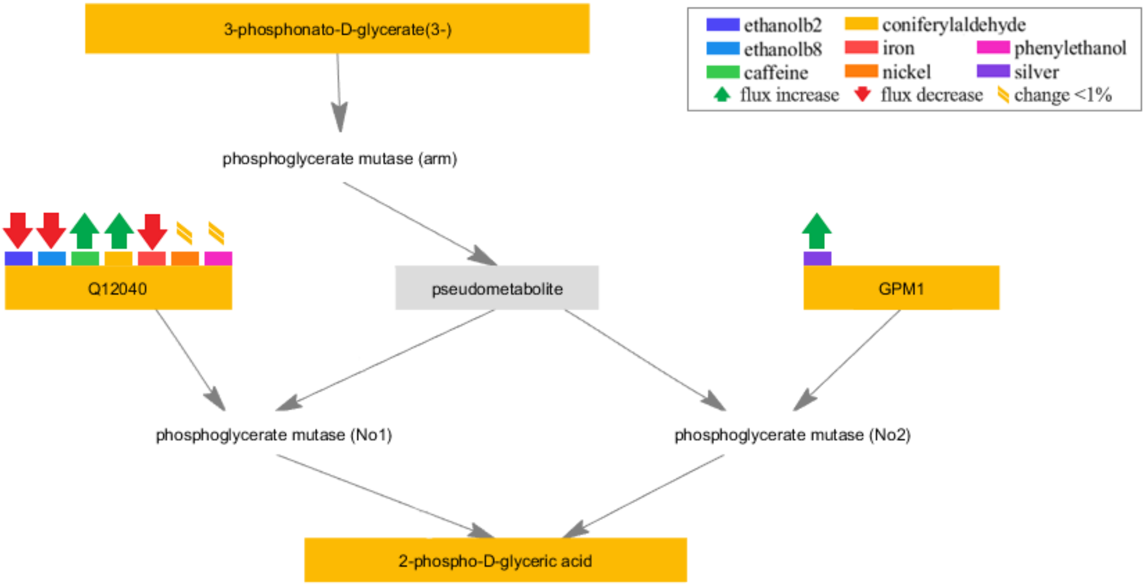


Figure 5.5. Map of phosphoglycerate mutase catalyzed reaction as it used in simulations. Corresponding genes used for evolved strains are shown in colors, and flux changes compared to reference strain model simulations are indicated as increased or decreased with arrows.

5.2. Strain Survivability

In-silico knock-out simulations showed that the single deletions of the enzymes L-homoserine-O-acetyltransferase (MET2), heme A:farnesyltransferase (COX10), gamma-aminobutyrate transaminase or 4-aminobutyrate aminotransferase (UGA1), succinate semialdehyde dehydrogenase (UGA2), protoporphyrinogen oxidase (HEM14), thiamine-phosphate diphosphorylase and hydroxyethylthiazole kinase (THI6), and 5’-methyl thioribose-1-phosphate isomerase (MRI1) are essential for the growth of all evolved strains except caffeine resistant strain which is totally unaffected by the deletions of mentioned enzymes. From those, THI6 that is required for thiamine biosynthesis is already reported essential in the literature [146], however the caffeine resistant strain was able to grow at even higher rates without it.

Contrarily, the caffeine resistant strain decreased its growth rate by half when the enzymes 3-phosphoglycerate kinase (PGK1), β-subunit of heterooctameric phosphofructokinase (PFK2), triose phosphate isomerase (TPI1), and phosphoglucose isomerase (PGI1) are blocked (knocked-out), whereas the other strains could survive without getting affected much.

Another contrast is observed when the uroporphyrinogen-III synthase (HEM4) is deleted. Uroporphyrinogen III synthase catalyzes the conversion of preuroporphyrinogen to uroporphyrinogen-III in the fourth step in heme biosynthesis [147]. Although it is not reported as essential in literature, its deficiency can result in the diseases in the human homolog [148]. In survivability analysis simulations, ethanol B8, caffeine, and phenylethanol resistant strains did not affected from its deletion, however other strains could not grow at all.

Single deletions showed that all the evolved strains including the reference strain could not produce biomass at high rates when S-adenosylmethionine synthetase (SAM1) and protein O-mannosyltransferase (PMT5). SAM1 catalyzes the transfer of adenosyl group from ATP to L-methionine, forming S-adenosyl-L-methionine [149], and PMT5 catalyzes the release of mannose residues from dolichyl D-mannosyl phosphate in the endoplasmic reticulum [150]. Although these enzymes are not found as essential, biomass production occurred despite the low rates, they are found as highly required by all strains therefore forming a common behavior across evolved strains.

5.3. Simulations vs. Experimental Results

Turanlı-Yıldız et al., obtained two evolved clones that could tolerate up to 12% (v/v) ethanol, namely B2 and B8 strains, under increasing ethanol levels [151]. Apart from the important findings on triggered diploidization during adaptation, their transcriptome analyses revealed that the most enriched genes were related to carbohydrates storage metabolism, however, only B2 strain alone exhibited a higher glycogen accumulation. They have also found that the abundances of mitochondrial proteins were decreased in B8 strain, suggesting a reduced respiratory activity. On top of that, findings on higher abundances of ribosomal proteins, amino acid metabolism, and glycolysis supported the idea of a higher fermentation levels.

Simulations show that, under the same environmental conditions provided, B2 strain is able to grow at slightly higher rates compared to B8 strain, 0.347 h^{-1} and 0.337 h^{-1} respectively, while B8 strain is able to produce ethanol at a higher rate 10.667 mmol/gDWh^{-1} compared to 9.934 mmol/gDWh^{-1} for B2 strain. The main difference observed in simulations between B2 and B8 strains is the regulation of citrate in the TCA cycle, specifically the transport reaction

$$citrate_c + isocitrate_m \xleftrightarrow[\text{Rev: B8 (0.319 mmol/gDWh}^{-1}\text{)}]{\text{Fwd: B2 (0.218 mmol/gDWh}^{-1}\text{)}} citrate_m + isocitrate_c \tag{5.4}$$

where the indicators c is for the cytoplasmic, and m is for the mitochondrial metabolites. Transportation of citrate from mitochondria to cytoplasm in B8 strain agrees with the experimental findings on the idea of reduced respiratory activity, and this idea is supported with the ferrocytochrome-c:oxygen oxidoreductase (oxidative phosphorilation) fluxes where B2 strain has higher flux rate 18.867 mmol/gDWh^{-1}, compared to 17.439 mmol/gDWh^{-1} on B8 strain.

Although there is not a clear indicator on the glycogen synthase on simulations, the reason behind a slightly higher growth rate for B2 strain could arise from the carbohydrate pseudoreaction

$$\begin{aligned} 0.74851 \text{ (1-3)-beta-D-glucan} + 0.25009 \text{ (1-6)-beta-D-glucan} + \\ 0.36141 \text{ glycogen} + 0.71094 \text{ mannan} + 0.13828 \text{ trehalose} \rightarrow \text{carbohydrate} \end{aligned} \tag{5.5}$$

required for biomass. However, it must be noted that only 0.003 mmol/gDWh^{-1} difference is observed between B2 and B8 clones for the glycogen (starch) synthase reaction where UDP-D-glucose is converted into glycogen.

