EVALUATING AND IMPROVING STOICHIOMETRICALLY CONSTRAINED MODELS OF YEAST METABOLISM FOR APPLICATION TO DESIGN OF METABOLIC ENGINEERING STRATEGIES

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Benjamin Douglas Heavner

August 2012

EVALUATING AND IMPROVING STOICHIOMETRICALLY CONSTRAINED MODELS OF YEAST METABOLISM FOR APPLICATION TO DESIGN OF METABOLIC ENGINEERING STRATEGIES

Benjamin Douglas Heavner, Ph. D.

Cornell University 2012

This document presents the results of efforts to apply the Yeast Consensus Reconstruction model of the *Saccharomyces cerevisiae* metabolic network to develop a metabolic engineering strategy for industrial strain improvement. Following a review of the development of mathematical models of metabolism, it describes an evaluation of the Consensus Reconstruction. We find that the computational reconstruction of this portion of metabolism differs from the biochemistry of this pathway as described in the literature. Our efforts to correct these discrepancies are described in Chapter 4. The updated model improves both the accuracy of the metabolic reconstruction and the prediction of viability and auxotrophy phenotypes, thus demonstrating that literature-based curation is a technique which can be successfully applied to improve the model.

Chapter 5 describes an *in silico* screen for formate-producing yeast mutants. By working to reproduce an *in silico* screen previously conducted using the iND750 model, we found that the computational prediction of formate-producing yeast mutants is sensitive to implementation details and reaction constraints when using either the iND750 or the Yeast model. Our results suggest that comparative analysis of constraint based models is a useful tool for improving models of the yeast metabolic network. The main text concludes with a summary and discussion

of future research opportunities in Chapter 6. MATLAB scripts which enable evaluation of model predictive accuracy and demonstrate model applications such growth simulation and mutant library screening are included as appendices.

This work contributes to the ongoing effort to develop systems biotechnology tools which will enable the rational design of new microbial strains, and which may enable broader industrial-scale application of biotechnology. The fields of computational biology and systems biotechnology are young, and abundant opportunity remains to develop and apply this technology to meet human needs.

BIOGRAPHICAL SKETCH

Ben Heavner was born on December 1, 1976 in Seattle, WA, the third child of Dr. James Edward Heavner and Rev. Betsey Jean Clark Heavner. He graduated from Lubbock High School in Lubbock, TX in 1995, and earned a Bachelor of Science degree with double majors in Mathematics and Natural Sciences with an emphasis in Physics from the University of Puget Sound in 1999. Ben earned a Master of Science degree in Civil and Environmental Engineering from the University of Colorado, Boulder in 2004, and matriculated at Cornell University in 2006.

Ben is married to Gretchen Lee Watson Heavner, who is a doctoral candidate in Cornell's Department of Civil and Environmental Engineering. Their son, Elias Jacob Heavner, was born in 2011. Following graduation, the Heavners plan to move to Seattle, WA where Ben will continue researching the computational reconstruction of metabolism at the Institute for Systems Biology as a member of Dr. Nathan Price's research group.

Dedicated with love to Gretchen and Eli

ACKNOWLEDGMENTS

This work was not completed in isolation, and thanks are due to many people for their constructive advice, mentorship, fellowship, and support during its completion. Thanks to:

- my committee members for their suggestions and intellectual guidance; particularly to Professor L. P. Walker, whose interest in stoichiometrically constrained metabolic models served as an impetus for my work, and whose continuing mentorship enabled the completion of this dissertation;
- my research collaborators in the Walker and Henry research groups, and to those with whom I co-authored publications;
- the faculty and staff of the Department of Biological and Environmental Engineering. Daily interactions have made our department home, and your assistance in navigating the administrative requirements of graduate studies at Cornell greatly smoothed the process;
- fellow members of The Brewing Conspiracy for friendships gained, adventures shared, and challenges overcome. I hope to maintain these friendships, and know that I will carry their positive influences far beyond my time of active research at Cornell;
- to my immediate and extended family for your love, support, and for setting a high bar for me.

Especially heartfelt thanks to Gretchen, for building this adventure with me.

TABLE OF CONTENTS

CHAPTER 1 - INTRODUCTION

Work to characterize, harness, and improve the use of microbial metabolism for human benefit is an ongoing effort, involving the disparate techniques of biological scientists, process engineers, and increasingly, computational experts. Archeological evidence suggests that this work has been ongoing, perhaps since the dawn of human civilization: fermentation to produce bread or beer may have motivated the domestication of cereal grains [1]. Fermentation research led to the modern science of biochemistry [2], and the genome of *Saccharomyces cerevisiae* was the first eukaryotic genome sequence to be completed [3]. In our effort to maximize yeast's productive capacity for industrial-scale biotechnology applications, we have refined fermentation vessels and processes, optimized media, and used selective breeding methods for strain improvement. Most recently, in the post-genome era, we have begun working to apply the knowledge gained through techniques of molecular biology and genetics to develop predictive computer models of yeast physiology to implement rational strain design and synthetic biology approaches.

An illustrative example is provided by research efforts to improve the industrial production of ethanol via fermentation for use as a liquid transportation fuel, thus reducing our society's dependence upon petroleum [4]. In working to improve the production of first-generation biofuels, researchers and process engineers have taken a variety of approaches in attempts reduce fermentation process time and to improve product yield. Efforts have included the use of various substrate feeding modes [5]; novel fermentation process designs, such as those involving multiple vessels with cell recycling [6]; supplementing with nutrients [7] or vitamins [8]; and optimizing temperature [9]; oxygen concentration [10]; initial inoculum size [11]; and combinations of these factors [12]. This strategy of fermentation process optimization has been

very successful at the laboratory scale, with results including a fermentation of a high-sugar wheat media to produce 21.5% (v/v) ethanol in approximately 175 hr without nutritional supplement [11]. However, the production of such a high ethanol concentration required large quantities of dry yeast as the inoculums, a technique which causes extended lag times in the absence of yeast preconditioning [13]. Such increased fermentation time would increase industrial scale capital costs significantly due to fermentation vessel size requirements. There are other obstacles to industrial-scale application of successful laboratory-scale approaches: the inoculum required to replicate these results at a large scale would be prohibitively expensive, and the organism is subject to a wide variety of stresses in the industrial context, including osmotic, temperature, and pH stresses which arise from process considerations [14]. Other supplementation and process control strategies face similar economic challenges when scaled up. What's more, strategies of altering nutritional or process approaches are not sufficient for the industrial production of compounds which are not endogenously biosynthesized. Thus, it is hoped that improved bioprocesses and biocatalysts may be obtained more economically by modifying the fermentative organism itself through metabolic engineering or synthetic biology approaches.

Traditional genetic approaches to yeast strain development and improvement include breeding programs [15] and generating and screening libraries of mutants for a desired phenotype [16]. These approaches are time- and labor-intensive – drawbacks which must be overcome for rapid development of industrial biotechnology. With the emergence of large datasets in the post-genomic era, scientists and engineers have built upon the computational foundations built by earlier efforts to model metabolism to develop new high-throughput and *in silico* approaches to strain improvement, approaches which have been termed "Systems

Biotechnology" [17].

As described in the next chapter, the ideas of mathematically modeling single reactions have been extended to metabolic pathways, and scaled up to model entire metabolic networks. In the face of challenges to organism-scale kinetic models, researchers have developed other approaches, such as using steady state assumptions and applying mathematical programming algorithms to stoichiometrically constrained computational reconstructions of the metabolic reaction network – an approach that has been termed "flux balance analysis" [18], or FBA.

With FBA providing the mathematical foundation for the analysis of reconstructed metabolic networks, researchers began encoding the hard-gained knowledge represented by metabolic pathway maps into computer formats amenable to simulation. Ongoing efforts by different research groups have led to a proliferation of such models – including at least 16 different genome-scale models of yeast metabolism (Table 2.2). Further, efforts to reduce the need for extensive manual review of model details and to increase the rate at which such models can be built have led to the development of software pipelines to automate the process of building metabolic models from genomic data [19]. As they were developed, these automated processes were quickly applied to create stoichiometrically-constrained models for hundreds of organisms, and have now been extended in hopes of generating metabolic models as quickly as new genomes are sequenced (there are 1541 models in the Model SEED database at time of writing [20]).

The development of reconstruction and simulation methods has been driven by the desire to facilitate application. Indeed, the increasing range of published applications of FBA to stoichiometrically constrained metabolic models has been the subject of at least two recent review articles which summarize their application for contextualizing high-throughput data,

metabolic engineering, model-directed discovery, interrogating multi-species relationships, interpretations of phenotypic screens, analysis of network properties and studies of evolutionary processes [21, 22]

The research described in this dissertation is an effort to validate and improve the computational reconstruction of yeast metabolism so that a yeast model may be applied to inform metabolic engineering. In beginning this work, my objective was to apply an existing model of yeast metabolism [23] to develop strategies to enhance the production of compounds which have previously been linked to improve stress tolerance. However, when I reviewed the biochemical reactions included in the Consensus Reconstruction of Yeast Metabolism [23, 24], I found that this model did not accurately reconstruct the metabolic pathways involved in the biosynthesis of complex sphingolipids (Chapter 3, [25]). Thus, I expanded my objectives to:

1. Conduct literature-based validation of the reconstruction of sphingolipid metabolism included in the Consensus Reconstruction of Yeast Metabolism
2. Update the existing model by incorporate any changes required to correct errors in the reconstruction
3. Evaluate the impact of these changes on the predictive accuracy of the model
4. Apply the updated model to develop metabolic engineering strategies for strain improvement
5. Evaluate any differences in strategies suggested by the updated model with predictions of previously published models

I found that correcting errors in the reconstruction of established biochemistry improved model predictions of gene essentiality and knockout-induced auxotrophies (Chapter 4, [26]). Following these corrections to the model, I reviewed published examples of applying constraint-

based models of yeast metabolism to implement metabolic engineering strategies. I re-implemented an *in silico* screen of simulated knockout mutants for strains predicted to produce formate, and found that my simulations made different predictions than earlier yeast metabolic models. This comparative approach of evaluating phenotype and metabolic flux predictions using different metabolic models also proved to be a valuable tool for further model refinement (Chapter 5).

This work documents that there remains a need for careful manual curation of the biochemistry included in stoichiometrically constrained models of yeast metabolism, and that a curation approach based upon manual review of relevant literature is a successful strategy for improving the predictive accuracy of such models. I observe that model-based metabolic engineering tools are sensitive to details of implementation, but can be successfully applied to strain improvement. Thus, this dissertation contributes to the ongoing effort to develop systems biotechnology tools which will enable the rational design of new microbial strains, and which may enable broader industrial-scale application of biotechnology.

REFERENCES

1. Braidwood RJ, Sauer JD, Helbaek H, Mangelsdorf PC, Cutler HC, Coon CS, Linton R, Steward J, Oppenheim AL: **Symposium: Did Man Once Live by Beer Alone?** *American Anthropologist* 1953, **55**:515–526.

2. Cornish-Bowden A: *New beer in an old bottle : Eduard Buchner and the growth of biochemical knowledge*. Valencia: Universitat de València; 1998.

3. Mewes HW, Albermann K, Bähr M, Frishman D, Gleissner A, Hani J, Heumann K, Kleine K, Maierl A, Oliver SG, Pfeiffer F, Zollner A: **Overview of the yeast genome**. *Nature* 1997, **387**:7–65.

4. **Cellulosic Ethanol: Benefits and Challenges** [http://genomicscience.energy.gov/biofuels/benefits.shtml].

5. Kana G, Oloke JK, Lateef A: **Novel feeding strategies for Saccharomyces cerevisiae DS2155, using glucose limited exponential fedbatch cultures with variable specific growth rates (μ)**. *African Journal of Biotechnology* 2007, **6**:1122–1127.

6. Ben Chaabane F, Aldiguier AS, Alfenore S, Cameleyre X, Blanc P, Bideaux C, Guillouet SE, Roux G, Molina-Jouve C: **Very high ethanol productivity in an innovative continuous two-stage bioreactor with cell recycle**. *Bioprocess Biosyst Eng* 2006, **29**:49–57.

7. Reddy LVA, Reddy OVS: **Improvement of ethanol production in very high gravity fermentation by horse gram (Dolichos biflorus) flour supplementation**. *Lett Appl Microbiol* 2005, **41**:440–444.

8. Alfenore S, Molina-Jouve C, Guillouet SE, Uribelarrea JL, Goma G, Benbadis L: **Improving ethanol production and viability of Saccharomyces cerevisiae by a vitamin feeding strategy during fed-batch process**. *Applied Microbiology and Biotechnology* 2002, **60**:67–72.

9. Jones AM, Ingledew WM: **Fuel alcohol production: optimization of temperature for efficient very-high-gravity fermentation**. *Appl. Environ. Microbiol* 1994, **60**:1048–1051.

10. Nagodawithana TW, Castellano C, Steinkraus KH: **Effect of dissolved oxygen, temperature, initial cell count, and sugar concentration on the viability of Saccharomyces cerevisiae in rapid fermentations**. *Applied and Environmental Microbiology* 1974, **28**:383.

11. Thomas KC, Ingledew WM: **Production of 21%(v/v) ethanol by fermentation of very high gravity (VHG) wheat mashes**. *Journal of Industrial Microbiology and Biotechnology* 1992, **10**:61–68.

12. Jones HL, Margaritis A, Stewart RJ: **The combined effects of oxygen supply strategy, inoculum size and temperature profile on very-high-gravity beer fermentation by Saccharomyces cerevisiae**. *Journal of the Institute of Brewing* 2007, **113**:168–184.

13. Bellissimi E, Ingledew WM: **Metabolic acclimatization: preparing active dry yeast for fuel ethanol production**. *Process Biochemistry* 2005, **40**:2205–2213.

14. Attfield PV: **Stress tolerance: The key to effective strains of industrial baker's yeast**. *Nat Biotechnol* 1997, **15**:1351–1357.

15. Romano P, Soli MG, Suzzi G, Grazia L, Zambonelli C: **Improvement of a Wine Saccharomyces cerevisiae Strain by a Breeding Program**. *Appl. Environ. Microbiol.* 1985, **50**:1064–1067.

16. Rowlands RT: **Industrial strain improvement: rational screens and genetic recombination techniques**. *Enzyme and Microbial Technology* 1984, **6**:290–300.

17. Lee SY, Lee D-Y, Kim TY: **Systems biotechnology for strain improvement**. *Trends in Biotechnology* 2005, **23**:349–358.

18. Orth JD, Thiele I, Palsson BØ: **What is flux balance analysis?** *Nature Biotechnology* 2010, **28**:245–248.

19. Notebaart RA, Van Enckevort FH., Francke C, Siezen RJ, Teusink B: **Accelerating the reconstruction of genome-scale metabolic networks**. *BMC bioinformatics* 2006, **7**:296.

20. Henry CS, DeJongh M, Best AA, Frybarger PM, Linsay B, Stevens RL: **High-throughput generation, optimization and analysis of genome-scale metabolic models**. *Nat Biotechnol* 2010, **28**:977–982.

21. Feist AM, Palsson BO: **The growing scope of applications of genome-scale metabolic reconstructions using Escherichia coli**. *Nat Biotechnol* 2008, **26**:659–667.

22. Oberhardt MA, Palsson BØ, Papin JA: **Applications of genome-scale metabolic reconstructions**. *Mol Syst Biol* 2009, **5**.

23. Herrgard MJ, Swainston N, Dobson P, Dunn WB, Arga KY, Arvas M, Buthgen N, Borger S, Costenoble R, Heinemann M, Hucka M, Le Novere N, Li P, Liebermeister W, Mo ML, Oliveira AP, Petranovic D, Pettifer S, Simeonidis E, Smallbone K, Spasie I, Weichart D, Brent R, Broomhead DS, Westerhoff HV, Kurdar B, Penttila M, Klipp E, Palsson BO, Sauer U, Oliver SG, Mendes P, Nielsen J, Kell DB: **A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology**. *Nat Biotech* 2008, **26**:1155–1160.

24. Dobson PD, Jameson D, Simeonidis E, Lanthaler K, Pir P, Lu C, Swainston N, Dunn WB, Fisher P, Hull D, Brown M, Oshota O, Stanford NJ, Kell DB, King RD, Oliver SG, Stevens RD, Mendes P: **Further developments towards a genome-scale metabolic model of yeast**. *BMC Syst Biol* 2010, **4**:145.

25. Heavner BD, Henry SA, Walker LP: **Evaluating Sphingolipid Biochemistry in the Consensus Reconstruction of Yeast Metabolism**. *Industrial Biotechnology* 2012, **8**:72–78.

26. Heavner BD, Smallbone K, Barker B, Mendes P, Walker LP: **Yeast 5 - an expanded reconstruction of the Saccharomyces Cerevisiae metabolic network**. *BMC Systems Biology* 2012, **6**:55.

CHAPTER 2 – LITERATURE REVIEW

Improving industrial yeast strains

Stress tolerance is critically important for yeast strains in industrial biotechnology application because "in industrial practices, yeast never encounters the physiologically optimum environment and instead is exposed to a variety of stresses" which reduce their productivity [1]. Our understanding of the mechanisms by which yeast adapts to stress has improved over time. In 1974, Hayashida et al. found that yeasts acquired enhanced ethanol tolerance when grown in the presence of a fungal mycelium extract traditionally used in the production of sake, concluding that "the formation of a high concentration of alcohol in sake mash is related to the lipid metabolism of yeast" [2]. Hayashida expanded upon this work in a series of articles, showing that supplementation with the proteolipid extract increased fermentative activity and "alcohol-durability" [3]; that the proteolipid extract consisted primarily of phospholipids and protein; that the phospholipids included phosphatidylcholine (PC), "a sphingolipid", and lysophosphatidylcholine (LPC) [4]; and that supplementation with PC and ergosteryl oleate resulted in effects similar to supplementation with the fungal extract [5]. Other supplements have also been found to increase yeast's tolerance to stress caused by increasing ethanol concentrations in their environment, including linoleic acid and oleic acid [6] and inositol [7, 8].

More recently, research has documented endogenously produced inositol in genetically modified strains has the same effect of improving ethanol tolerance as inositol supplemented from exogenous sources, though stress tolerance of such strains is further enhanced by additional inositol supplementation [9]. Thus, metabolic engineering of yeast strains by redirecting metabolic flux to the production of compounds which have been documented to improve stress tolerance may be a fruitful approach for designing better-performing industrial yeasts. However,

the process of rationally designing such metabolic engineering strategies remains a challenge to the development such stress-tolerant strains.

Computational context for the development of industrial yeast strains

Successful development of predictive models of yeast metabolism will permit more effective industrial application of yeast-based biotechnology. Such models of metabolism would enable a rational design approach to genetic engineering for developing industrial yeast strains, an approach which has the potential to provide industrial biotechnology solutions to current constraints on microbe-dependent production of fuels or other chemicals [10]. Historical approaches to organizing emerging knowledge of the biochemistry of metabolism into models can be separated into two general classes: 1) qualitative or conceptual maps of the chemical pathways through which substrates are transformed by cells to produce new biomass; and 2) quantitative mathematical models which describe the dynamics of metabolism. The ability to use the conceptual information embodied in pathway maps for quantitative analysis of biological systems is a fairly recent innovation. The emergence of this modeling approach is described below.

Emergence of the reductionist approach

Attempts to develop mathematical tools to describe metabolism have taken both top-down and bottom-up approaches. Top-down approaches consider metabolism as a "black box" and attempt to predict gross physiological behavior without consideration of the underlying mechanism (an example of this approach is predicting optimal feed rates for cultures based upon dissolved oxygen measurements in industrial biotechnology applications [11]). Bottom-up

models are based upon the reductionist approach which attempts to develop holistic kinetic models from mechanistic descriptions of individual reactions. Top-down models describing the gross kinetics of cellular physiology without consideration of the chemical reactions which constitute metabolism have long been successfully applied by engineers in an industrial context for processes such as wastewater treatment with activated sludge [12]. However, because such models do not describe internal metabolic processes, they are of only limited use for providing insights for rational design strategies for manipulating metabolic networks to improve substrate utilization or product formation.

Models which may be applied to the challenge of testing rational design strategies have arisen from efforts to model metabolism in 19th and 20th century which focused on developing mechanistic models from a reductionist kinetic approach. Development of such models continues [13]. However, as our understanding of *in vivo* enzyme kinetics has advanced, it has become clear that this reductionist approach is currently insufficient for describing emergent properties of systems-level functions and the control of metabolism [14]. Further, it has become apparent that enzyme kinetic characteristics in conditions applied for most *in vitro* characterization experiments may vary greatly from *in vivo* conditions. Fully characterizing *in vivo* kinetics has proven to be an extremely challenging problem.

As recently as 2006, Minton [15] demonstrated our lack of knowledge of *in vivo* kinetics when his commentary concluded with the question "By how much do biochemical reactions within cells differ from those in test tubes?" As described in the "Kinetics of multi-reaction systems" section of this chapter, a variety of approaches has been applied to the challenge of

assessing uncertainty of *in vivo* kinetics, including various parameter estimation and sensitivity analysis approaches.

While it had been previously thought that the reductionist approach of characterizing the kinetics of individual enzyme-mediated reactions was more amenable to developing computational models than the conceptual mapping approach of drawing chemical pathway maps, there has recently been a reconsideration of the potential to create predictive quantitative models from data embodied in pathway maps. Using an approach rooted in network theory and systems analysis, a modeling framework of applying constraints has been developed that has predictive capabilities without full kinetic characterization of the reactions such maps represent. In this framework, the model does not depend upon complete kinetic information, but any available kinetic information can serve as a constraint on the set of feasible metabolic states. Such constrained network models serve as the foundation for the proposed research, and are described in the "Stoichiometric modeling" section of this chapter.

Kinetics of enzyme-mediated reactions

Improved understanding of yeast metabolism has been a scientific goal since before the 19th century debate of whether fermentation was a chemical process or solely a biological function – a debate described as being between "vitalists" and biochemists [16]. When the Buchner brothers observed fermentation of sucrose by cell extract [17], they laid the foundation for our current understanding of enzyme-mediated reactions and their role in converting environmental resources to cell biomass, energy, and waste products. Although enzymology was an established specialty prior to the 1890's, it had been limited in scope because all known

enzymes at that time were exo-enzymes which catalyzed hydrolysis reactions [18]. In fact, when the term "enzyme" was coined in 1877 [19], was used to describe the catalyzing agent for reactions occurring outside of organisms. Neither theories describing the mechanism of action nor mathematical models for enzyme kinetics had been well established at the time the Buchners were investigating fermentation. Further, the key role of enzymes in catalyzing metabolic reactions had not been conclusively demonstrated prior to their work. Through their cell-free fermentation, the Buchners demonstrated a link between enzyme activity and the chemical reactions of metabolism, and are thus generally credited with laying the cornerstone for the new science of biochemistry. With the understanding that metabolism consisted of a series of enzymatically-catalyzed reactions, biochemists had a new tool – enzymology – with which to build a theoretical framework for metabolic pathways. Other tools are essential for mapping metabolic pathways, but enzymology was the starting point for developing mathematical models of metabolism.

Search for a rate-limiting step

From the beginning, enzymologists embraced a reductionist approach to understanding the kinetics of metabolic pathways by characterizing the kinetics of enzymes which had been extracted from cells and purified. This approach was explicitly justified in an influential 1905 paper by British plant physiologist Frederick Blackman [20]. Beginning with the axiom "When a process is conditioned as to its rapidity by a number of separate factors, the rate of the process is limited by the pace of the 'slowest' factor", Blackman argued that the enzyme-catalyzed reactions of metabolism should be considered individually to determine the rate-limiting factor:

> "Regarding the cell, as we now may, from the metabolic point of view, as a congeries of enzymes, a colloidal honeycomb of katalytic agents, as many in number as there are cell-functions, and each capable of being isolated and made to do its particular work alone in vitro, we look for light on the action of chemical stimuli in the cell to their effect on the action of isolated enzyme in vitro."

Such reasoning was extended to developing mathematical models of metabolism. If the kinetics of each enzyme in a metabolic pathway could be described mathematically, the rate-limiting enzyme would determine the kinetics of an entire metabolic pathway. It was thought that, the interdependence of various metabolic pathways could be explored mathematically, and the rate-limiting pathway would control the kinetics of a cell's metabolism. Thus, metabolism could be mathematically described with the same formulation as the kinetics of the key rate-limiting enzyme. Presumably, metabolic rates were controlled by the key rate-limiting enzymes– if the kinetics of the rate-limiting step could be influenced, it would create a mechanism to affect the rate of the entire pathway [21].

The Michaelis-Menten formulation

Development of mathematical models of enzyme kinetics proceeded throughout the 20th century as enzymologists worked to extract, purify, and quantify the action of key enzymes. By 1902, Henri [22] and Brown [23] had independently devised kinetic models for enzyme-catalyzed reactions which were based on the idea of the formation of an enzyme-substrate complex, relying upon the assumption that substrates reach enzymes by diffusion at rates that are governed by chemical mass-action kinetics. This theoretical approach was given a firmer experimental foundation in 1913 by L. Michaelis and M. Menten [24]. In 1925, Briggs and Haldane applied a quasi-steady state assumption that the concentration of enzyme-substrate

intermediates is constant, and rederived Michaelis and Menten's kinetic model into to the now-classic Michaelis-Menten kinetic equation for irreversible enzyme-mediated reactions [25]. The Michaelis-Menten equation relates the substrate concentration (a) to the initial (v_o) and limiting (v_{max}) rates of the reaction, via the Michaelis constant, K_M (equation 2.a.1).

$$v_0 = \frac{v_{max} a}{K_M + a} \qquad \text{2.a.1}$$

The Michaelis-Menten equation remains a key mathematical model for describing the kinetics of irreversible reactions catalyzed by purified enzymes. However, attempting to apply the Michaelis-Menten equation to the regulated reactions of metabolic pathways can be problematic. For example, metabolic pathways must be able to rapidly respond to subtle environmental changes to maintain homeostasis inside a cell. Michaelis-Menten kinetics do not reflect this property. As Cornish-Bowden discusses, "a simple calculation shows that the rate is 0.1v when $a = K_M/9$, and that it is 0.9v when $a = 9K_M$. In other words, an enormous increase in substrate concentration, 81-fold, is required to bring about a comparatively modest increase in rate from 10% to 90% of the limit" [21]. In other words, Michaelis-Menten kinetics alone do not adequately describe the dynamics observed in metabolic pathway responses to environmental changes.

Further problems with Michaelis-Menten

Michaelis-Menten kinetics are often, but not universally, appropriate for describing the kinetics of a single enzyme. Even early in the 20th century, single-enzyme reactions had been documented with behavior that departed from Michaelis-Menten kinetics. In his measurements

of oxygen binding to hemoglobin, Bohr observed a cooperative effect in which binding of a single oxygen molecule with hemoglobin facilitated the binding of more hemoglobin [26] – a process now known as cooperative binding. Using empirical methods, Hill derived an equation (equation 2.a.2) to describe the kinetics of such reactions with the introduction of a coefficient to express the degree of cooperativity of an enzyme [27].

$$\mathrm{v} = \frac{V a^h}{K_{0.5}^h + a^h} \qquad \text{2.a.2}$$

Note that the Hill equation is equivalent to the Michaelis-Menten equation in the trivial case that the Hill coefficient, *h*, is 1. Hill did not hypothesize about the mechanism of cooperative binding. It was not until the mid 1960's that theoretical models of allosteric effects, in which the binding of a substrate or cofactor alters the conformation of an enzyme, which in turn affects the affinity of other binding sites, were elaborated to describe cooperative binding [28, 29].

The mathematical framework for describing the kinetics of enzyme-mediated reactions has continued to advance and become more complex. Equations have been developed to describe the kinetics of reversible reactions, reactions with product and substrate inhibition, enzyme complex formation, temperature and pH effects, and other situations observed in biochemical pathways [21]. Continuing the paradigm of the early enzymologists, much of the effort of characterizing enzyme kinetics has focused on individual enzyme-mediated reactions, and the bulk of enzyme kinetic data has been collected by enzymologists working with extracted and purified enzymes. However, there has also been increasing interest in developing mathematical

models of the kinetics of enzyme catalyzed reactions in their biological context.

Kinetics of multi-reaction systems

A key feature of the biological context for enzymatically catalyzed reactions is that they are linked together in the multi-reaction systems that make up metabolic pathways. Efforts to develop mathematical models describing the multi-reaction pathways of metabolism have grown from and progressed in tandem with the development of a theoretical framework for the kinetics of purified enzymes. Like the efforts to develop kinetic models for reactions catalyzed by individual enzymes, research efforts to model multi-reaction systems are ongoing. Biotechnologists have been particularly interested in multi-reaction pathway models which could be used to evaluate strategies for increasing the production of a target metabolite. As described below, although early models of metabolic pathways focused on the reductionist approach espoused by Blackman [20], evidence was mounting by the mid 1950's that the analysis of metabolism required consideration of emergent systems properties which would not be evident in experiments following the reductionist approach of characterizing the kinetics of individually purified enzymes.