Sürmeli et al., obtained yeast populations that can survive at high caffeine levels [152]. Contrary to expectations from literature, their evolved strain, Caf905-2, did not show any inhibitory effects on the growth during stress selection. Transcriptome analysis on the obtained evolved strain, showed that Cytochrome c isoform 2 (CYC7) was the most upregulated gene. It is known that CYC7 is an electron carrier of the mitochondrial intermembrane space, and it is expressed under hypoxic conditions. In our simulations, we were able to catch the over-expression of CYC7 on flux variability analysis (Table 4.24). As mentioned before, caffeine resistant strain reaches to highest growth rate among evolved strains, and it was able to maintain energy without oxidative phosphorylation.

As it can be seen from the phenotype phase planes in Figure 4.15, caffeine resistant strain is highly sensitive oxygen availability and cannot grow if the oxygen uptake rate is higher than 25 mmol/gDWh^{-1}. Flux balance analysis shows that caffeine resistant strain reaches its maximum available growth rate of 0.468 h^{-1} when the oxygen uptake rate is 0.235 mmol/gDWh^{-1}. In other words, the model chooses not to take oxygen from outside if the carbon supply is unlimited in the media, it simulates hypoxic conditions for best outcome. However, if the glucose uptake is forced to 10 mmol/gDWh^{-1}, model takes more oxygen at the rate 4.757 mmol/gDWh^{-1}.

In their work, it has also been reported the induction of genes involved in glycogen and trehalose metabolism under caffeine stress. However, since the genome-scale modelling does not simulate stress conditions (i.e., no caffeine molecule is provided into the defined medium), this finding is not observed in the simulations, suggesting that this change could be regulated on the metabolic level, not on the transcriptomics level. Additionally, upregulation of SNQ2 was also not captured in the simulations, due to lack of its gene association in the Yeast8 model.

Transcriptomic changes in a coniferyl aldehyde resistant yeast population, BH13, is reported by Hacısalihoğlu et al., and differential regulations after adaptive evolution on the NAD(P)-dependent aldehyde dehydrogenases are revealed. They reported that all members of aldehyde dehydrogenases (ALD) except for ALD5 were upregulated, and ALD upregulation were previously observed in the literature [153]. Additionally, oxidoreductase enzymes such as BDH2, YPL113C, YJR096W; and putative aryl-alcohol dehydrogenases such as YPL088W and AAD15 were reported as upregulated in coniferyl aldehyde resistant strain. Here, in flux variability analysis, ALD4 upregulation is observed in coniferyl aldehyde resistant strain, accompanied with the caffeine and phenylethanol resistant strains.

In the coniferyl aldehyde resistant strain, Hacısalihoğlu et al. reported that the glucose uptake and metabolism were enhanced even without the coniferyl aldehyde stress in the media [154]. This observation is explained by the upregulation of the genes encoding hexose transporters, and the enzymes involved in glycolysis, namely HXK1, GLK1, TDH1, GPM2, ERR1, and PYK2. Increased glucose uptake compared to wild-type model, and high flux range on flux variability analysis for HXK1, GLK1, and PYK2 in coniferyl aldehyde resistant strain confirms *in-silico* simulations to the experimental findings.

Similar to the results of ADL family, GLK1 enzyme has a higher flux range in caffeine and phenylethanol resistant strains next to coniferyl aldehyde resistant strain. Interestingly, for the flux range of PYK2, iron resistant strain accompanies coniferyl aldehyde and caffeine resistant strains instead of phenylethanol resistant strain.

Iron resistant *S. cerevisiae* mutant, M8FE, is obtained by Balaban et al with cross resistance feature to other metals [72], and their findings suggested that the resistance to the metals might be related to the downregulation of PHO84 gene, encoding a high-affinity inorganic phosphate transporter and also a low-affinity manganese transporter. Unfortunately, phosphate uptake rates show no significant difference on *in-silico* simulations. In the experiments, it is assumed that high amounts of iron would lead to oxidative stress that damage the cells, however, ROS amounts of the evolved M8FE strain were lower compared to reference strain. This phenomenon explained with the microarray results, where the upregulation on the oxidative stress response genes were observed. This upregulation is assumed to achieved in order to reduce oxidative levels in the cell. In the simulations, iron resistant strain was able to grow at the same rate with the reference strain model when the maximum glucose uptake is allowed (Table 4.4). However, when we consider ATP production, despite having high glucose uptake and growth rates as reference strain model, iron resistant strain has much lower fluxes through ATP producing reactions compared to other models (Figure 4.22). Although there is no *in-silico* confirmation, Balaban et al. suggests that the iron resistant strain prefers to save its energy as trehalose in the cell.

Their study also reported upregulations on the glycogen phosphorylase gene, GPH1; and phosphoglucomutase, PGM2. Here, flux variability analyses showed a high range for all evolved strains for the enzyme PGM2, catalyst of the conversion from glucose-1-phosphate to glucose-6-phosphate, except in the silver resistant strain. That being said, the upregulation in the PGM2 might possibly be related to DNA replication stress, and not a specific upregulation for the metal stress.

Küçükgöze et al. obtained nickel resistant strain of *S. cerevisiae*, called M9, with applications of ethyl methane sulfonate mutagenesis, and pulse stress of $NiCl_2$ in the media [73]. Their results show upregulations on metal homeostasis, especially iron metabolism, and stress response genes upon nickel exposure. Since we were able to integrate only 28 differentially expressed genes to build nickel resistant strain, its simulation results were similar to reference strain model simulations. Despite this, the nickel resistant strain showed 2.4-fold increase on the flux for MET2 gene which codes for homoserine O-acetyltransferase, followed by 1.5-fold increase on PFK2 (ATP-dependent 6-phosphofructokinase subunit beta), PDX1 (mitochondrial pyruvate dehydrogenase complex protein X component), and ARG7 (mitochondrial arginine biosynthesis bifunctional protein ArgJ).

Lastly, the silver resistant strain 2E were obtained by Terzioğlu et al. by using evolutionary engineering [75]. They showed the strain 2E has cross-resistance to copper, as both transition metals are similar, and copper-binding protein encoding genes are upregulated (such as CUP1-1, CUP1-2, and CTR3). Interestingly, they showed that their silver resistant strain has different mitochondrial activities compared to previous studies by Horstmann et al. [155], and Marquez et al. [156] in terms of mitochondrial activities. It has been shown by previous studies that silver stress causes decreased cellular respiration, and increased oxidative stress. However, Terzioğlu et al. reported an upregulation on set of genes encoding proteins for electron transfer system (such as CYC7, COX7, NDI1, and SDH1) indicating a higher activity on aerobic respiration. Additionally, GPX1, GRX1, and GRX2 genes were also reported as upregulated.

Although differential expressions of the study of Terzioğlu et al. were used to obtain silver resistant strain, our analysis results were in agreement with the previous findings. Simulations show zero flux through oxidative phosphorylation in mitochondria, despite having the second highest glucose uptake and ATP production flux rates. Genes that are upregulated in the differential analysis, namely GPX1, GRX1 and GRX2, were not used by the silver resistant strain, i.e., carried zero fluxes.

6. CONCLUSION

In this study, constraint-based genome-scale metabolic modelling techniques are used to analyze multiple mutant strains that have gained resistance against different stress conditions by adaptive laboratory evolution. Total of 9 different metabolic models for the evolved ethanol-, caffeine-, coniferylaldehyde-, iron-, nickel-, phenylethanol-, and silver-resistant strains, and the reference strain are reconstructed from the consensus *Saccharomyces cerevisiae* model Yeast8. In the reconstruction process, reactions are enzymatically constrained, and the differential expressional profiles are integrated to gather more biologically relevant results. FBA, FVA, MOMA, PhPP, robustness, survivability and random sampling analyses are conducted in computational environment for each mutant strain, and the results are compared both individually and comparatively.