Pathway effects

In 1952, three papers were published [30–32] which demonstrated that an end product of a metabolic pathway could inhibit reactions early in the pathway, though the mechanism was unclear at that time. In 1953, Adelberg and Umbarger [33] observed that the presence of valine in the growth medium of certain *E. coli* mutants inhibited the formation of α-ketoisovalerate, a

valine precursor. From this, they concluded that "an end product, in this case valine, can regulate the rate of its own biosynthesis". By 1956, Umbarger had demonstrated that in *E. coli*, isoleucine specifically inhibited the first enzyme-catalyzed reaction in the pathway for its biosynthesis from threonine [34]. He described this process using the terminology of systems control theory, calling it a "negative feedback mechanism". Such a feedback mechanism would not be observed in experiments with a single purified enzyme – multiple pathway enzymes or a media containing many compounds would need to be present to observe the feedback effect in a typical enzyme kinetic experiment. While such an experiment could be set up, the observed feedback effects demonstrated the existence of parameters which alter enzyme kinetics exist in metabolic systems. Such parameters had not previously been considered in the *in vitro* experimental conditions of classical enzymology.

Consideration of metabolic flux

Umbarger's concept of feedback suggested that metabolism could be considered from the viewpoint of dynamical systems analysis – that "cellular dynamics" should be a field of study akin to the evaluation mechanical and electronic systems, described using systems of simultaneous nonlinear ordinary and partial differential equations. Viewing the control of metabolic flux[1] in the context of cellular dynamics represented a major conceptual shift. In this

[1] As described by [16], the word *flux* has multiple meanings. In some contexts, it is used for the system property corresponding to the rate of reaction of a single enzyme (such as "the flux through glucokinase in central metabolism"). In other contexts, it is used to describe the rate of reaction through a multi-enzyme pathway (such as

new paradigm, rates of enzyme-mediated reactions may be considered local properties which refer to enzymes isolated from the system, but steady-state fluxes and metabolite concentrations (and their control) are the systemic properties of critical interest [21]. Thus, the study of the control of metabolic flux requires examination of the entire biochemical system, not just the component pieces of individual reactions.

In 1965, Joseph Higgins wrote an article to "develop these ideas to the point where they may provide a more analytical basis and broader viewpoint for the future development and understanding of cellular dynamics" [35]. Higgins suggested that the control of metabolic flux be considered as a function of "Structural Control Variables" (SCV), which he used as a general term for parameters that might influence the flux through a given reaction sequence, such as rate constants, enzyme concentrations, or substrate concentrations. He suggested that the portion of control allocated to a particular SCV could be calculated by measuring the change in flux that resulted from a change in one of the SCVs. He called this parameter the "control strength" of the SCV. Higgins hypothesized that if SCVs were ranked by control strength, one might be found to be "many times greater than the others." In such cases, the SCV with greatest control strength would be said to have control of the flux relative to other variables. However, refining Blackman's concept of a "rate-limiting factor", Higgins noted that the existence of a SCV with the greater control strength was not the same as a "rate limiting" enzyme, as had been generally

---

"the net flux of carbon to biomass formation"). For an unbranched pathway in steady state, the flux through each reaction is the same as the pathway flux.

conceived: “[i]n a normal (constant) stationary state the dynamical variables adjust themselves so that all steps in the sequence will have the same flux …. Consequently, no one step can be considered as rate limiting …. It is preferable to refer to a “rate controlling” step” [35]. Further, he cautioned that the possibility of a dominant SCV did not mean that control was necessarily limited to a single SCV when he wrote, “there need not be any great difference in the control strengths implying that no one step has dominant rate control.”

Elaborating upon his suggestion that the control strength was a key parameter in describing the dynamics of cellular metabolism, Higgins suggested that enzyme concentration is a “natural choice” for the SCV for metabolic reactions. This “natural choice” leads to an obvious approach to measuring control strengths: varying the amount of enzyme in a cellular extract and measuring the resulting change in flux. Successful identification of a rate controlling SCV would be particularly useful for biotechnology, since it would provide a method for increasing metabolic flux to a desired product: overexpression of the gene coding for the rate controlling enzyme. However, though the study of cellular dynamics continued to advance, it would be many years before this strategy was tested, and was found to achieve mixed results. Further advances in the analysis of control would give a theoretical basis for other approaches to metabolic engineering.

Examining the control of metabolism

Cellular dynamics research continued after Higgins’ work to provide a stronger analytical basis for the field, with particular focus on both the control of metabolism and dynamic modeling of metabolism. Consideration of SCVs led to Metabolic Control Analysis (MCA), an approach

rooted in sensitivity analysis and established 1973 by Kacser and Burns [36], and independently by Heinrich and Rapoport [37]. In their paper, Kacser and Burns first summarized previous research concerning the control of biochemical systems as an attempt to synthesize the information represented by metabolic maps with the information discovered through enzymology. They argued that this synthesis lead to "a large set of simultaneous non-linear equations for which there is no explicit solution." Computers could be used to solve particular cases, but more often than not, simulations simply confirmed assumptions about the system, and did not "yield a general theory of control." They argued that the concept of a 'controlling enzyme' was based upon different criteria in different studies, and the field was hampered by imprecise definitions regarding different aspects of control. In an attempt to work towards a general theory of control, they suggested a more quantitative approach to analyzing the sensitivity of enzyme mediated reactions and metabolic fluxes to changes in external parameters, an approach which came to be known as MCA, "a theoretical bridge from the kinetic properties of enzymes to the systems properties of metabolism" [38].

Kacser and Burns began their consideration of the control of flux by differentiating between *parameters* – constraints that can be controlled in an experimental situation (such as enzyme concentration or concentration of a substrate or inhibitor in the media) – and *variables*, which represent "levels of metabolites and the diverse molecular forms arising from them". In other words, the "pools" of metabolites and the fluxes into and out of these pools are the variables that are controlled by various system parameters. With this distinction between parameters and variables, they built upon Higgins' earlier suggestion that the dynamics of

metabolism must be analyzed holistically. As Kacser and Burns argued, "flux is a systemic property and questions of its control cannot be answered by looking at one step in isolation – or even each step in isolation."

From their definitions of parameters and variables, Kacser and Burns began formalizing MCA by proposing three "vague questions" which they considered to be "what many people have in mind when talking about control." They demonstrated that the questions could be restated more precisely, thus becoming amenable to quantitative analysis. These "vague questions" were: 1) What is the quantitative influence of one parameter on a variable?; 2) How much can any single enzyme be controlled?; and 3) How do changes in pool concentrations affect the rate of a reaction? The authors illustrated these questions with an example (Figure 2.a.1), which served as the basis for their development of a quantitative approach to sensitivity analysis of parameters in metabolic systems.

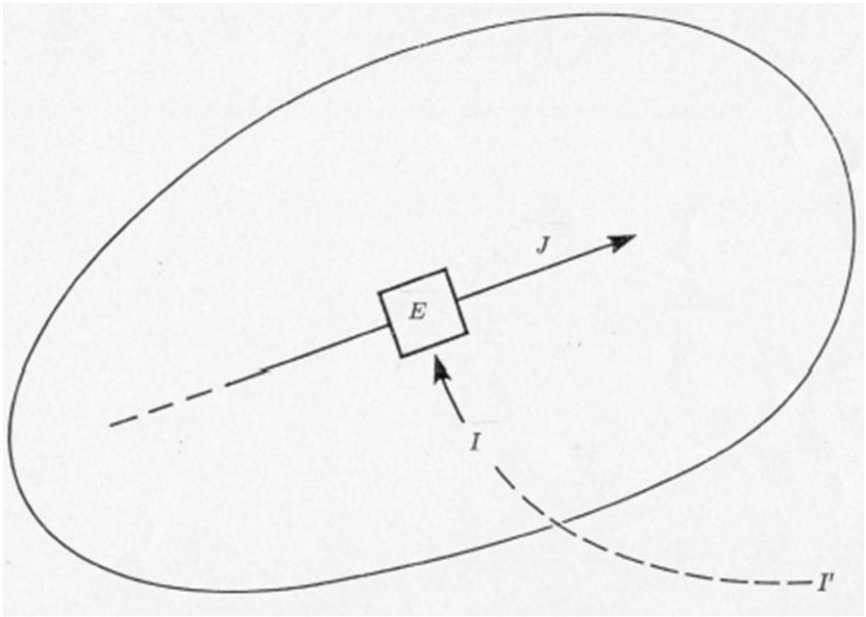


Figure 2.a.1: A thought experiment on the control of flux.

Assume that the flux, *J*, carried by an enzyme-mediated step, *E*, is influenced by the level of an internal parameter, such as an inhibitor, *I*. If *I* is dependent only upon the external concentration of the inhibitor, *I'*, how does varying *I'* change *J*? (from [27])

To enable quantitative analysis of the first vague question (What is the quantitative influence of one parameter on a variable?), Kacser and Burns proposed the use of the "Response Coefficient", which is now designated as R_I^J (using the notation standardized in 1985 [39]). The Response Coefficient is a simple ratio of the normalized variation of a flux, *J*, to a normalized change in a parameter, *I*:

$$\mathrm{R}_I^J \approx \frac{\delta J}{J} \Big/ \frac{\delta I}{I} \qquad \text{2.a.3}$$

Kacser and Burns considered the Response Coefficient to be "an overall measure of the

control exerted" on the flux by the parameter *I* at its level in the system at the time of interest. Continuing their formalization of the "vague questions", they next described how this overall measure could be decomposed into two factors, which captured the essence of the second and third questions (How much can any single enzyme be controlled? and How do changes in pool concentrations affect the rate of a reaction?). The first factor is the response of the enzyme in isolation to variation of a parameter, which they quantified with a "Controllability Coefficient", $\varepsilon_P^{\kappa,i}$. Like the Response Coefficient, the Controllability Coefficient is a ratio – in this case of the normalized variation in a reaction rate to a normalized change in a parameter, *I*:

$$\varepsilon_P^{\kappa,i} \approx \frac{\delta v}{v} \Big/ \frac{\delta I}{I} \qquad \text{2.a.4}$$

The Controllability Coefficient is nonlinear over a range of parameters, so Kacser and Burns argue that it is important to determine it for a given enzyme using substrate and product concentrations held precisely at the steady-state levels obtained in an organism (which is not an easy thing to do experimentally!). The Controllability Coefficient describes the sensitivity of a particular enzyme to particular changes in the environment in otherwise fixed system conditions. It is noted that the rate change measured to determine the Controllability Coefficient is generally different from – and usually higher – than the system flux change which would be measured in determining the Response Coefficient. This difference is due to systemic effects, discussed below. Their formulation of the "Controllability Coefficient" has not been widely adopted in subsequent MCA research, and is now considered a specific form of "Elasticity", a term which was also introduced by Kacser and Burns to describe the effects on variables of substances

generated in the system, which would themselves also be classified as variables, as opposed to parameters. For Kacser and Burns, the distinction between a controllability coefficient and an elasticity was that the controllability coefficient referred to the response of an isolated enzyme to a parameter which was determined externally to the system under consideration, and elasticity referred to the response of an isolated enzyme to a metabolite whose concentration was set by the metabolic system itself, a variable. In more recent studies, this distinction has not been emphasized [40].

The second factor is the response of the whole system to changes of the enzyme, which they quantified with the "Flux Control Coefficient", C_i^J. The Flux Control Coefficient is another ratio – of normalized flux change to normalized change in enzyme concentration (the flux change due to a change in enzyme concentration can be considered equivalent to a flux change caused by an increased concentration of an inhibitor):

$$\mathrm{C}_i^J \approx \frac{\delta J}{J} \Big/ \frac{\delta e_i}{e_i} \qquad \text{2.a.5}$$

Thus the Flux Control Coefficient is a quantification of the system's sensitivity to changes in its components.

The decomposition of the Response Coefficient into the Controllability Coefficient (or elasticity) and the Flux Control Coefficient was not merely semantic. Kacser and Burns showed that the Response Coefficient is in fact the product of the Controllability Coefficient and the Flux Control Coefficient:

$$\mathrm{R}_I^J = C_i^J \varepsilon_P^{\kappa,i} \qquad \text{2.a.6}$$

Thus, while both $\varepsilon_P^{\kappa,i}$ and C_i^J can be evaluated independently, neither is sufficient for explaining the control exerted on the flux by a given parameter, such as enzyme activity, gene copy number, or inhibitor concentration in media, except in extreme cases. As Kacser and Burns wrote, "This relationship reveals that *in vitro* studies can be misleading in predicting the effects in the *in vivo* situation without a knowledge of Flux Control Coefficients. The Flux Control Coefficient is a system property independent of whether any effectors act on the enzyme or what their strength is."

Kacser and Burns extended their consideration of Flux Control Coefficients from single reactions to evaluating the control of pathways, leading to a derivation of the Flux Summation Theorem, which states "if all the enzymes that can affect a particular metabolic flux in a cell or a metabolic system are taken and the values of their control coefficients on that flux added up, the sum comes to 1" [41]:

$$\sum_{1 \to n} C_i^J = 1 \qquad \text{2.a.7}$$

The Flux Summation Theorem has many practical implications. Kacser and Burns pointed out that it suggests that in a chain of enzyme-mediated reactions, either one (or few) enzymes approach full control, or none of them do (and thus control is distributed across the pathway). Evaluating which case holds for a particular pathway requires quantitative analysis. Another corollary of the theorem is that for a system with a sufficiently large number of enzymes, almost all of the enzymes could be said to be present "in excess" because the quantity or activity of any single enzyme could be reduced without an appreciable effect on the flux. On

the other hand, if any single enzyme is drastically reduced, that step may change from having little control to having a great deal of control – its Flux Control Coefficient would increase as it became a bottleneck. Though Kacser and Burns did not explicitly state it in their early work, this formulation of the Summation Theorem suggests a strategy for biotechnological attempts at modifying metabolic fluxes: since the Flux Control Coefficient is most likely to be large (and thus the enzyme of interest will be controlling) when the enzyme concentration is very low, gene deletion-based strategies are more likely to be effective for modifying metabolic fluxes than overexpression approaches, which would have a smaller impact on pathway flux due to the low Flux Control Coefficient of most enzymes.

In their seminal paper, Kacser and Burns also discussed some limitations of their consideration of Metabolic Control: they had only considered the steady state, which applies to either constant volume or expanding systems (expanding systems can have a steady state because while volume increases exponentially, so do enzymes, pools, and fluxes, so concentrations remain time-invariant); and they did not consider compartmentalization that is not associated with organelles, which would affect the "effective concentration surrounding the *in vivo* enzyme".

Expanding and applying MCA

The field of Metabolic Control Analysis was built upon the foundation laid by Kacser and Burns in the early 1970s by an array of theoretical and experimental approaches throughout the 1970's, and 1980's. Theoretical advances included the generalization of Flux Control Coefficients to time-invariant systems [42]; new approaches to the Summation Theorem [43],

and its application to branching pathways [43–45]; applications to cycling and moiety conservation [46, 47]; and restatement in a matrix-oriented approach [48, 49], which was generalized [50].

As Kacser and Burns demonstrated with the decomposition of the response coefficient (Equation 2.a.6), MCA requires quantitative information about both kinetic parameters of individual reactions and of the systemic effects described by control coefficients, C_i^J. Since characterizing enzyme kinetic parameters has been a focus of enzymology from the field's origins in the 19th century, the new experimental work required for application of MCA has focused upon measuring C_i^J values for enzymes in pathways. A variety of experimental approaches have been applied [41]. Approaches have included altering gene dosage in *Neurospora crassa* to investigate fluxes of arginine synthesis [51]; measuring flux differences arising because of various allelic forms of certain enzymes in diploid organisms including *Drosophila melanogaster* [52], *Escherichia coli* [53], *Clarkia xantiana* [54], *Arabidopsis thaliana* [55]; altering gene dose through the use of plasmids [56–58]; using antisense RNA inhibition [59, 60]; measuring natural flux changes due to environmental changes [61]; and titrations by either purified enzymes [62] or specific inhibitors [63]. By the early 1990s, the field of MCA was sufficiently established for a series of review articles that describe many of these advances [40, 64, 65]. However, as the field became established, practical difficulties grew apace.

Experimental challenges to MCA

A first significant challenge to practical MCA is that measurement of elasticity ($\varepsilon_P^{\kappa,i}$)

requires careful quantification of an enzyme's response to small perturbations in a static environment which is identical to its *in vivo* microenvironment. While the differences between *in vivo* and *in vitro* chemical microenvironments have a long history of investigation [66], the apparent necessity of detailed kinetic information for MCA made the distinction between an aqueous and a "crowded" intracellular environment critically important. Such differences include macromolecular crowding [67, 68]; intracellular compartmentalization due to organelles and other spatial effects such as "macromolecular confinement" (Minton 2006); and differing thermodynamic regimes due to low water content (Clegg 1984). By 1984, it was abundantly clear that the water content in a cell "exhibits physical properties that differ from those of dilute solutions" and so "the microenvironment in which metabolic activity actually occurs in cells is not equivalent to the conditions used to determine the properties of enzymes in vitro and which are extensively used to model metabolism" [69].

The hypothesis that enzymatically catalyzed reactions in cells do not depend upon freely diffusing substrates was supported by experiments performed by Kempner and Miller, in which they found cells that were centrifuged to separate proteins from a soluble phase remained viable although there were no macromolecules remaining in the lighter fraction [70]. In describing this work, Clegg wrote, "these studies suggest that few macromolecules may be free in solution in intact cells" and thus even small molecules might be "structure associated", rather than freely diffusing [69]. Further, an assumption that chemical reaction rates are a continuous function may not be true in the case of reactions involving a small number of enzymes, in which case the dynamics of the reactions may instead be discrete. In other words, the mass-action, diffusion

based model of chemical kinetics which underlie even simple models of enzyme kinetics may not hold in the internal metabolic environment of intact cells [71].

The magnitude and significance of changes in enzyme activity due to changes in local chemical environment is not a well-studied field. As recently as 2006, Minton wrote that "greater than order-of-magnitude increases in association rate and equilibrium constants attributable to background interactions have been observed" [15]. New kinetic models have been suggested, including the addition of corrective factors to Michaelis-Menten [72], the power-law approximation [73], and the development of "fractal-like kinetics" [74]. The latter have been simulated with automata on a mathematical grid in a particularly interesting work by Schnell and Tuner [71], but many alternative modeling frameworks require a greater level of detailed knowledge of the local conditions than the classical approaches. Although it is known that "the limited solvent capacity is a relevant constraint acting on *S. cerevisiae* at physiological growth conditions" [75], the impact of such alternative kinetic frameworks upon MCA has not been fully explored in the literature.

Data management challenges to MCA

A second challenge to practical application of MCA arose with the explosion of data following genome sequencing efforts. On one hand, increasing computational power means some larger systems of differential equations can be solved, which has allowed the development of kinetic models on the scale of pathways, such as the sphingolipid pathway model presented by Alvarez-Vasquez [13]. On the other hand, however, such models have not yet scaled to whole organisms. Gene sequencing generated lists of putative gene products and putative enzymes for

more than 180 species between 1995 and 2004 [76]. No kinetic models exist at the scale of even one of these species. Further, many of the putative enzymes assigned through sequencing efforts have had no classical biochemical measurements. Until such measurement occurs, kinetic modeling of reactions they catalyze is "virtually impossible" [16].

In face of the challenge of uncertainty related to the accurate determination of kinetic parameters, efforts to develop predictive models of multi-reaction systems have followed various paths, including parameter estimation approaches and a re-evaluation of the potential to make predictive models from pathway maps without well-defined kinetic information. Both of these approaches have yielded significant results from the perspective of applying modeling techniques to metabolic engineering - recent parameter sensitivity analyses have demonstrated that metabolic networks are "sloppy" by nature, and that parametric sensitivity can vary widely across parameters, an implication that has led to the suggestion that "modelers should focus on predictions rather than parameters" [77]. Reconsideration of mass-balance constraints on multi-reaction systems has given rise to stoichiometric modeling, which has potential for significant application in the field of metabolic engineering.

Stoichiometric modeling

The stoichiometric modeling approach serves as a framework for quantitative analysis of the information contained in biochemical pathway maps. It arose over the course of two decades, beginning with a conceptual separation of the structural constraints on a dynamic system imposed by mass balance, and the dynamic constraints imposed by kinetics. Among the earliest examinations of the dynamic effects of stoichiometric constraints was the development of a

stoichiometric model that exhibited the nonlinear behavior previously explained by allosteric regulation in central carbon metabolism of red blood cells [78]. Despite this demonstration of the significance of stoichiometric constraints, it was some time before the techniques of stoichiometric modeling were developed and gained wider acceptance. Between 1981 and 1994, stoichiometric modeling was given a formal mathematical foundation, and preliminary results established it as a valid modeling framework. After 1994, continuing research led to a great expansion of stoichiometric models in application and scale, as well as attempts to standardize notation, annotation of data sources, and nomenclature.

Separating kinetic and stoichiometric constraints

Consideration of the dynamic mass balance of a simple network can help understand the separation of stoichiometric and kinetic constraints that gave rise to stoichiometric modeling. For the purposes of this example, consider a reaction network consisting of three metabolites and six fluxes (Figure 2.a.2).

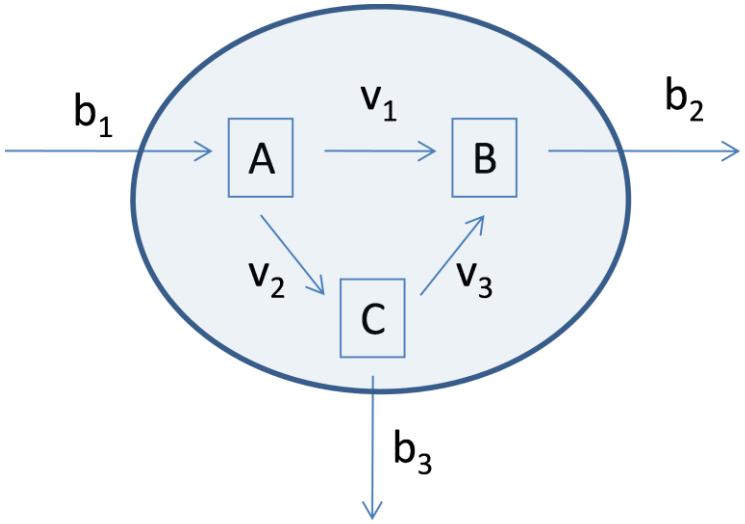


Figure 2.a.2: A simple chemical network.

A simple network consisting of metabolite pools A, B, and C, with internal fluxes v_1, v_2, and v_3, and external fluxes b_1, b_2, and b_3. Based on [70].

A dynamic mass balance may be written for each component of a network by summing the fluxes into and out of each pool. For the simple network under consideration, this mass balance would give rise to three differential equations (Equation 2.a.8).

$$\begin{aligned} \frac{dA}{dt} &= -v_1 - v_2 + b_1 \\ \frac{dB}{dt} &= v_1 + v_3 - b_2 \\ \frac{dC}{dt} &= v_2 - v_3 - b_3 \end{aligned} \qquad \text{2.a.8}$$

This system of equations can be rewritten in matrix form (Equation 2.a.9)

$$\begin{bmatrix} \frac{dA}{dt} \\ \\ \frac{dB}{dt} \\ \\ \frac{dC}{dt} \end{bmatrix} = \begin{bmatrix} -1 & -1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \text{2.a.9}$$

Or, more succinctly (Equation 2.a.10)

$$\frac{dX}{dt} = S \cdot v \quad \text{2.a.1 0}$$

Since there are more fluxes than metabolites, equation 2.b.10 is generally underdetermined – there may not be a unique solution. In other words, there may be more than one valid flux through the reaction network (Figure 2.a.3).

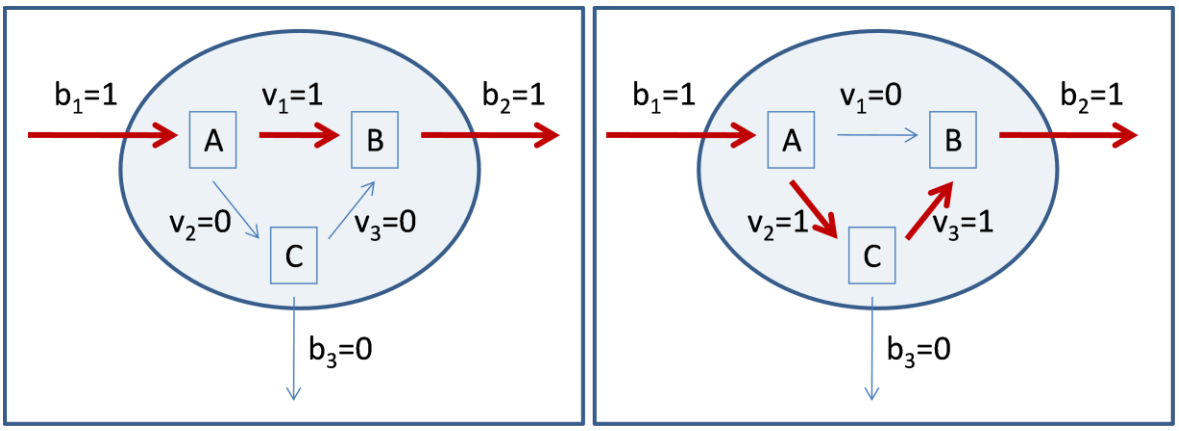


Figure 2.a.3: Multiple valid fluxes.

Since the system is underdetermined, there may be multiple valid fluxes through the network. Active fluxes are shown with bold red arrows. In both cases pictured, $\frac{dX}{dt}$ is **0**, and no metabolites accumulate within the cell. Note that there may also be other permissible fluxes, such as may arise when b_3 is not 0.

As described in the "Kinetics of enzyme-mediated reactions" section of this chapter, the primary focus of metabolic modeling prior to 1992 was describing the time-varying kinetics of enzyme-mediated reactions, which are embodied in the vector *v* of equation 2.a.10.

Although the development of kinetic models was a primary focus of efforts to simulate metabolism, it must be noted that this was not the only approach being explored. In a pioneering, but subsequently under-cited work, Howard M. Shapiro drew upon methods used to build economic models to formulate an input-output model of metabolism [79]. Shapiro's work anticipated future developments through discussion of the use of mathematical programming. As discussed below, a similar approach to modeling metabolism would be re-invented two more times before becoming popularized as flux balance analysis of constraint-based stoichiometric

models.

Theoretical foundations and mathematical advances

By the 1980s emergence of computers enabled the development of approaches for large-scale examination of constraints imposed by the stoichiometric matrix, *S*. An early theoretical foundation for stoichiometric modeling of metabolism was presented in 1981 [80]. In this work, Clark presents "a complete parametric description of all steady states on the closed domain of the general stoichiometric dynamical system." Although this paper describes much of the theoretical basis of stoichiometric network analysis as it is currently practiced, it is a fairly abstract work, and it would be 13 years before some of its ideas were reformulated and popularized through application to large-scale modeling of biological networks. Using differential geometry, Clark provides a proof that the complete set of steady states for stoichiometric dynamical systems can be generated for a range of rate constants and concentrations. This theoretical approach provides "a general mathematical framework for investigating a chemical network's static properties." Among other observations, Clark notes that for all stoichiometric dynamical systems, there exists an equivalent system of equations for which all the steady state solutions exist in a "convex polyhedral cone" – a key description of the solution space that was later used extensively in a popular textbook on stoichiometric modeling [81]. Clark defines the *stoichiometric matrix* as a matrix whose elements are the stoichiometric coefficients of the reactions being examined, and a *stoichiometric network* as a class of matrices and functions that can be described by the general equations of motion:

$$\dot{X} = v \cdot u(X) \qquad \text{2.a.1}$$
$$\dot{k} = 0 \qquad 1$$

(where $\dot{X}$ is the time derivative of concentrations, v is the stoichiometric matrix, u(X) is the "reaction velocity vector", and k is the kinetic rate constant). Clark notes that these equations of motion determine a flow on the network, and that the steady-state flows exist in the left null space of the stoichiometric matrix.

Additional advances to the field of stoichiometric modeling were achieved through examination of the mathematical implications of a full-rank stoichiometric matrix. In a paper that demonstrated the ability of stoichiometrically constrained models to describe various topologies of metabolic networks, Hofmeyr examined the steady-state modeling of four "metabolic structures" – loops, cycles, linear, and branched chains [82]. In this paper, Hofmeyer shows that conservation constraints and flux relationships can be deduced through analysis of the stoichiometric matrix. Although this approach requires a full-rank matrix, Hofmeyer limited the number of required measurements by lumping reactions to reduce degrees of freedom.

Another key mathematical advance was achieved in 1988 when Reder generalized the summation and connectivity theorems from Metabolic Control Analysis, showing that they could be derived from invertible stoichiometric matrices [50]. In this paper, Reder argues that the structural properties alone can provide enough information about the system to describe key behaviors, demonstrating that from a dynamical systems perspective, the conservation relationships of stoichiometry represent a linear "first invariant" of the system. From the basis of the stoichiometric matrix, Reder derives the existence of constant pools of metabolites. Then, in

order to apply the philosophy of metabolic control theory, which examines small variations from the steady state, finds the Jacobian matrix of an invertable stoichiometric matrix, and derives expressions for the familiar MCA concepts of elasticity, summation, and control. Thus, Reder gave formal mathematical justification for the statement that "much information can be derived from the study of the metabolic networks structure alone".