Despite the lack of metabolomics and fluxomics data, findings of the study contribute to the research in adaptation mechanism by providing a more systematic view compared to a sole differentially expression analysis. Even with the low coverage of transcriptional integrations, multiple targets of common or divergent behaviors of mutant strains are reported and discussed.

One of the main findings of the study is that the protein allocation is highly intertwined with the energy metabolism. As the kinetic parameters integrated into metabolic models allow to follow each flux distribution without needing any *ad hoc* constraints, the only difference making part of each model was the protein allocation which models decide to use preferred enzymes for the best outcome. As models used their proteins the most efficient way in terms of to fulfill objective (such as biomass production), the converging difference became the different behaviors in the energy related reactions, especially in the coenzyme (NAD, NADP, FAD) utilizing reactions, and the glutathione metabolism. Secondly, sensitivities to the exchange metabolites differed in each strain, such as oxygen, biotin, riboflavin and flavin mononucleotide.

Most of the metabolites that are found as target were closely related to stress response of yeast, therefore commonalities within strains were treated as a result of the stress response, and divergent findings were considered as adaptive effects. Lastly, targets that affect the composition of the lipids, and therefore cell wall formation, was another noteworthy finding of the study since the connection between the two were previously reported multiple times.

Although our methods tried to capture the adaptive metabolism in a more systematic broader way compared to the differential gene expression analysis; the results were still too abundant to investigate one by one. In the near future, with more precise data made available in ALE experiments, the essence of the adaptive metabolism could easily be captured using metabolic modelling techniques.

REFERENCES

1. Futuyma, D. J. and T. R. Meagher, “Evolution, science and society: evolutionary biology and the national research agenda.”, *California Journal of Science Education*, Vol. 1, No. 2, pp. 19–32, 2001.

2. Hird, S. M., “Evolutionary biology needs wild microbiomes”, *Frontiers in Microbiology*, Vol. 8, p. 725, 2017.

3. Nielsen, J. and J. D. Keasling, “Engineering cellular metabolism”, *Cell*, Vol. 164, No. 6, pp. 1185–1197, 2016.

4. Garland, T. and M. R. Rose, *Experimental evolution: concepts, methods, and applications of selection experiments*, University of California Press Berkeley, CA, 2009.

5. Dragosits, M. and D. Mattanovich, “Adaptive laboratory evolution–principles and applications for biotechnology”, *Microbial Cell Factories*, Vol. 12, No. 1, p. 64, 2013.

6. McDonald, M. J., “Microbial Experimental Evolution–a proving ground for evolutionary theory and a tool for discovery”, *EMBO Reports*, Vol. 20, No. 8, 2019.

7. Winkler, J., L. H. Reyes and K. C. Kao, “Adaptive laboratory evolution for strain engineering”, *Systems Metabolic Engineering*, pp. 211–222, Springer, 2013.

8. Conrad, T. M., N. E. Lewis and B. Ø. Palsson, “Microbial laboratory evolution in the era of genome-scale science”, *Molecular Systems Biology*, Vol. 7, No. 1, 2011.

9. Soyer, O. S. and M. A. O’Malley, “Evolutionary systems biology: what it is and why it matters”, *BioEssays*, Vol. 35, No. 8, pp. 696–705, 2013.

10. Long, C. P. and M. R. Antoniewicz, "How adaptive evolution reshapes metabolism to improve fitness: recent advances and future outlook", *Current Opinion in Chemical Engineering*, Vol. 22, pp. 209–215, 2018.

11. Shepelin, D., A. S. L. Hansen, R. Lennen, H. Luo and M. J. Herrgård, "Selecting the best: evolutionary engineering of chemical production in microbes", *Genes*, Vol. 9, No. 5, p. 249, 2018.

12. Palsson, B., "Adaptive laboratory evolution this tool for studying molecular mechanisms and evolutionary dynamics follows bacteria as they adapt to environmental stress", *Microbe*, Vol. 6, No. 2, p. 69, 2011.

13. Phaneuf, P. V., D. Gosting, B. O. Palsson and A. M. Feist, "ALEdb 1.0: a database of mutations from adaptive laboratory evolution experimentation", *Nucleic Acids Research*, Vol. 47, No. D1, pp. D1164–D1171, 2019.

14. Wong, B. G., C. P. Mancuso, S. Kiriakov, C. J. Bashor and A. S. Khalil, "Precise, automated control of conditions for high-throughput growth of yeast and bacteria with eVOLVER", *Nature Biotechnology*, Vol. 36, No. 7, pp. 614–623, 2018.

15. LaCroix, R. A., B. O. Palsson and A. M. Feist, "A model for designing adaptive laboratory evolution experiments", *Applied and Environmental Microbiology*, Vol. 83, No. 8, 2017.

16. Kitano, H., "Systems biology: a brief overview", *Science*, Vol. 295, No. 5560, pp. 1662–1664, 2002.

17. Bellouquid, A. and M. Delitala, *Mathematical modeling of complex biological systems*, Springer, 2006.

18. Kremling, A., *Systems biology: mathematical modeling and model analysis*, Chapman and Hall/CRC, 2013.

19. Bruggeman, F. J. and H. V. Westerhoff, “The nature of systems biology”, *Trends in Microbiology*, Vol. 15, No. 1, pp. 45–50, 2007.

20. Shahzad, K. and J. J Loor, “Application of top-down and bottom-up systems approaches in ruminant physiology and metabolism”, *Current Genomics*, Vol. 13, No. 5, pp. 379–394, 2012.

21. Thiele, I. and B. Ø. Palsson, “A protocol for generating a high-quality genome-scale metabolic reconstruction”, *Nature Protocols*, Vol. 5, No. 1, p. 93, 2010.

22. Orth, J. D., I. Thiele and B. Ø. Palsson, “What is flux balance analysis?”, *Nature Biotechnology*, Vol. 28, No. 3, p. 245, 2010.

23. Vallino, J. J. and G. Stephanopoulos, “Carbon flux distributions at the glucose 6-phosphate branch point in Corynebacterium glutamicum during lysine overproduction”, *Biotechnology Progress*, Vol. 10, No. 3, pp. 327–334, 1994.

24. Varma, A., B. W. Boesch and B. O. Palsson, “Biochemical production capabilities of Escherichia coli”, *Biotechnology and Bioengineering*, Vol. 42, No. 1, pp. 59–73, 1993.

25. Feist, A. M., M. J. Herrgård, I. Thiele, J. L. Reed and B. Ø. Palsson, “Reconstruction of biochemical networks in microorganisms”, *Nature Reviews Microbiology*, Vol. 7, No. 2, p. 129, 2009.

26. Pitkänen, E., P. Jouhten, J. Hou, M. F. Syed, P. Blomberg, J. Kludas, M. Oja, L. Holm, M. Penttilä, J. Rousu *et al.*, “Comparative genome-scale reconstruction of gapless metabolic networks for present and ancestral species”, *PLoS Computational Biology*, Vol. 10, No. 2, p. e1003465, 2014.

27. Kerkhoven, E. J., P.-J. Lahtvee and J. Nielsen, “Applications of computational modeling in metabolic engineering of yeast”, *FEMS Yeast Research*, Vol. 15, No. 1, pp. 1567–1364, 2014.

28. Chen, N., I. J. del Val, S. Kyriakopoulos, K. M. Polizzi and C. Kontoravdi, "Metabolic network reconstruction: advances in in silico interpretation of analytical information", *Current Opinion in Biotechnology*, Vol. 23, No. 1, pp. 77–82, 2012.

29. Durot, M., P.-Y. Bourguignon and V. Schachter, "Genome-scale models of bacterial metabolism: reconstruction and applications", *FEMS Microbiology Reviews*, Vol. 33, No. 1, pp. 164–190, 2008.

30. Dikicioglu, D., B. Kırdar and S. G. Oliver, "Biomass composition: the "elephant in the room" of metabolic modelling", *Metabolomics*, Vol. 11, No. 6, pp. 1690–1701, 2015.

31. Machado, D. and M. Herrgård, "Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism", *PLoS Computational Biology*, Vol. 10, No. 4, p. e1003580, 2014.

32. Ramkrishna, D. and H.-S. Song, "Dynamic models of metabolism: Review of the cybernetic approach", *AIChE Journal*, Vol. 58, No. 4, pp. 986–997, 2012.

33. Kim, T. Y., S. B. Sohn, Y. B. Kim, W. J. Kim and S. Y. Lee, "Recent advances in reconstruction and applications of genome-scale metabolic models", *Current Opinion in Biotechnology*, Vol. 23, No. 4, pp. 617–623, 2012.

34. Stephanopoulos, G., "Metabolic fluxes and metabolic engineering", *Metabolic Engineering*, Vol. 1, No. 1, pp. 1–11, 1999.

35. Stephanopoulos, G., "Synthetic biology and metabolic engineering", *ACS Synthetic Biology*, Vol. 1, No. 11, pp. 514–525, 2012.

36. Österlund, T., I. Nookaew and J. Nielsen, "Fifteen years of large scale metabolic modeling of yeast: developments and impacts", *Biotechnology Advances*, Vol. 30, No. 5, pp. 979–988, 2012.

37. Pinzon, W., H. Vega, J. Gonzalez and A. Pinzon, "Mathematical Framework Behind the Reconstruction and Analysis of Genome Scale Metabolic Models", *Archives of Computational Methods in Engineering*, pp. 1–14, 2018.

38. Reimers, A.-M. and A. C. Reimers, "The steady-state assumption in oscillating and growing systems", *Journal of Theoretical Biology*, Vol. 406, pp. 176–186, 2016.

39. Thiele, I. and B. Ø. Palsson, "Bringing genomes to life: the use of genome-scale in silico models", *Introduction to Systems Biology*, pp. 14–36, Springer, 2007.

40. Price, N. D., J. L. Reed and B. Ø. Palsson, "Genome-scale models of microbial cells: evaluating the consequences of constraints", *Nature Reviews Microbiology*, Vol. 2, No. 11, p. 886, 2004.

41. Lewis, N. E., H. Nagarajan and B. O. Palsson, "Constraining the metabolic genotype–phenotype relationship using a phylogeny of in silico methods", *Nature Reviews Microbiology*, Vol. 10, No. 4, pp. 291–305, 2012.

42. Gélinas, P., "Inventions on baker's yeast strains and specialty ingredients", *Recent Patents on Food, Nutrition & Agriculture*, Vol. 1, No. 2, pp. 104–132, 2009.

43. Goffeau, A., J. Park, I. T. Paulsen, J.-L. JONNIAUX, T. Dinh, P. Mordant and M. H. SAIER JR, "Multidrug-resistant transport proteins in yeast: complete inventory and phylogenetic characterization of yeast open reading frames within the major facilitator superfamily", *Yeast*, Vol. 13, No. 1, pp. 43–54, 1997.

44. Dujon, B., "The yeast genome project: what did we learn?", *Trends in Genetics*, Vol. 12, No. 7, pp. 263–270, 1996.

45. Botstein, D., S. A. Chervitz and M. Cherry, "Yeast as a model organism", *Science*, Vol. 277, No. 5330, pp. 1259–1260, 1997.

46. Barnett, J. A., “A history of research on yeasts 1: work by chemists and biologists 1789–1850”, *Yeast*, Vol. 14, No. 16, pp. 1439–1451, 1998.

47. Barnett, J. A., “A history of research on yeasts 2: Louis Pasteur and his contemporaries, 1850–1880”, *Yeast*, Vol. 16, No. 8, pp. 755–771, 2000.

48. Larsson, C., I.-l. Påhlman and L. Gustafsson, “The importance of ATP as a regulator of glycolytic flux in Saccharomyces cerevisiae”, *Yeast*, Vol. 16, No. 9, pp. 797–809, 2000.

49. Díaz-Ruiz, R., N. Avéret, D. Araiza, B. Pinson, S. Uribe-Carvajal, A. Devin and M. Rigoulet, “Mitochondrial Oxidative Phosphorylation Is Regulated by Fructose 1, 6-Bisphosphate A Possible Role in Crabtree Effect Induction?”, *Journal of Biological Chemistry*, Vol. 283, No. 40, pp. 26948–26955, 2008.

50. ter Kuile, B. H. and H. V. Westerhoff, “Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway”, *FEBS Letters*, Vol. 500, No. 3, pp. 169–171, 2001.

51. Daran-Lapujade, P., M. L. Jansen, J.-M. Daran, W. van Gulik, J. H. de Winde and J. T. Pronk, “Role of transcriptional regulation in controlling fluxes in central carbon metabolism of Saccharomyces cerevisiae A chemostat culture study”, *Journal of Biological Chemistry*, Vol. 279, No. 10, pp. 9125–9138, 2004.

52. Liu, Z. and R. A. Butow, “A transcriptional switch in the expression of yeast tricarboxylic acid cycle genes in response to a reduction or loss of respiratory function”, *Molecular and Cellular Biology*, Vol. 19, No. 10, pp. 6720–6728, 1999.

53. Coleman, J. and J. Bhattacharjee, “Regulation of citrate synthase activity ofSaccharomyces cerevisiae”, *Antonie van Leeuwenhoek*, Vol. 41, No. 1, pp. 249–256, 1975.

54. Gadde, D. M. and M. T. McCammon, “Mutations in theIDH2Gene Encoding

the Catalytic Subunit of the Yeast NAD+-Dependent Isocitrate Dehydrogenase Can Be Suppressed by Mutations in theCIT1Gene Encoding Citrate Synthase and Other Genes of Oxidative Metabolism", *Archives of Biochemistry and Biophysics*, Vol. 344, No. 1, pp. 139–149, 1997.

55. Bakker, B. M., K. M. Overkamp, A. J. van Maris, P. Kötter, M. A. Luttik, J. P. van Dijken and J. T. Pronk, "Stoichiometry and compartmentation of NADH metabolism in Saccharomyces cerevisiae", *FEMS Microbiology Reviews*, Vol. 25, No. 1, pp. 15–37, 2001.

56. Zitomer, R. S., D. L. Montgomery, D. L. Nichols and B. D. Hall, "Transcriptional regulation of the yeast cytochrome c gene", *Proceedings of the National Academy of Sciences*, Vol. 76, No. 8, pp. 3627–3631, 1979.