This mathematical sentiment was advanced again in 1988, when Clark (the same author who provided the foundational mathematical theory paper in 1981) described the use of stoichiometric analysis for examining qualitative features of dynamic systems, such as whether a steady state is globally attracting, or whether a chaotic or oscillatory domain is a global attractor [83]. As in the 1981 paper, Clark again describes the geometry of the solution space of the underdetermined system of equations which make up the model as a "convex polyhedral cone" which is created by linear combinations of the "significant overall reactions" that make up its edges. Through mathematical derivation, Clark shows that sensitivity analysis can determine which rate constants have the greatest effect on period and amplitude of chemical oscillators, and concludes that "stoichiometric analysis is an extremely powerful approach for learning about all aspects of the dynamics of nonlinear systems, in which stoichiometric constraints dominate the dynamics. The approach allows one to invalidate mechanisms using only qualitative information about steady-state concentrations and flows, without any information about rate constants."

Working with stoichiometric constraints

The consideration of dynamic mass balances and subsequent partitioning of constraints on metabolic models into static and dynamic descriptors gave rise to the stoichiometric matrix

and its associated equations. As methods for applying stoichiometrically constrained models were developed, two primary approaches were investigated for selecting unique solutions from the family of all possible solutions to the underdetermined equation 2.a.10. Some researchers sought to limit the degrees of freedom by lumping reactions or excising portions of the metabolic network from the model. Others sought to apply techniques of mathematical optimization. Both approaches contributed advances to the understanding of stoichiometric modeling.

Solving the underdetermined problem

An early application of lumping reactions to apply large-scale mass-balance constraints was to determine "the maximal allowable butanol (or any other product) yield", which "is determined by both thermodynamic constraints and the biochemical topology" [84]. To model key yields, Papoutsakis derives a generalized "fermentation equation" that accounts for balances of ATP and reducing energy by exploring the linear dependencies of metabolic pathways. Paoutsakis suggests that the fermentation equation could be used for the development of "gateway sensors" which measure one parameter from which other key parameters can be deduced.

Papoutsakis and Meyer then extend their earlier work on a "fermentation equation" to butanediol and mixed-acid fermentations [85]. As in their earlier paper, they derived a system of equations by lumping reactions into common algebraic expressions. After noting that the stoichiometrically-derived fermentation equations can be written in matrix form as $Lu = v$, the authors suggest that this equation can be used "to calculate the 'best' values of certain desired fermentation parameters (according to the least-squares method, for example) by employing

more experimental data values". As a stoichiometrically-constrained model, they emphasize that their equations are independent of any particular kinetics of biomass growth or product formation.

Rather than using the least-squares approach of Papoutsakis and Meyer, another approach was to reduce the degrees of freedom to a manageable number, and determining the value of remaining variables experimentally. This approach was applied to one of the first stoichiometric models of *S. cerevisiae* [86], which was based upon the stoichiometric growth equations for ethanol and glucose metabolism. The model used 3 different stoichiometric equations for the growth of *S. cerevisiae*: two describing oxidative pathways, and one describing reductive pathways. Based upon elemental biomass composition, the authors found that they needed to experimentally measure 1 component for each pathway in order to solve the system of equations. When they did so, the model had good agreement with experimental results.

In addition to reducing the degrees of freedom in the stoichiometric matrix by lumping or removing equations and using least squares techniques, statistical methods were also employed to integrate stoichiometric model predictions with experimental measurements. In 1987, Papoutsakis and Meyer's 1984-1985 work was applied to analyzing bioreactor data by Chemical Engineers at Drexel University [87]. Rather than using Papoutsakis and Meyer's least squares approach, Tsai and Lee applied the statistical maximum likelihood method to the stoichiometric model to identify any gross measurement errors in bioreactor data collection. A key insight from their work was in a description of the solution space of possible metabolic states in a stoichiometrically constrained model. They suggest that metabolism can "be expressed as a

linear combination of all pathway reactions."

The most advanced stoichiometric model prior to the decisive demonstration of the utility of optimization-based methods described below was published in 1992 [88]. The authors evaluated the metabolic network of *C. glutamicum* with a stoichiometrically constrained mass balance modeling approach in which they solved the dynamic mass balance equation with a steady state assumption that intracellular metabolites do not accumulate, and by lumping equations into "singular groups" which accounted for all dependent reactions so that the stoichiometric matrix was full-rank. Thus, the authors developed a model of batch fermentation based upon stoichiometry rather than dynamics.

The optimization approach

While some researchers worked to reduce the dimensionality of the stoichiometric matrix so that they could find unique solutions to the model, others worked to apply other methods to solving an underdetermined system of equations. A biochemistry instructor named MR Watson presented software intended for teaching with metabolic maps in 1984. In this under-appreciated paper [89], Watson re-invented many aspects of Shapiro's input-output modeling approach [79] and described his application of the Simplex algorithm [90] to find the best solution to the under-determined system of equations represented by equation 2.a.10. This approach would be re-invented again before becoming popularized as "Flux Balance Analysis" (FBA).

Mathematically, the linear programming approach may be stated as follows:

$$\text{Maximize } c^T x \qquad \text{2.a.1}$$

$$\text{Subject to } \mathrm{Ax} \leq b, x \geq 0 \qquad 2$$

Where $\mathrm{c}^T x$ represents a criterion, or "objective function" which is to be minimized or maximized, and $\mathrm{Ax} \leq b, x \geq 0$ represent constraints which are applied to the model. The process of evaluating the best objective function for the application of linear optimization techniques is an ongoing area of research, though the objective of maximizing biomass has become generally accepted for determining the maximal growth rate for a cell in a given environment [91]. In his model, Watson suggests using "minimum free-energy dissipation" as the objective function. In considering his novel approach to metabolic modeling, Watson writes "there seems no reason in principle … to prevent the tailoring of the models to the species level if this is desired." Following a procedure that is now regarded as the standard FBA approach, Watson's approach involves listing reactions to build a stoichiometric matrix, and solving for the optimal flux through the network defined by the stoichiometric constraints with the Simplex method of linear programming. Notably, this approach uses matrix representation of stoichiometry; requires no net accumulation or depletion of intermediates; estimates relative contributions of competing pathways; and generates sequences of maps representing metabolic adaptations. Although Watson's application of the simplex algorithm is equivalent to the modern approach of FBA, his application was limited in scope, and it took some time to formalize and popularize this approach.

By 1986, Watson had expanded upon his earlier efforts to develop an instructional program to teach students about metabolic pathways [92]. Watson describes that it is common to consider molar yields of ATP from various pathways as an instructional tool to appreciate the

balance between reactions that provide and consume energy. Realizing that this idea of balancing production and consumption need not be limited to energy, he extended the idea of calculating yields which demonstrated the balanced consumption and production of other metabolites. However, as manual calculation of such balances quickly becomes a laborious process, he suggests the application of computer algorithms. As he writes, "what initially seemed to be a balancing of inputs against outputs has broadened out into a modeling procedure" [92]. In Watson's paper, he describes a program he has written for balancing 33 different metabolites. He conceives of metabolism as an optimization process in which an organism opens and closes pathways via regulation as it adjusts to its environment – a "continual search for new optimal states". Citing [93], Watson describes metabolism as "like an economy, in which chemical resources are automatically disposed to the best effect." Thus, perhaps considering the tools used by economists, Watson reasserts his suggestion that a suitable framework for an optimization and flow based model of cellular metabolism is linear programming. He describes how to construct such a model: 1) break the system up into a number of elementary functions which he calls "activities" – enzyme catalyzed reactions could be an example of an activity; 2) develop a statement of constraints; 3) select an objective function – as in his earlier paper, Watson again uses "minimal free energy loss" as his objective function; and 4) use the simplex algorithm to find an optimal solution.

In addition to describing his method, Watson offers a brief discussion of the philosophical difference between his modeling approach and earlier kinetic models: he considers his approach to be discrete while others are continuous. He suggests that computers can be

applied to problems such as modeling metabolism in two different ways: since they can rapidly perform computations, they can be applied to extending the "familiar calculations from chemical and enzyme kinetics to multi-enzyme systems"; alternatively, computers can be viewed as fundamentally new tools, which allow for new types of models. Watson considers his model to be fundamentally different, a model in which "transitions between configurations are governed by an objective function."

Applying optimization to metabolic pathways

In 1986, the same year that Watson elaborated upon his optimization-based approach, Fell and Small, who had previously published work on Metabolic Control Analysis (see "Kinetics of multi-reaction systems"), applied Watson's linear programming approach to examine stoichiometric constraints on fat synthesis in rat adipose tissue [94]. They describe Watson's approach as one which "selects the particular solution that minimizes some cost … or maximizes some yield … from amongst the possible solutions that are consistent with the requirements of balance". Using a metabolic simulation and control analysis software package called SCMP, they programmatically built the stoichiometry matrices to describe "over 50 reactions and intermediates, even after condensation of some of the processes into a single overall reaction." Significantly, Fell and Small note that with the application of linear programming, "it would also be possible to build in limitations on the flux allowed through individual steps to accommodate information on maximal enzyme activities." Perhaps because of Fell's significant previous work on enzyme control, this paper is more widely cited than Shapiro's pioneering 1969 work or Watson's seminal papers from 1984 and 1986.

The next publication of the optimization approach appeared in 1989, when in an effort to apply "the cybernetic perspective", Majewski and Domach used the linear programming approach to describe the phenomenon of acetate overflow observed in the metabolism of the bacterium *E. coli* [95]. This model represented something of a hybrid approach, since they use a lumped network formulation similar to the approach used by Papoutsakis and Meyer [85], but solve it with linear programming using an objective of maximal ATP synthesis and constraints on electron transport and the capacity of the Krebs cycle.

Applying optimization to cell-scale models

Between 1992 and 1994, the Palsson group (then at the University of Michigan) published a series of papers firmly establishing the application of linear optimization to stoichiometrically constrained networks as a valuable modeling technique for considering metabolic fluxes, including analysis of fluxes in intermediary metabolism. These papers began by describing the mathematical framework of this approach, moved to development of models of hybridoma cells and the metabolism of *E. coli*, and cumulated in a 1994 review article describing the practical application of "Metabolic Flux Balancing", which came to be known as flux balance analysis (FBA).

These papers mark the first application of Watson's approach of applying the simplex algorithm to find an optimal to a specific organism at a scale encompassing multiple portions of metabolism, with added emphasis on features of linear programming that Watson did not address, such as the concepts of shadow prices and reduced costs. In addition to finding theoretical maximum growth rates of hybridoma cells on amino acids, they use the shadow

prices to identify the relative value of cofactors and metabolic intermediates, thus finding which compounds cause the growth to be limited. The first paper [96] describes the mathematical basis for the approach. This paper develops the modern nomenclature of flux balance analysis, and was the first comprehensive assessment of the value and limitations of applying linear optimization to modeling metabolic networks. The authors begin by stating the dynamic mass balance formulation for steady state metabolism:

$$\mathrm{S}\cdot v(x) = \mathbf{b} \qquad \text{2.a.1}$$

3

Where S is the stoichiometric matrix, *v(x)* is the vector of reaction rates, and **b** is the vector of consumption and production rates and biosynthetic fluxes. This equation is equivalent to equation 2.a.10. The authors make the distinction between kinetic and static portions of this equation, stating that "a large body of literature is available discussing the applications of this model equation to different systems, and the emphasis is usually placed on the enzyme mechanisms, given v(X), and the parameters found therin." In contrast, Savinell and Palsson chose to focus on the stoichiometric constraints embodied in S, seeking to "derive as much information as possible" with "no information on reaction rate expressions or kinetic parameters".

To justify the use of linear programming techniques, Savinell and Palsson comment upon the fact that the dynamic equations are underconstrained: "the stoichiometry of the metabolic network does not uniquely specify the fluxes through the cell's pathways." Thus, the challenge of applying linear programming techniques is to select an objective that is postulated to underlie

a cell's behavior. Noting a benefit of linear programming, the authors discuss two auxiliary quantities that are computed by the simplex algorithm that may be interpreted as biologically meaningful concepts: the shadow prices and the reduced cost. Shadow prices quantify the sensitivity of the objective function with respect to each constraint, and reduced costs can be interpreted as the sensitivity of the objective function with respect to "non-basic fluxes".

The first cell-scale model: hybridoma

Like previous modelers, Savinell and Palsson lumped some linear pathways of their reconstruction of the hybridoma metabolic network into single reactions in their 1992 work, though they did so to reduce computational overhead rather than to obviate the need for linear programming techniques. Cytoplasmic and mitochondrial pools were treated separately (that is, the model was compartmentalized), and the network consisted of 83 reactions and 42 metabolites. The model included the metabolism of monoclonal antibody synthesis, nucleotide synthesis, phospholipid synthesis, polysaccharide synthesis, and energy metabolism. In this early application of linear programming, Savinell and Palsson evaluated three objective functions: 1) optimize energy efficiency by minimizing ATP production; 2) optimize use of substrates by minimizing total nutrient uptake; and 3) optimize redox metabolism by minimizing NADH production. While they found that "no single objective function gave results that completely describe the actual behavior of the hybridoma cell, certain objectives do give results characteristic of particular aspects of hybridoma cell behavior" [97]. Specifically, they found that optimizing energy efficiency predicted glutamine utilization well, as well as a preference for amino acid uptake as a major nutrient source. Optimizing substrate use did not mimic key

characteristics of hybridoma metabolism, but minimizing NADH production predicted the reduced yields, low oxygen uptake, high glutamine uptake, and high alanine production observed in hybridoma cells. Significantly, the authors noted that applying linear optimization allowed them to "predict how the deletion of a gene for a particular enzyme, or the inhibition of an enzyme, would affect the cell's performance".

In the second paper [97], the authors examine the application of the linear optimization approach to the metabolism of the hybridoma cell in more detail, including comparison to experimental results obtained during exponential growth of a hybridoma culture. In this work, the authors used maximized growth rate as the objective function, and noted that the model predicted a higher maximal growth rate than the culture attained. They reasoned that the culture was limited by factors other than flux through the metabolic network, such as the rate at which DNA is replicated. However, they suggested that adding boundaries establishing the maximum and minimum allowable fluxes through various reactions would limit the maximum growth rate predicted by the model to better agree with measured rates. Another approach the authors pursued was to fix the growth rate at the experimentally measured value, and examine fluxes with a different objective function – in this case, the authors examined maximizing and minimizing ATP production. Using the ATP maximization function, the model predicted secretion of alanine at the measured growth rate, a result that had been observed *in vitro*.

By examining the shadow prices for intermediates at different model conditions, the researchers were able to evaluate which nutrients were "most valuable" in those conditions, and thus made predictions about limiting nutrients. The authors examined the effects of varying

glutamine uptake rates, glucose uptake rates, antibody production rates, and dissolved oxygen concentration. Together, these two papers firmly established linear optimization as a valuable tool for applying models based on stoichiometrically constrained networks.