57. de Vries, S. and C. A. Marres, "The mitochondrial respiratory chain of yeast. Structure and biosynthesis and the role in cellular metabolism", *Biochimica et Biophysica Acta (BBA)-Reviews on Bioenergetics*, Vol. 895, No. 3, pp. 205–239, 1987.

58. Cimini, D., K. R. Patil, C. Schiraldi and J. Nielsen, "Global transcriptional response of Saccharomyces cerevisiae to the deletion of SDH3", *BMC Systems Biology*, Vol. 3, No. 1, p. 17, 2009.

59. Lagunas, R. and C. Gancedo, "Role of phosphate in the regulation of the Pasteur effect in Saccharomyces cerevisiae", *European Journal of Biochemistry*, Vol. 137, No. 3, pp. 479–483, 1983.

60. Kaliterna, J., R. A. Weusthuis, J. I. Castrillo, J. P. Van Dijken and J. T. Pronk, "Transient responses of Candida utilis to oxygen limitation: regulation of the Kluyver effect for maltose", *Yeast*, Vol. 11, No. 4, pp. 317–325, 1995.

61. Scheffers, W., "Stimulation of fermentation in yeasts by acetoin and oxygen",

Nature, Vol. 210, No. 5035, pp. 533–534, 1966.

62. Mans, R., J.-M. G. Daran and J. T. Pronk, “Under pressure: evolutionary engineering of yeast strains for improved performance in fuels and chemicals production”, *Current Opinion in Biotechnology*, Vol. 50, pp. 47–56, 2018.

63. Ho, P.-W., S. Swinnen, J. Duitama and E. Nevoigt, “The sole introduction of two single-point mutations establishes glycerol utilization in Saccharomyces cerevisiae CEN. PK derivatives”, *Biotechnology for Biofuels*, Vol. 10, No. 1, p. 10, 2017.

64. Bracher, J. M., E. de Hulster, C. C. Koster, M. van den Broek, J.-M. G. Daran, A. J. van Maris and J. T. Pronk, “Laboratory evolution of a biotin-requiring Saccharomyces cerevisiae strain for full biotin prototrophy and identification of causal mutations”, *Applied and Environmental Microbiology*, Vol. 83, No. 16, 2017.

65. Smith, J., E. Van Rensburg and J. F. Görgens, “Simultaneously improving xylose fermentation and tolerance to lignocellulosic inhibitors through evolutionary engineering of recombinant Saccharomyces cerevisiaeharbouring xylose isomerase”, *BMC Biotechnology*, Vol. 14, No. 1, p. 41, 2014.

66. Voordeckers, K., J. Kominek, A. Das, A. Espinosa-Cantu, D. De Maeyer, A. Arslan, M. Van Pee, E. van der Zande, W. Meert, Y. Yang *et al.*, “Adaptation to high ethanol reveals complex evolutionary pathways”, *PLoS Genetics*, Vol. 11, No. 11, p. e1005635, 2015.

67. Fletcher, E., A. Feizi, M. M. Bisschops, B. M. Hallström, S. Khoomrung, V. Siewers and J. Nielsen, “Evolutionary engineering reveals divergent paths when yeast is adapted to different acidic environments”, *Metabolic Engineering*, Vol. 39, pp. 19–28, 2017.

68. Gaxiola, R., I. F. De Larrinoa, J. M. Villalba and R. Serrano, “A novel and

conserved salt-induced protein is an important determinant of salt tolerance in yeast.", *The EMBO Journal*, Vol. 11, No. 9, pp. 3157–3164, 1992.

69. Turanlı-Yıldız, B., L. Benbadis, C. Alkım, T. Sezgin, A. Akşit, A. Gökçe, Y. Öztürk, A. T. Baykal, Z. P. Çakar and J. M. François, "In vivo evolutionary engineering for ethanol-tolerance of Saccharomyces cerevisiae haploid cells triggers diploidization", *Journal of Bioscience and Bioengineering*, Vol. 124, No. 3, pp. 309–318, 2017.

70. Sürmeli, Y., C. Holyavkin, A. Topaloğlu, M. Arslan, H. İ. Kısakesen and Z. P. Çakar, "Evolutionary engineering and molecular characterization of a caffeine-resistant Saccharomyces cerevisiae strain", *World Journal of Microbiology and Biotechnology*, Vol. 35, No. 12, pp. 1–16, 2019.

71. Hacısalihoğlu, B., C. Holyavkin, A. Topaloğlu, H. İ. Kısakesen and Z. P. Çakar, "Genomic and transcriptomic analysis of a coniferyl aldehyde-resistant Saccharomyces cerevisiae strain obtained by evolutionary engineering", *FEMS Yeast Research*, Vol. 19, No. 3, p. foz021, 2019.

72. Balaban, B. G., Ü. Yılmaz, C. Alkım, A. Topaloğlu, H. İ. Kısakesen, C. Holyavkin and Z. P. Çakar, "Evolutionary Engineering of an Iron-Resistant Saccharomyces cerevisiae Mutant and Its Physiological and Molecular Characterization", *Microorganisms*, Vol. 8, No. 1, p. 43, 2020.

73. Küçükgöze, G., C. Alkım, Ü. Yılmaz, H. İ. Kısakesen, S. Gündüz, S. Akman and Z. P. Çakar, "Evolutionary engineering and transcriptomic analysis of nickel-resistant Saccharomyces cerevisiae", *FEMS Yeast Research*, Vol. 13, No. 8, pp. 731–746, 2013.

74. Holyavkin, C., *Evolutionary engineering of phenylethanol-resistant Saccharomyces cerevisiae*, Ph.D. Thesis, Istanbul Technical University, Graduate School of Science, Engineering and Technology, 2013.

75. Terzioğlu, E., C. Alkım, M. Arslan, B. G. Balaban, C. Holyavkin, H. İ. Kısakesen, A. Topaloğlu, Ü. Yılmaz Şahin, S. Gündüz Işık, S. Akman *et al.*, "Genomic, transcriptomic and physiological analyses of silver-resistant Saccharomyces cerevisiae obtained by evolutionary engineering", *Yeast*, Vol. 37, No. 9-10, pp. 413–426, 2020.

76. DeRisi, J. L., V. R. Iyer and P. O. Brown, "Exploring the metabolic and genetic control of gene expression on a genomic scale", *Science*, Vol. 278, No. 5338, pp. 680–686, 1997.

77. Cho, R. J., M. Fromont-Racine, L. Wodicka, B. Feierbach, T. Stearns, P. Legrain, D. J. Lockhart and R. W. Davis, "Parallel analysis of genetic selections using whole genome oligonucleotide arrays", *Proceedings of the National Academy of Sciences*, Vol. 95, No. 7, pp. 3752–3757, 1998.

78. Förster, J., I. Famili, P. Fu, B. Ø. Palsson and J. Nielsen, "Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network", *Scientific Reports*, Vol. 13, No. 2, pp. 244–253, 2003.

79. Förster, J., I. Famili, B. Ø. Palsson and J. Nielsen, "Large-scale evaluation of in silico gene deletions in Saccharomyces cerevisiae", *OMICS A Journal of Integrative Biology*, Vol. 7, No. 2, pp. 193–202, 2003.

80. Famili, I., J. Förster, J. Nielsen and B. O. Palsson, "Saccharomyces cerevisiae phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network", *Proceedings of the National Academy of Sciences*, Vol. 100, No. 23, pp. 13134–13139, 2003.

81. Duarte, N. C., M. J. Herrgård and B. Ø. Palsson, "Reconstruction and validation of Saccharomyces cerevisiae iND750, a fully compartmentalized genome-scale metabolic model", *Genome Research*, Vol. 14, No. 7, pp. 1298–1309, 2004.

82. Lopes, H. and I. Rocha, "Genome-scale modeling of yeast: chronology, applications and critical perspectives", *FEMS Yeast Research*, Vol. 17, No. 5, 2017.

83. Kuepfer, L., U. Sauer and L. M. Blank, "Metabolic functions of duplicate genes in Saccharomyces cerevisiae", *Genome Research*, Vol. 15, No. 10, pp. 1421–1430, 2005.

84. Nookaew, I., M. C. Jewett, A. Meechai, C. Thammarongtham, K. Laoteng, S. Cheevadhanarak, J. Nielsen and S. Bhumiratana, "The genome-scale metabolic model iIN800 of Saccharomyces cerevisiae and its validation: a scaffold to query lipid metabolism", *BMC Systems Biology*, Vol. 2, No. 1, p. 71, 2008.

85. Mo, M. L., B. Ø. Palsson and M. J. Herrgård, "Connecting extracellular metabolomic measurements to intracellular flux states in yeast", *BMC Systems Biology*, Vol. 3, No. 1, p. 37, 2009.

86. Zomorrodi, A. R. and C. D. Maranas, "Improving the iMM904 S. cerevisiae metabolic model using essentiality and synthetic lethality data", *BMC Systems Biology*, Vol. 4, No. 1, p. 178, 2010.

87. Herrgård, M. J., N. Swainston, P. Dobson, W. B. Dunn, K. Y. Arga, M. Arvas, N. Blüthgen, S. Borger, R. Costenoble, M. Heinemann *et al.*, "A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology", *Nature Biotechnology*, Vol. 26, No. 10, p. 1155, 2008.

88. Dobson, P. D., K. Smallbone, D. Jameson, E. Simeonidis, K. Lanthaler, P. Pir, C. Lu, N. Swainston, W. B. Dunn, P. Fisher *et al.*, "Further developments towards a genome-scale metabolic model of yeast", *BMC Systems Biology*, Vol. 4, No. 1, p. 145, 2010.

89. Heavner, B. D., K. Smallbone, B. Barker, P. Mendes and L. P. Walker, "Yeast 5– an expanded reconstruction of the Saccharomyces cerevisiae metabolic network",

BMC Systems Biology, Vol. 6, No. 1, p. 55, 2012.

90. Heavner, B. D., K. Smallbone, N. D. Price and L. P. Walker, "Version 6 of the consensus yeast metabolic network refines biochemical coverage and improves model performance", *Database*, Vol. 2013, 2013.

91. Aung, H. W., S. A. Henry and L. P. Walker, "Revising the representation of fatty acid, glycerolipid, and glycerophospholipid metabolism in the consensus model of yeast metabolism", *Industrial Biotechnology*, Vol. 9, No. 4, pp. 215–228, 2013.

92. Lu, H., F. Li, B. J. Sánchez, Z. Zhu, G. Li, I. Domenzain, S. Marcišauskas, P. M. Anton, D. Lappa, C. Lieven *et al.*, "A consensus S. cerevisiae metabolic model Yeast8 and its ecosystem for comprehensively probing cellular metabolism", *Nature Communications*, Vol. 10, No. 1, pp. 1–13, 2019.

93. Bro, C., B. Regenberg, J. Förster and J. Nielsen, "In silico aided metabolic engineering of Saccharomyces cerevisiae for improved bioethanol production", *Metabolic Engineering*, Vol. 8, No. 2, pp. 102–111, 2006.

94. Burgard, A. P., P. Pharkya and C. D. Maranas, "Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization", *Biotechnology and Bioengineering*, Vol. 84, No. 6, pp. 647–657, 2003.

95. Ng, C., M.-y. Jung, J. Lee and M.-K. Oh, "Production of 2, 3-butanediol in Saccharomyces cerevisiae by in silico aided metabolic engineering", *Microbial Cell Factories*, Vol. 11, No. 1, p. 68, 2012.

96. Xu, G., W. Zou, X. Chen, N. Xu, L. Liu and J. Chen, "Fumaric acid production in Saccharomyces cerevisiae by in silico aided metabolic engineering", *PloS One*, Vol. 7, No. 12, p. e52086, 2012.

97. Sun, Z., H. Meng, J. Li, J. Wang, Q. Li, Y. Wang and Y. Zhang, "Identification of novel knockout targets for improving terpenoids biosynthesis in Saccharomyces

cerevisiae", *PloS One*, Vol. 9, No. 11, p. e112615, 2014.

98. Chowdhury, R., A. Chowdhury and C. Maranas, "Using gene essentiality and synthetic lethality information to correct yeast and CHO cell genome-scale models", *Metabolites*, Vol. 5, No. 4, pp. 536–570, 2015.

99. Arslan, M., C. Holyavkin, H. İ. Kısakesen, A. Topaloğlu, Y. Sürmeli and Z. P. Çakar, "Physiological and transcriptomic analysis of a chronologically long-lived Saccharomyces cerevisiae strain obtained by evolutionary engineering", *Molecular Biotechnology*, Vol. 60, No. 7, pp. 468–484, 2018.

100. Nilsson, A. and J. Nielsen, "Metabolic trade-offs in yeast are caused by F1F0-ATP synthase", *Scientific Reports*, Vol. 6, p. 22264, 2016.

101. Sánchez, B. J., C. Zhang, A. Nilsson, P.-J. Lahtvee, E. J. Kerkhoven and J. Nielsen, "Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints", *Molecular Systems Biology*, Vol. 13, No. 8, 2017.

102. Huber, W., V. J. Carey, R. Gentleman, S. Anders, M. Carlson, B. S. Carvalho, H. C. Bravo, S. Davis, L. Gatto, T. Girke *et al.*, "Orchestrating high-throughput genomic analysis with Bioconductor", *Nature Methods*, Vol. 12, No. 2, p. 115, 2015.

103. Davis, S. and P. S. Meltzer, "GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor", *Bioinformatics*, Vol. 23, No. 14, pp. 1846–1847, 2007.

104. Ritchie, M. E., B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi and G. K. Smyth, "limma powers differential expression analyses for RNA-sequencing and microarray studies", *Nucleic Acids Research*, Vol. 43, No. 7, pp. e47–e47, 2015.

105. Edwards, J. S., R. Ramakrishna and B. O. Palsson, "Characterizing the metabolic

phenotype: a phenotype phase plane analysis", *Biotechnology and Bioengineering*, Vol. 77, No. 1, pp. 27–36, 2002.

106. Tack, I. L., P. Nimmegeers, S. Akkermans, I. Hashem and J. F. Van Impe, "Simulation of Escherichia coli dynamics in biofilms and submerged colonies with an individual-based model including metabolic network information", *Frontiers in Microbiology*, Vol. 8, p. 2509, 2017.

107. Thiele, I., R. M. Fleming, A. Bordbar, J. Schellenberger and B. Ø. Palsson, "Functional characterization of alternate optimal solutions of Escherichia coli's transcriptional and translational machinery", *Biophysical Journal*, Vol. 98, No. 10, pp. 2072–2081, 2010.