Extending FBA: *E. coli*

The Palsson group built upon these early efforts at stoichiometric models with applications beyond the hybridoma cell. In a series of papers in 1993 and 1994, Varma and Palsson described their work building and evaluating a stoichiometrically constrained model of *E. coli*, which they evaluated using linear optimization techniques. In describing their approach, Varma and Palsson noted that stoichiometric analysis had the potential for both exploring the boundaries of achievable metabolic performance and "systematizing knowledge of metabolic systems" [98]. They noted that the linear optimization approach could provide "a rational basis for modifying cellular metabolism", which "is the subject of the recently defined field of metabolic engineering" [99] since it would allow the determination of the optimal flux for a desired goal, "such as the commercial production of a metabolite" [100], thereby identifying the important reactions which could be subjected to genetic engineering to achieve that goal. Varma and Palsson noted three limitations to the stoichiometric approach to metabolic modeling at that time: 1) there was no consideration of regulation of enzymes; 2) there was no accounting for concentrations of metabolites within the cell; and 3) there was no incorporation of thermodynamic information in the form of rate expressions - time is not a variable in the model, thus no predictions can be made with regard to time constraints of cellular processes.

Compiling a metabolic database

Since, as they noted at the time of their work, "no single source exists that compiles and documents the active metabolic pathways, with their associated stoichiometry, in *E. coli*" [98], they began by developing a "biochemical database" of comprehensive information about the biochemical pathways in the organism. They divided the reactions leading to synthesis of a bacterial cell into "fueling, biosynthetic, polymerization and assembly reactions". Their first paper [98] examined the ability of the flux balance model to produce individual biosynthetic precursors.

In this paper, they extracted information from the biochemical database to "formulate a stoichiometric model of the bacterial network of fueling reactions," which they reduced to 53 reactions and 30 metabolites by applying a set of rules to lump reactions. They used linear optimization to define the solution space of the under-determined system of equations represented by this mass balance system, and applied various objective functions to investigate the capabilities of this "reconstructed network" to maximally produce 12 biosynthetic precursors and three key cofactors.

Biomass formation

Following the approach described by Watson [92], Varma and Palsson next considered the balance of metabolite supply and demand. After demonstrating that a stoichiometrically constrainted network evaluated with linear optimization techniques could be used to model the production of biosynthetic precursors and cofactors, Varma and Palsson studied the ability of the reconstructed network to meet a balanced set of demands intended to simulate growth

requirements in their second paper [100]. Varma and Palsson defined an overall reaction to represent the growth requirements for *E. coli*:

$$\sum d_M \cdot M \overset{v_{gro}}{\rightarrow} biomass \qquad 2.a.1$$

4

where M represents biosynthetic precursors and cofactors, d_M represents the amount in which they are needed per unit of biomass, and v_{gro} is the flux to biomass. For the weightings represtend by d_M, Varma and Palsson used values previously established in studies on cellular composition. Although cellular biomass composition varies in response to changing environmental conditions (for example, membrane lipid composition alters in response to stress [101]), Varma and Palsson used a fixed biomass composition as the primary demand for metabolites. Varma and Palsson suggested that the optimal solution for the production of this defined "biomass" could be found with the linear optimization approach, thus establishing the use of a "biomass function" as a sink of metabolites, and the maximization of this flux as a possible objective function. By applying the optimization approach, Varma and Palsson computed a theoretical yield of biomass, then evaluated which factors constrained the growth through examination of shadow prices. Varma and Palsson observed that fluxes computed in the optimal solution for biomass yield differed from experimentally measured fluxes, and so performed an analysis to determine the sensitivity of the optimal solution to changes in fluxes, metabolic demands (ie, biomass composition), P/O ratio, and maintenance energy requirements. Their sensitivity analysis approach relied upon computing logarithmic sensitivity coefficients that correspond directly to Kacser and Burns' Control Coefficients (discussed in "Kinetics of

multi-reaction systems"). Varma and Palsson found that constraints on selected fluxes had little effect on the optimal biomass yield, nor did variation in requirements for any single precursor or cofactor, but varying the P/O ratio (a measurement of the number of high-energy phosphate bonds formed by the electron transfer system) and ATP maintenance requirements could have effects upon the maximum biomass yield. In comparing model performance to *in vivo* measurements, the authors report "remarkably close agreement".

Applying reconstructed metabolic models

In three more papers, Varma and Palsson applied the *E. coli* model to the metabolic engineering problem of predicting biochemical production capabilities [99], conducted a series of experiments for *in vivo* validation of quantitative flux predictions from the *in silico* model [102], and expanded upon the parametric sensitivity analysis [103]. In the latter paper, they simulated aerobic chemostat, batch, and fed-batch experiments, and found that the model was sensitive to parameters describing metabolic capacity and relatively insensitive to parameters describing metabolic requirements for growth. They noted the good fortune that the parameters that have higher sensitivities are more readily measured in *in vivo* experimental settings.

Maturation of the FBA approach

After demonstrating the utility of flux-balance models, Varma and Palsson published a review article summarizing the FBA approach, which requires only knowledge of metabolic stoichiometry, including energy and metabolite balances [104]. Their article defended an optimization approach using linear programming, but also noted approaches that depend upon

reducing the dimensionality of the stoichiometric matrix so that it can be explicitly solved. With the publication of this review article, the application of metabolic flux balancing to problems in the field of metabolic engineering was firmly established. As will be described in the next section, subsequent work included expanding metabolic network reconstructions, developing networks for other organisms, applying reconstructed metabolic network models to a variety of basic science and engineering problems, and on standardizing data formats for sharing information and reproducing previous results.

Genome-scale metabolic reconstructions

The process of manually reconstructing metabolic networks by listing reactions and representing that information in a stoichiometric matrix continued and advanced, including efforts to model a variety of organisms, including *S. cerevisiae.* Early reconstructions of the yeast metabolic network included a 1995 reconstruction with 81 reactions and 88 metabolites [105], and a 1997 reconstruction with 37 reactions and 43 compounds [106]. However, metabolic reconstructions of *E. coli*, were the most extensive reconstructions throughout the 1990s. A significant 1997 *E. coli* reconstruction had 289 metabolites and 300 reactions [107].

Automating reconstruction

In 1995, the first bacterial genome – that of *Haemophilus influenza* Rd - was sequenced [108], and biology transformed from "a data-poor to a data-rich environment" [109]. Stoichiometric reconstructions offered a tool to manage "the large volume of genome-scale data that is being produced and made available in databases on the World Wide Web" [110], and so

the emergence of genomic information almost immediately affected the process used to "reconstruct whole-cell metabolic networks for sequenced organisms" [111].

The emergence of genomic information enabled a new protocol for developing reconstructed metabolic networks [112]. Instead of starting with biochemistry texts to compile a metabolic database, a draft reconstruction can now be programmatically constructed from genome annotation and public databases, then refined through an iterative process of editing, supplementing with auxiliary data sources, and testing the draft reconstruction (Figure 2.a.4). This time-intensive process of "curating" a draft reconstruction necessarily forces close examination of metabolic function, and serves both to structure existing biochemical knowledge and to highlight areas needing additional basic research. A metabolic function required for viability may not be included in early genomic annotation, but the functional necessity may be revealed through metabolic network reconstruction. Thus, metabolic reconstruction can provide hypothesized functions for unannotated open reading frames. Efforts to improve automated the reconstruction of metabolic networks are ongoing, including the recent development of method of "comparative model reconstruction", in which a draft metabolic network is built from the reconstructed network of a related organism, using synteny between the species annotated genome sequences [113].

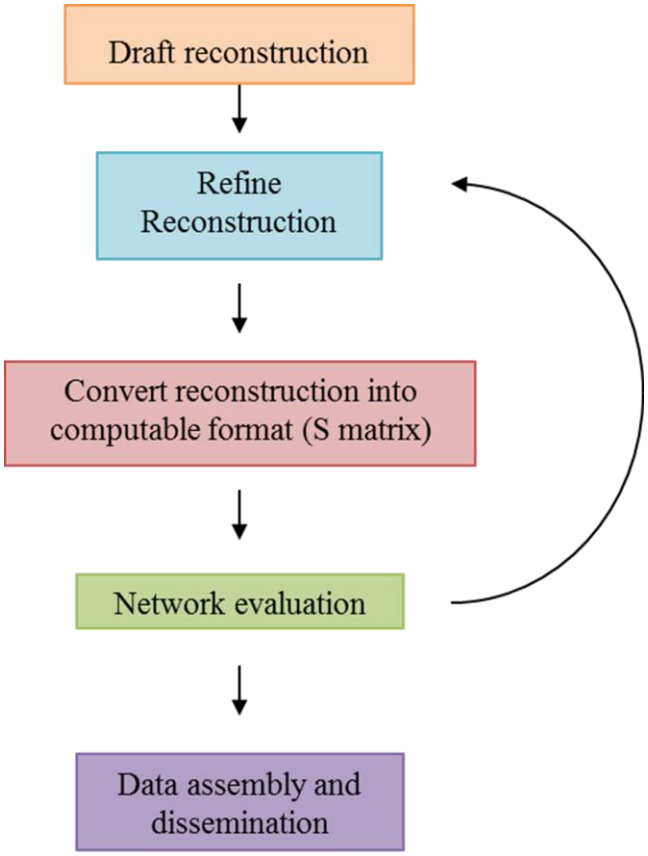


Figure 2.a.4: Overview of metabolic reconstruction procedure.

With the emergence of annotated genomic information and public databases, a Draft reconstruction can now be generated programmatically. Note that stages 2-4 are iterated until model predictions are similar to observed phenotypes or data sources are exhausted. Figure based upon [103].

Yeast genome-scale reconstructions

The first metabolic reconstruction which applied the protocol of building a draft reconstruction from annotated genomic information was published in 1999, and constituted the

first “genome scale” metabolic reconstruction [114]. This model, of *Haemophilus influenza* Rd, included 488 reactions and 343 reactions and demonstrated that “the synthesis of *in silico* metabolic genotypes from annotated genome sequences is possible” [114]. The construction of the metabolic model directly contributed to the genomic research because FBA methods were used to analyze pathways and the reconstruction was “used to reconcile and curate the sequence annotation by identifying reactions whose function was not supported” and “to predict gene products that should be co-regulated and perhaps co-expressed” [115]. The model was applied to engineering problems by predicting necessary constituents of minimal media, and essential genes were identified through an *in silico* knockout procedure (ibid).

With this demonstration of applying annotated genomic information as the basis for reconstructing metabolic networks, the method was applied to other organisms as the sequences became available, including *E. coli* strain MG1655 [116] and, in 2003, the first genome-scale reconstruction of a eukaryote, *Saccharomyces cerevisiae* [117]. Following a naming convention proposed in 2004 [118], this first genome-scale reconstruction of the yeast metabolic network came to be known as iFF708, since it is an *in silico* model developed by Forster and Famili, and includes reactions catalyzed by 708 gene products. iFF708 included 1175 reactions and 584 metabolites, and accounted for ~16% of the characterized open reading frames of the yeast genome at that time. The authors of iFF708 constructed the model from genomic, biochemical, and physiological information sourced from pathway databases, biochemistry textbooks, the annotated genome, and relevant literature through a “nonautomated and iterative decision-making process” [117]. To enable FBA analysis, iFF708 included a biomass composition (Table

2.1) that defined a drain of metabolites from the system. FBA was applied to the model and used to predict growth phenotypes. FBA produced *in silico* results that "were consistent with observed phynotypic functions for ~70-80% of the conditions considered" [119]. Subsequently, a large-scale evaluation of the effect of 599 single gene deletions was performed with this model [120]. In 87.8% of the cases, the *in silico* results were in agreement with published experimental observations. The authors suggested that reasons for false predictions of knockout results included errors in 1) model media composition; 2) substituted biomass components; 3) incomplete biochemical information; and 4) the lack of regulatory information in the model.

As the first genome-scale reconstruction of a eukaryotic metabolic network, iFF708 has proven to be a very useful model, and continues to be applied to research questions (see [121], [122], and [123] for examples of recent applications of this model). It has also served as the basis for a series of updated and expanded genome-scale reconstructions of yeast metabolism (Table 2.2).

Though this document focuses on the application of reconstructed yeast metabolic networks, the process of developing reconstructed networks for other species has continued concurrently. By 2009, there were more than 50 reconstructions, including representative species from the domains Eukaryota, Bacteria, and Archaea [124]. Reconstructions of plant metabolic networks have been underrepresented, but recent reconstructions have been published for *Arabidopsis* [125] and Barley Seeds [126].

Table 2.1: iFF708 biomass composition.

Reproduced from [117].

Metabolite	mmo	Metabolite	mmole
Amino acids [127]		**Ribonucleotides** [127]	
Alanine	0.45	AMP	0.046
Arginine	0.16	CMP	0.045
Asparagine	0.10	GMP	0.046
Aspartate	0.29	UMP	0.060
Cysteine	0.00		
Glutamine	0.10	**Deoxyribonucleotides** [128]	
Glutamate	0.30	DAMP	0.004
Glycine	0.29	DCMP	0.002
Histidine	0.06	DGMP	0.002
Isoleucine	0.19	DTMP	0.004
Leucine	0.29		
Lysine	0.28	**Lipids** [129],	
Methionine	0.05		
Phenylalanine	0.13	**Sterols** [130]	
Proline	0.16		
Serine	0.18	Triacylglycerol	0.007
Threonine	0.19	Ergosterol	0.001
Tryptophane	0.02	Zymosterol	0.002
Tyrosine	0.10	Phosphatidate	0.001
Valine	0.26	Phosphatidylcholine	0.006
		Phosphatidylethanolamine	0.004
Carbohydrates [132]		Phosphatidylinositol	0.005
Glycogen	0.51	Phosphatidylserine	0.002
Trehalose	0.02		
Mannan	0.80		
Other	1.13		

Emerging yeast reconstructions

As the authors of iFF708 noted, "the *in silico* model-building procedure at the genome scale is an iterative and ongoing process during which the content of a model grows, and thus the

scope of properties that can be computed widens" [119]. Subsequent researchers embraced this iterative approach, and efforts to expand the reconstruction of yeast metabolism led to the release of many new models as groups released reconstructions based on iFF708 (Table 2.2). The first model based on iFF708 was iND750 [118]. iND750 expanded some of the lumped reactions from iFF708 to account for individual reactions, thus accounting for more genes (750 genes, compared to iFF708's 708). iND750 expanded iFF708's compartmentalization from three compartments (cytosol, mitochondria, and extracellular space) to eight compartments by locating reactions in five additional compartments (peroxisome, nucleus, golgi apparatus, vacuole, and endoplasmic reticulum). The addition of these compartments required an increased inclusion of transport reactions – iND750 added 297 new transport reactions to those derived from iFF708. iND750 also added explicit association between genes and reactions in the form of a gene-protein-reaction association table. However, despite the greater number of compartments, improved annotation, and increased number of metabolites and reactions, the lethality predictions for gene knockout simulations was lower than those of iFF708 (iND750 reported 82.6% agreement with large-scale knockout studies, and iFF708 reported 87.8% agreement [120]).