108. Mahadevan, R. and C. Schilling, "The effects of alternate optimal solutions in constraint-based genome-scale metabolic models", *Metabolic Engineering*, Vol. 5, No. 4, pp. 264–276, 2003.

109. Reed, J. L. and B. Ø. Palsson, "Genome-scale in silico models of E. coli have multiple equivalent phenotypic states: assessment of correlated reaction subsets that comprise network states", *Genome Research*, Vol. 14, No. 9, pp. 1797–1805, 2004.

110. Gudmundsson, S. and I. Thiele, "Computationally efficient flux variability analysis", *BMC Bioinformatics*, Vol. 11, No. 1, p. 489, 2010.

111. Kiatsupaibul, S., R. L. Smith and Z. B. Zabinsky, "An analysis of a variation of hit-and-run for uniform sampling from general regions", *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, Vol. 21, No. 3, p. 16, 2011.

112. Saa, P. A. and L. K. Nielsen, "ll-ACHRB: a scalable algorithm for sampling the feasible solution space of metabolic networks", *Bioinformatics*, Vol. 32, No. 15, pp. 2330–2337, 2016.

113. Schellenberger, J. and B. Ø. Palsson, "Use of randomized sampling for analysis of metabolic networks", *Journal of Biological Chemistry*, Vol. 284, No. 9, pp. 5457–5461, 2009.

114. Heirendt, L., S. Arreckx, T. Pfau, S. N. Mendoza, A. Richelle, A. Heinken, H. S. Haraldsdóttir, J. Wachowiak, S. M. Keating, V. Vlasov *et al.*, "Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v. 3.0", *Nature Protocols*, Vol. 14, No. 3, pp. 639–702, 2019.

115. Wang, H., S. Marcišauskas, B. J. Sánchez, I. Domenzain, D. Hermansson, R. Agren, J. Nielsen and E. J. Kerkhoven, "RAVEN 2.0: A versatile toolbox for metabolic network reconstruction and a case study on Streptomyces coelicolor", *PLoS Computational Biology*, Vol. 14, No. 10, p. e1006541, 2018.

116. Van Hoek, P., J. P. Van Dijken and J. T. Pronk, "Effect of specific growth rate on fermentative capacity of baker's yeast", *Appl. Environ. Microbiol.*, Vol. 64, No. 11, pp. 4226–4233, 1998.

117. Jeske, L., S. Placzek, I. Schomburg, A. Chang and D. Schomburg, "BRENDA in 2019: a European ELIXIR core data resource", *Nucleic Acids Research*, Vol. 47, No. D1, pp. D542–D549, 2019.

118. Bordel, S., R. Agren and J. Nielsen, "Sampling the solution space in genome-scale metabolic networks reveals transcriptional regulation in key enzymes", *PLoS Computational Biology*, Vol. 6, No. 7, p. e1000859, 2010.

119. Button, D., "Biochemical Basis for Whole-Cell Uptake Kinetics: Specific Affinity, Oligotrophic Capacity, and the Meaning of the Michaelis Constant", .

120. Penninckx, M. J., "An overview on glutathione in Saccharomyces versus non-conventional yeasts", *FEMS Yeast Research*, Vol. 2, No. 3, pp. 295–305, 2002.

121. Kepes, F. and R. Schekman, "The yeast SEC53 gene encodes phosphomannomu-

tase.”, *Journal of Biological Chemistry*, Vol. 263, No. 19, pp. 9155–9161, 1988.

122. Oltmanns, O. and A. Bacher, “Biosynthesis of riboflavine in Saccharomyces cerevisiae: the role of genes rib 1 and rib 7”, *Journal of Bacteriology*, Vol. 110, No. 3, pp. 818–822, 1972.

123. Arnold, I., M. F. Bauer, M. Brunner, W. Neupert and R. A. Stuart, “Yeast mitochondrial F1F0-ATPase: the novel subunit e is identical to Tim11”, *FEBS Letters*, Vol. 411, No. 2-3, pp. 195–200, 1997.

124. Trembath, M., B. Monk, G. Kellerman and A. Linnane, “Biogenesis of mitochondria 36”, *Molecular and General Genetics MGG*, Vol. 141, No. 1, pp. 9–22, 1975.

125. Tokatlidis, K., T. Junne, S. Moes, G. Schatz and B. S. Glick, “Translocation arrest of an intramitochondrial sorting signal next to Tim11 at the inner-membrane import site”, *Nature*, Vol. 384, No. 6609, pp. 585–588, 1996.

126. Meunier, B., N. Fisher, S. Ransac, J.-P. Mazat and G. Brasseur, “Respiratory complex III dysfunction in humans and the use of yeast as a model organism to study mitochondrial myopathy and associated diseases”, *Biochimica Et Biophysica Acta (BBA)-Bioenergetics*, Vol. 1827, No. 11-12, pp. 1346–1361, 2013.

127. Athenstaedt, K. and G. Daum, “Biosynthesis of phosphatidic acid in lipid particles and endoplasmic reticulum of Saccharomyces cerevisiae”, *Journal of Bacteriology*, Vol. 179, No. 24, pp. 7611–7616, 1997.

128. Jiménez, A., M. A. Santos and J. L. Revuelta, “Phosphoribosyl pyrophosphate synthetase activity affects growth and riboflavin production in Ashbya gossypii”, *BMC Biotechnology*, Vol. 8, No. 1, pp. 1–12, 2008.

129. Schneiter, R., A. T. Carter, Y. Hernando, G. Zellnig, L. M. Schweizer and M. Schweizer, “The importance of the five phosphoribosyl-pyrophosphate syn-

thetase (Prs) gene products of Saccharomyces cerevisiae in the maintenance of cell integrity and the subcellular localization of Prs1p", *Microbiology*, Vol. 146, No. 12, pp. 3269–3278, 2000.

130. Ugbogu, E. A., S. Wippler, M. Euston, E. N. Kouwenhoven, A. P. de Brouwer, L. M. Schweizer and M. Schweizer, "The contribution of the nonhomologous region of Prs1 to the maintenance of cell wall integrity and cell viability", *FEMS Yeast Research*, Vol. 13, No. 3, pp. 291–301, 2013.

131. De Smet, C. H., E. Vittone, M. Scherer, M. Houweling, G. Liebisch, J. F. Brouwers and A. I. de Kroon, "The yeast acyltransferase Sct1p regulates fatty acid desaturation by competing with the desaturase Ole1p", *Molecular Biology of the Cell*, Vol. 23, No. 7, pp. 1146–1156, 2012.

132. Hasslacher, M., A. Ivessa, F. Paltauf and S. Kohlwein, "Acetyl-CoA carboxylase from yeast is an essential enzyme and is regulated by factors that control phospholipid metabolism", *Journal of Biological Chemistry*, Vol. 268, No. 15, pp. 10946–10952, 1993.

133. Morris, C. P., F. Lim and J. C. Wallace, "Yeast pyruvate carboxylase: gene isolation", *Biochemical and Biophysical Research Communications*, Vol. 145, No. 1, pp. 390–396, 1987.

134. Tsibris, J. C., D. B. McCormick and L. D. Wright, "Studies on the binding and function of flavin phosphates with flavin mononucleotide-dependent enzymes", *Journal of Biological Chemistry*, Vol. 241, No. 5, pp. 1138–1143, 1966.