In an effort to improve the agreement, and to apply FBA to questions of gene duplication function, Kuepfer et al. revised iFF708 to produce another model, iLL672 [133]. iLL672 improved the ability to predict viability of single knockouts with FBA to "95%-98%" through a combination of reformulating biomass composition, removing 110 "dead-end metabolites" which were not connected to other portions of the metabolic reconstruction, and adding reactions based

on new biochemical knowledge to connect 33 other metabolites. Along with other modifications, the improved biomass definition of iLL762 was incorporated to iND750 to form an updated model, iMM904 [134]. In revising iND750 to create iMM904, the authors also revisited gene-protein-reaction associations, updated reactions with information based upon new biochemical discoveries, and expanded lipid, transport, and carbohydrate subsystems.

From this proliferation of yeast metabolic reconstructions, a group of 34 researchers worked to build a common, "consensus" network reconstruction, which came to be Yeastnet version 1.0 [135]. The Yeastnet reconstruction combined iMM904 and iLL672, integrating them with common metabolite names and reaction description nomenclature. To avoid ambiguous metabolite names, a significant effort was expended to annotate molecules in the model with references to persistent databases and other external resources. As the goal of the Yeastnet reconstruction was to provide a common inventory of metabolites and network structure as a shared resource, it did not include a biomass function, leaving the biomass definition to researchers who hoped to evaluate the Yeastnet reconstruction with FBA. Yeastnet v1 was limited by the network reconstructions that it was based upon. Since neither iMM904 nor iLL672 included comprehensive descriptions of lipid pathways, Yeastnet did not either.

In an attempt to correct this deficiency, another metabolic reconstruction was created as a "scaffold to query lipid metabolism. This model, based upon iFF708, was called iIN800 [136]. Like the iLL672 reconstruction, iIN800 included a refined biomass function in addition to its greatly expanded description of lipid pathways, particularly the biosynthesis and elongation of fatty acids. It also included expanded representation of sphingolipid synthesis, but did not

differentiate between very long-chain fatty acid moieties (see Chapter 3). iIN800 also expanded coverage of reactions involving esterification of sterols, degradation of lipids. iIN800 included 17.2% of the characterized open reading frames in the yeast genome at the time of its publication, and was the most comprehensive reconstruction at that time.

The updates of iIN800 were integrated to the Yeastnet model to create Yestnet version 2.0 [137]. Yeastnet has continued to be updated, and the current release of this consensus reconstruction is Yeastnet version 4.02 [138]. Modifications that I have suggested as a result of validation of the Yeastnet version 3.0 reconstruction of sphingolipid metabolism will be integrated into an upcoming release, with the addition of 136 new reactions (the results of the validation work are discussed in Section 4, "Preliminary results"). The Yeastnet reconstruction of lipid metabolism has not yet been fully validated, but with the recent documentation of 120 distinct phospholipids in yeast [139], it is likely that further expansion will be necessary.

Table 2.2: Reconstructed yeast metabolic networks.

Following the first genome-scale reconstruction, iFF708, researchers embraced an iterative process of updating models based upon previous reconstructions, a process which led to the formation of many models. These models were re-integrated to form the Yeastnet model, which is centrally curated and accepts suggestions for refinements from interested research groups

Model Name	Reactions	Chemical Species*	ORFs	Compartments	Based upon	Reference
-	81	88		1		[105]
-	99	98				[140]
-	37	43		3		[106]
-						[141]
-	64	67				[142]
iFF708	1175	733	708	3		[119]
iND750	1489	646	750	8	iFF708	[118]
iLL672	1038	636	672	2	iFF708	[133]
iMH805/775	1489	646	805	8	iND750	[143]
iJH732	1152	935	732	2	iND750	[144]
iIN800	1446	1013	800		iFF708	[136]
iMM904	1412	1228	904	8	iND750	[134]
Yeastnet	1857	2153	832	15	iLL672 iMM904	[135]
Yeastnet 2.0	2576	2491		16	Yeastnet iIN800	
Yeastnet 3.0	2830	3057		16	Yeastnet 2.0	
Yeastnet 4.0	2342	2657		16	Yeastnet 3.0	
iAZ900	1597	1398	900	8	iMM904	[145]
Yeastnet 5.1	2110	1655	918	16	Yeastnet 4.0 iAZ900	[146]
Yeastnet 5.30	2051	1602	904	16	Yeastnet 5.1	Chapter 6

Applications of yeast metabolic models to metabolic engineering

To date, the majority of papers relating to stoichometrically constrained models of yeast metabolism have provided commentary, model development, or other theoretical approaches. Despite the proliferation of yeast models and review papers discussing their development, there have been comparatively few reports of successful efforts to apply such models to *in vivo* metabolic engineering efforts. These efforts include: an effort to manipulate redox metabolism to increase ethanol yield [147], development of a gene-deletion strategy to manipulate C1 metabolism [148], and an effort to improve the production of sesquiterpenes [149]. The relative paucity of successful attempts to apply these models to metabolic engineering has led some to question whether the underlying method of flux balance analysis can describe eukaryotic metabolism as well as it works for simpler organisms [145, 150]. However, since yeast model development is an ongoing effort, this question seems premature.

Challenges encountered reconstructing the yeast metabolic network

Flux balance analysis of stoichiometrically constrained metabolic models requires three components: accurate reconstruction of the biochemical reaction network of metabolism; a set of constraints on the reactions in the network; and an objective function to be maximized or minimized via linear programming. Each of these components has provided challenges to researchers working to improve computational models of yeast metabolism, and there remains opportunity to refine each component to improve model predictions and applicability.

Work presented in this thesis documents that review of published literature is an effective tool for improving both the reconstruction of the known yeast biochemical reaction network

(Chapters 3 and 4) and the constraints on that network (Chapter 5). The sensitivity of such models to reaction constraints has also been reported previously [151]. Additionally, many authors of reconstructions of yeast metabolic networks have commented upon the challenges and importance of accurately defining the biomass function. Two significant challenges in creating the biomass definition are 1) biomass composition changes in response to different environmental conditions (such as stress or nutrient deficiency), and 2) in some cases, organisms can substitute one chemical species for another to form a biomass composition. Since the components included in a biomass definition are considered a fixed requirement for growth in *in silico* simulations, each component must be produced by the metabolic network. There exist deletion mutants *in vivo* that are not able to produce an original biomass precursor, but the strains are viable because "other metabolites can replace these initial precursors or building blocks" [120], an effect which leads to incorrect viability predictions *in silico.* Examples of this problem include the substitution of phosphatidylethanolamine for phosphatidylcholine in lipid metabolism [120], and the use of fatty acids of different lengths as membrane constituents [136]. As the authors of iIN800 stated to justify their use of different biomass functions for different growth regimes (carbon or nitrogen limited), "more accurate biomass compositions lead to improved leathality predictions" (ibid).

Authors have taken different approaches to dealing with the challenges of accurately reconstructing the biochemical network, appropriately applying constraints, and selecting an appropriate biomass definition for the objective function. The primary approach to developing the biochemical reaction network is to refine the network of a previously existing model. Thus,

every genome-scale stoichiometrically constrained yeast metabolic model to date is derived from iFF708. However, a variety of gap-filling and computational procedures have been applied to improve model lethality predictions by adding reactions from other sources [145, 152, 153]. Similarly, computational approaches have been applied to modify constraints [154]. The differing constituents of biomass in different growth conditions are addressed by modeling a specific period of the cell cycle – usually exponential growth – and using carefully defined experimental results as the basis for model validation. Model developers have also taken different approaches to dealing with biomass constituents which may be substituted. In some cases, "Available experimental data often report the average composition of these compounds" [112] and so the model represents average compounds, rather than all possible compounds. In other cases, the model is simplified and certain biomass components, such as chitin and glycoprotein, are neglected [117]. Similarly, sphingolipid metabolism was not completely elucidated, and was not included in iFF708 [117], though "ceramide" was considered as two moieties, using C24:0 and C26:0 very long-chain fatty acids, in iND750 [118]. When Yeastnet was constructed, lipid pathways were lumped, pending further elucidation of the pathway [135]. The representation of these pathways was expanded for iIN800, which expanded reactions in fatty acid synthesis, elongation, activation and □-oxidation, sphingolipid synthesis, ergosterol esterification, and lipid degradation [136]. These reactions were incorporated into Yeastnet to create version 2 [137]. However, no currently published reconstruction of the yeast metabolic network includes a full representation of the current understanding of yeast lipid and sphingolipid metabolism.

Software tools for working with existing genome-scale reconstructed metabolic networks

The increase in model complexity and size that accompanied the availability of genome-scale data sets created new problems for researchers. Previously, smaller scale models published in peer-reviewed journals had been accompanied by instructions for recreating the work, which enabled easy review and application of published models. However, this approach suffered from a variety of drawbacks: models published in this fashion became unusable when simulation software was no longer supported, transferring models between different simulation or analysis tools was difficult, and subsequent users needed to obtain the same modeling environment as the original authors, making such models difficult to examine, test, and reuse [155]. The approach of publishing all the reactions described by a model in a journal paper became unworkable with the emergence of genome-scale reconstructed metabolic networks. These problems were addressed with the creation of syntax for encoding computational models of biochemical reaction networks, the Systems Biology Markup Language (SBML).

Sharing models: SBML

In Spring 2000, a group of researchers were invited to Caltech for a workshop entitled "Platforms for Systems Biology" [156]. The workshop was held to discuss the development of a software framework to facilitate sharing and evaluation of the computational models and approaches in the genomic era. A their first meeting in April 2000, the workshop participants decided to develop a simple computer language for representing and exchanging models between software tools. This language, which was to be based upon the eXtensible Markup Language (XML) [157], came to be known as the Systems Biology Markup Language (SBML) [155].

SBML is a structured text-based, machine-readable format that can be used to describe models and exchange data between various modeling software applications. SBML has since become "the de facto standard format for representing formal, quantitative and qualitative models at the level of biochemical reactions and regulatory networks" [158].

Using SBML: LibSBML

The establishment of SBML standardized the data format for sharing biochemical reaction network models. With this standard, software could be developed to allow a wide variety of applications to translate between SBML and their native data structures. Since SBML is based on XML, software developers could write custom XML parsers to achieve this task; however, an application programming interface (API) would facilitate use of SBML data without requiring low-level parsing of the information. This function was published with the release of LibSBML, an API "for reading, writing, manipulating and validating content expressed in the SBML format" [159]. LibSBML is written in C and C++, and includes bindings for a variety of other languages, including MATLAB. This binding allows users to write programs that read a SBML document into a MATLAB data structure for further analysis or simulation.

SBML with MATLAB: SBML and COBRA toolboxes

The commercially available software package MATLAB (http://www.mathworks.com) provides a robust computer environment for mathematical analysis, including graphical output. Core MATLAB functionality may be extended with toolboxes, which consist of bundles of short programs, or functions, that support a similar task (e.g. the Optimization Toolbox, the Statistics

Toolbox, and others). In an effort to better support the application of MATLAB to models published with SBML, researchers who had been involved in the development of LibSBML released the SBML Toolbox [160]. The SBML Toolbox was designed to provide the ability to read SBML files to a MATLAB data structure called MATLAB_SBML (Figure 2.a.5) and to write SBML files from that data structure. The SBML Toolbox is intended to serve as "a facilitator for the development of other functions and toolboxes" [160], and so only limited analytical functionality is included.

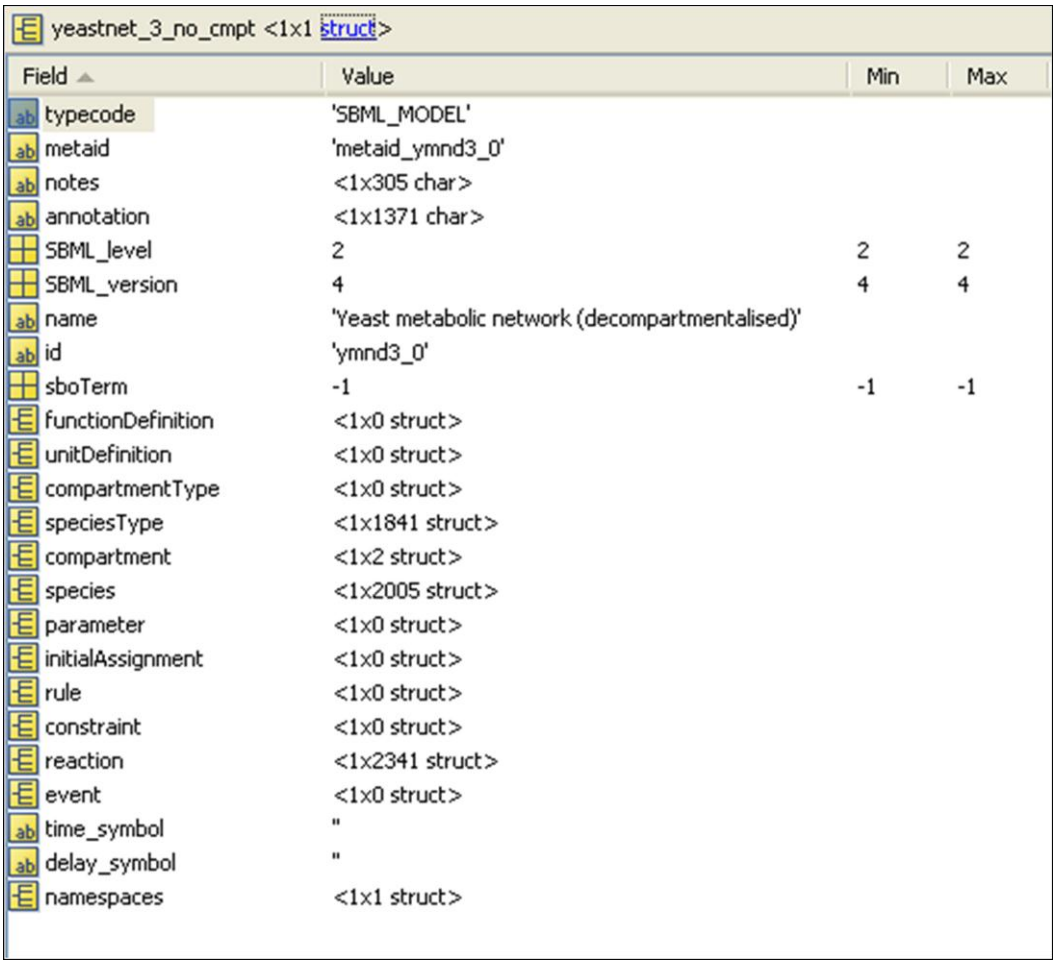
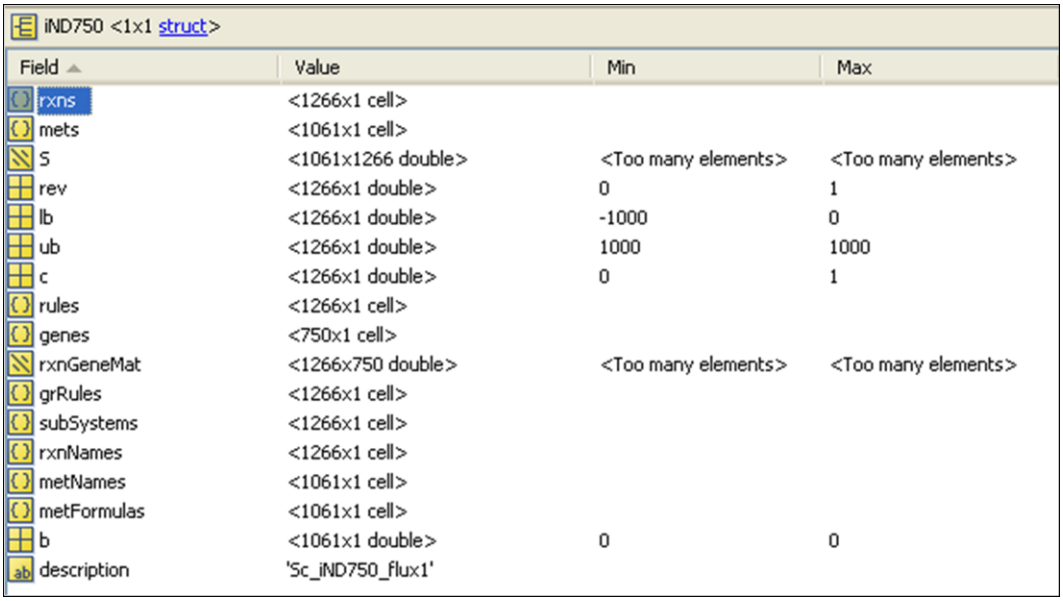


Figure 2.a.5: The MATLAB_SBML data structure.