135. Boucherié, H., N. Bataille, I. T. Fitch, M. Perrot and M. F. Tuite, "Differential synthesis of glyceraldehyde-3-phosphate dehydrogenase polypeptides in stressed yeast cells", *FEMS Microbiology Letters*, Vol. 125, No. 2-3, pp. 127–133, 1995.

136. Barber, R. D., D. W. Harmer, R. A. Coleman and B. J. Clark, "GAPDH as a

housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues", *Physiological Genomics*, Vol. 21, No. 3, pp. 389–395, 2005.

137. Yang, Y., W. Fan, L. Zhu, T. Zhao, L. Ma, Y. Wu, R. Ge and M. Fan, "Effects of hypoxia on mRNA expression of housekeeping genes in rat brain tissue and primary cultured neural cells", *Frontiers of Medicine in China*, Vol. 2, No. 3, pp. 239–243, 2008.

138. Sweeny, E. A., A. B. Singh, R. Chakravarti, O. Martinez-Guzman, A. Saini, M. M. Haque, G. Garee, P. D. Dans, L. Hannibal, A. R. Reddi *et al.*, "Glyceraldehyde-3-phosphate dehydrogenase is a chaperone that allocates labile heme in cells", *Journal of Biological Chemistry*, Vol. 293, No. 37, pp. 14557–14568, 2018.

139. Chuang, D.-M., C. Hough and V. V. Senatorov, "Glyceraldehyde-3-phosphate dehydrogenase, apoptosis, and neurodegenerative diseases", *Annu. Rev. Pharmacol. Toxicol.*, Vol. 45, pp. 269–290, 2005.

140. Hohmann, S., "PDC6, a weakly expressed pyruvate decarboxylase gene from yeast, is activated when fused spontaneously under the control of the PDC1 promoter", *Current Genetics*, Vol. 20, No. 5, pp. 373–378, 1991.

141. Tkach, J. M., A. Yimit, A. Y. Lee, M. Riffle, M. Costanzo, D. Jaschob, J. A. Hendry, J. Ou, J. Moffat, C. Boone *et al.*, "Dissecting DNA damage response pathways by analysing protein localization and abundance changes during DNA replication stress", *Nature Cell Biology*, Vol. 14, No. 9, pp. 966–976, 2012.

142. Pryde, F. E., T. C. Huckle and E. J. Louis, "Sequence analysis of the right end of chromosome XV in Saccharomyces cerevisiae: an insight into the structural and functional significance of sub-telomeric repeat sequences", *Yeast*, Vol. 11, No. 4, pp. 371–382, 1995.

143. Kornblatt, M., J. Richard Albert, S. Mattie, J. Zakaib, S. Dayanandan, P. Hanic-

Joyce and P. Joyce, “The Saccharomyces cerevisiae enolase-related regions encode proteins that are active enolases”, *Yeast*, Vol. 30, No. 2, pp. 55–69, 2013.

144. Ho, C. K., A. F. Lam and L. S. Symington, “Identification of nucleases and phosphatases by direct biochemical screen of the Saccharomyces cerevisiae proteome”, *PloS One*, Vol. 4, No. 9, p. e6993, 2009.

145. Grisolia, S. and J. Tecson, “Mercury-induced reversible increase in 2, 3-diphosphoglycerate phosphatase and concomitant decrease in mutase activity of animal phosphoglycerate mutases”, *Biochimica et Biophysica Acta (BBA)-Enzymology*, Vol. 132, No. 1, pp. 56–67, 1967.

146. Nosaka, K., H. Nishimura, Y. Kawasaki, T. Tsujihara and A. Iwashima, “Isolation and characterization of the THI6 gene encoding a bifunctional thiamin-phosphate pyrophosphorylase/hydroxyethylthiazole kinase from Saccharomyces cerevisiae.”, *Journal of Biological Chemistry*, Vol. 269, No. 48, pp. 30510–30516, 1994.

147. Amillet, J.-M. and R. Labbe-Bois, “Isolation of the gene HEM4 encoding uroporphyrinogen III synthase in Saccharomyces cerevisiae”, *Yeast*, Vol. 11, No. 5, pp. 419–424, 1995.

148. Tan, F.-C., Q. Cheng, K. Saha, I. U. Heinemann, M. Jahn, D. Jahn and A. G. Smith, “Identification and characterization of the Arabidopsis gene encoding the tetrapyrrole biosynthesis enzyme uroporphyrinogen III synthase”, *Biochemical Journal*, Vol. 410, No. 2, pp. 291–299, 2008.

149. Chiang, P. K. and C. GL, “Activation of methionine for transmethylation. Purification of the S-adenosylmethionine synthetase of bakers’ yeast and its separation into two forms.”, *The Journal of Biological Chemistry*, 1977.

150. Girrbach, V. and S. Strahl, “Members of the evolutionarily conserved PMT family of proteinO-mannosyltransferases form distinct protein complexes among them-

selves", *Journal of Biological Chemistry*, Vol. 278, No. 14, pp. 12554–12562, 2003.

151. Turanlı-Yıldız, B., L. Benbadis, C. Alkım, T. Sezgin, A. Akşit, A. Gökçe, Y. Öztürk, A. T. Baykal, Z. P. Çakar and J. M. François, "In vivo evolutionary engineering for ethanol-tolerance of Saccharomyces cerevisiae haploid cells triggers diploidization", *Journal of Bioscience and Bioengineering*, Vol. 124, No. 3, pp. 309–318, Sep. 2017.

152. Sürmeli, Y., C. Holyavkin, A. Topaloğlu, M. Arslan, H. İ. Kısakesen and Z. P. Çakar, "Evolutionary engineering and molecular characterization of a caffeine-resistant Saccharomyces cerevisiae strain", *World Journal of Microbiology and Biotechnology*, Vol. 35, No. 12, Nov. 2019.

153. Adeboye, P. T., M. Bettiga, F. Aldaeus, P. T. Larsson and L. Olsson, "Catabolism of coniferyl aldehyde, ferulic acid and p-coumaric acid by Saccharomyces cerevisiae yields less toxic products", *Microbial Cell Factories*, Vol. 14, No. 1, p. 149, 2015.

154. Hacısalihoğlu, B., C. Holyavkin, A. Topaloğlu, H. İ. Kısakesen and Z. P. Çakar, "Genomic and transcriptomic analysis of a coniferyl aldehyde-resistantSaccharomyces cerevisiaestrain obtained by evolutionary engineering", *FEMS Yeast Research*, Vol. 19, No. 3, Apr. 2019.

155. Horstmann, C., C. Campbell, D. S. Kim and K. Kim, "Transcriptome profile with 20 nm silver nanoparticles in yeast", *FEMS Yeast Research*, 2019.

156. Galván Márquez, I., M. Ghiyasvand, A. Massarsky, M. Babu, B. Samanfar, K. Omidi, T. W. Moon, M. L. Smith and A. Golshani, "Zinc oxide and silver nanoparticles toxicity in the baker's yeast, Saccharomyces cerevisiae", *PloS One*, Vol. 13, No. 3, p. e0193111, 2018.

APPENDIX A: RESULTS OF DIFFERENTIAL EXPRESSION ANALYSIS AND INTEGRATION INFORMATION

The content is given in the CD.

APPENDIX B: RECONSTRUCTED MODEL FILES

The content is given in the CD.

APPENDIX C: SOLUTION VECTORS OF ANALYSES

The content is given in the CD.