The MATLAB-SBML structure is a native MATLAB structure, created by reading a SBML file with the SBML Toolbox.

The SBML Toolbox enables programmatic manipulation and analysis of reconstructed metabolic networks in MATLAB. However, to conduct FBA, the list of reactions and species in the MATLAB_SBML data structure must be converted to the stoichiometric matrix and mathematical formulation of constraints required for the linear programming approach. Programs

to conduct this translation, along with a variety of analysis tools, were released as the COBRA Toolbox [161]. The acronym COBRA is derived from a description of the toolbox's purpose: COnstraint Based Reconstruction and Analysis. The COBRA Toolbox includes a variety of functions, including readcbmodel.m. readcbmodel.m uses the SBML Toolbox to read a SBML model, and translates it to a structure containing the S matrix, constraints and objectives for linear programming, and other information (Figure 2.a.6).



iND750 <1x1 struct>

Field	Value	Min	Max
rxns	<1266x1 cell>		
mets	<1061x1 cell>		
S	<1061x1266 double>	<Too many elements>	<Too many elements>
rev	<1266x1 double>	0	1
lb	<1266x1 double>	-1000	0
ub	<1266x1 double>	1000	1000
c	<1266x1 double>	0	1
rules	<1266x1 cell>		
genes	<750x1 cell>		
rxnGeneMat	<1266x750 double>	<Too many elements>	<Too many elements>
grRules	<1266x1 cell>		
subSystems	<1266x1 cell>		
rxnNames	<1266x1 cell>		
metNames	<1061x1 cell>		
metFormulas	<1061x1 cell>		
b	<1061x1 double>	0	0
description	'Sc_iND750_flux1'		

Figure 2.a.6: The COBRA toolbox data structure.

The COBRA structure is a MATLAB data structure translated from the MATLAB_SBML structure by the COBRA toolbox.

The COBRA Toolbox also includes code for adding and removing reactions from a reconstructed metabolic network, printing reaction formulas, analyzing network robustness, conducting gene deletion simulations, and other functions. The COBRA Toolbox supports the

use of different linear programming solvers for conducting FBA and dynamic FBA, which is used to simulate batch growth. Supported solvers include the GNU Linear Programming Kit (GLPK) [162], a free and open-source program for "solving large-scale linear programming (LP), mixed integer programming (MIP), and other related problems" [162].

Software dependencies

In summary, conducting FBA on publicly available reconstructed metabolic networks with MATLAB requires a variety of interdependent software tools (Figure 2.a.7).

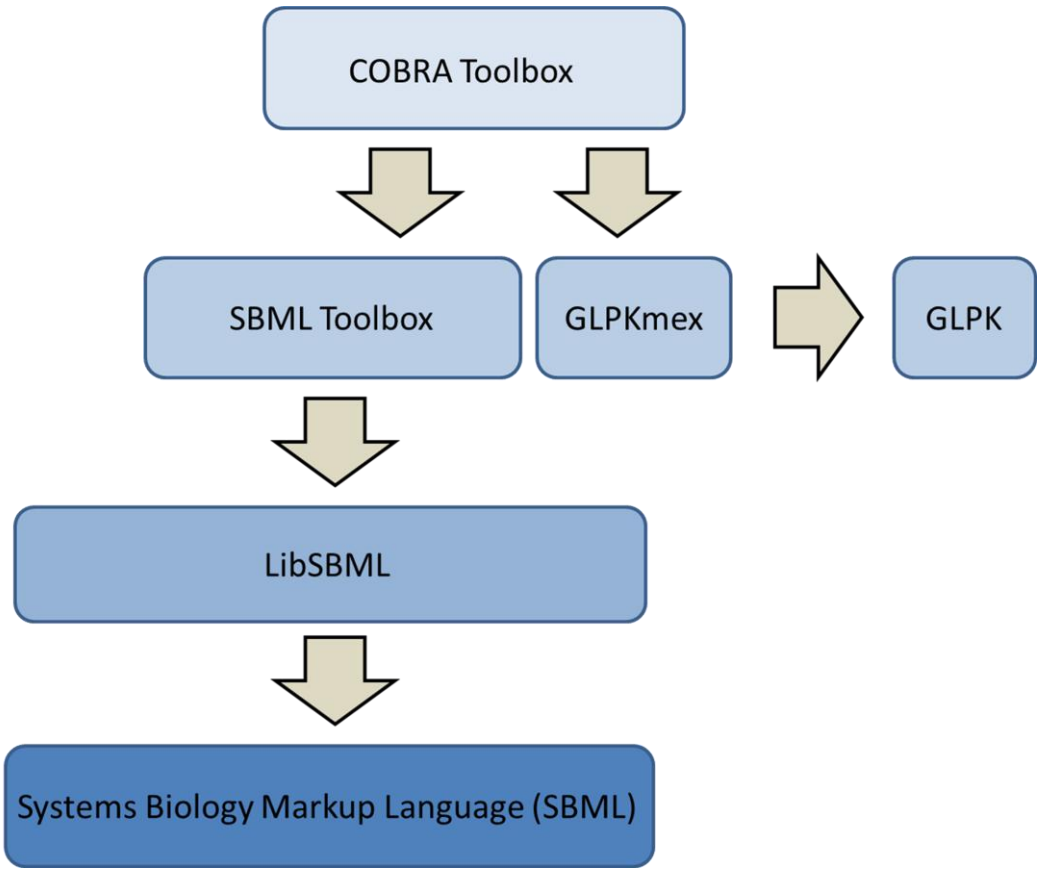


Figure 2.a.7: Software required for FBA with MATLAB.

The COBRA Toolbox uses functionality of the SBML Toolbox and the MATLAB interface to the GLPK linear programming solver (GLPKmex). In turn, GLPKmex depends upon GLPK, and the SBML Toolbox uses LibSBML to translate the SBML data format into a MATLAB data structure for manipulation and analysis.

A challenge arises because each of these software components (MATLAB, the SBML standard, LibSBML, the SBML Toolbox, GLPK, and the COBRA Toolbox) are maintained by different groups, and are thus updated at different times. As different versions are released, care

must be taken to ensure interoperability of the software components.

Summary

Industrial biofuel production would benefit from yeast strains that could more rapidly produce ethanol in high concentrations. Rational design of such strains through a genetic engineering approach benefits from the synthesis of biological knowledge into computational tools that can be applied for hypothesis generation, analysis of optimal genetic intervention, and to predict strain performance and media requirements. Knowledge of the biology of the fermentative yeast *Saccharomyces cerevisiae* and computer modeling techniques have now advanced sufficiently to merit the proposed investigation of applying rational design approaches devised and tested through the use of FBA to the improvement of ethanol stress tolerance of *S. cerevisiae*.

In vivo investigations of *Saccharomyces cerevisiae's* response to ethanol stress have come to focus on the dynamic process by which the organism remodels its plasma membrane. Early observations that the ratio of saturated and unsaturated fatty acids in the membrane have been refined to focus on the changing chemical species of phospholipid and other membrane constituents. PI has been identified as a key species related to ethanol tolerance, and the precursor inositol has been found to enhance ethanol tolerance of yeast when added to the medium exogenously or via biosynthesis stimulated through the deletion of the *OPI1* gene, which creates strains with the Opi^- phenotype. However, screens of deletion libraries have found that the Opi^- phenotype is also the result of a variety of other genetic interventions. Any of these other interventions may also increase ethanol tolerance – perhaps more so than the deletion of

OPI1 - but the effects or these interventions on factors such as growth rate and ethanol productivity have not been well documented. Additionally, the overproduction of other membrane constituents, such as complex sphingolipids, and other compounds such as trehalose, glycerol, and ergosterol may further increase ethanol tolerance.

Systems-level computational models are now available which can be used to evaluate the consequences of redirecting metabolic flux upon growth rate and ethanol productivity. These models have arisen following a century of attempts to understand and model metabolism, beginning with early efforts at characterizing the kinetics of individual enzymes. Attempts to model multi-enzyme systems gave rise to questions of metabolic control that led to a paradigm shift away from the reductionist approaches of enzymology to a systems-level approach to modeling metabolism. This approach incorporated modeling techniques from fields such as sensitivity analysis and parameter estimation, control theory, and cybernetics. By the 1980s, the explosion of data available in the post-genome age, enabled whole-cell (or "genome-scale") reconstruction of metabolic networks, which could be analyzed with a variety of approaches, including FBA. FBA is an optimization-based technique for analysis of metabolic fluxes arising from stoichiometric constraints on the reaction network, as well as other constraints which may be imposed upon the system. A significant benefit of applying FBA to reconstructed metabolic networks is that it is a modeling approach which can provide results without detailed kinetic data for each reaction included in the model (see "Stoichiometric modeling").

REFERENCES

1. Attfield PV: **Stress tolerance: The key to effective strains of industrial baker's yeast**. *Nat Biotechnol* 1997, **15**:1351–1357.

2. Hayashida S, Feng DD, Hongo M: **Function of the High Concentration Alcohol-producing Factor**. *Agricultural and Biological Chemistry* 1974, **38**:2001–2006.

3. Hayashida S, Feng DD, Hongo M: **Physiological properties of yeast cells grown in the proteolipid-supplemented media**. *Agricultural and Biological Chemistry* 1975, **39**:1025–1031.

4. Hayashida S, Feng DD, Ohta K, Chaitiumvong S, Hongo M: **Compositions and a role of Aspergillus oryzae-proteolipid as a high concentration alcohol-producing factor**. *Agricultural and Biological Chemistry* 1976, **40**:73–78.

5. Hayashida S, Ohta K: **Effects of phosphatidylcholine or ergosteryl oleate on physiological properties of Saccharomyces sake**. *Agricultural and Biological Chemistry* 1980, **44**:2561–2567.

6. Thomas DS, Rose AH: **Inhibitory effect of ethanol on growth and solute accumulation by Saccharomyces cerevisiae as affected by plasma-membrane lipid composition**. *Arch. Microbiol.* 1979, **122**:49–55.

7. Chi Z, Kohlwein SD, Paltauf F: **Role of phosphatidylinositol (PI) in ethanol production and ethanol tolerance by a high ethanol producing yeast**. *Journal of Industrial Microbiology and Biotechnology* 1999, **22**:58–63.

8. Furukawa K, Kitano H, Mizoguchi H, Hara S: **Effect of cellular inositol content on ethanol tolerance of Saccharomyces cerevisiae in sake brewing**. *Journal of Bioscience and Bioengineering* 2004, **98**:107–113.

9. Krause EL, Villa-Garcia MJ, Henry SA, Walker LP: **Determining the effects of inositol supplementation and the opi1 mutation on ethanol tolerance of Saccharomyces cerevisiae**. *Industrial Biotechnology* 2007, **3**:260.

10. U.S. Department of Energy: *Breaking the Biological Barriers to Cellulosic Ethanol: A Joint Research Agenda. A Research Roadmap Resulting from the Biomass to Biofuels Workshop, December 7–9, 2005, Rockville, Maryland*. Rockville, MD: 2006.

11. Yano T, Kurokawa M, Nishizawa Y: **Optimum substrate feed rate in fed-batch culture with the DO-stat method**. *Journal of Fermentation and Bioengineering* 1991, **71**:345–349.

12. Tchobanoglous G, Metcalf & Eddy: *Wastewater engineering : treatment and reuse*. 4th ed. /. Boston: McGraw-Hill; 2003.

13. Alvarez-Vasquez F: **Integration of kinetic information on yeast sphingolipid metabolism in dynamical pathway models**. *J Theor Biol* 2004, **226**:265–291.

14. Szallasi Z: *System modeling in cell biology : from concepts to nuts and bolts*. Cambridge Mass.: MIT Press; 2006.

15. Minton AP: **How can biochemical reactions within cells differ from those in test tubes?** *Journal of Cell Science* 2006, **119**:2863–2869.

16. Cornish-Bowden A: **Putting the systems back into systems biology**. *Perspectives in biology and medicine* 2006, **49**:475–489.

17. Buchner E: **Alkoholische Gährung ohne Hefezellen**. *Ber. Dtsch. Chem. Ges.* 1897, **30**:117–124.

18. Kohler R: **The background to Eduard Buchner's discovery of cell-free fermentation**. *Journal of the History of Biology* 1971, **4**:35–61.

19. Kühne W: **Über das Verhalten verschiedener organisirter und sog. ungeformter Fermente**. *Verhandlungen des naturhistorisch-medicinischen Vereins zu Heidelberg* 1877, **1**:190–193.

20. Blackman F: **Optima and Limiting Factors**. *Annals of botany* 1905, **19**:281–295.

21. Cornish-Bowden A: *Fundamentals of enzyme kinetics*. 3rd ed. London: Portland Press; 2004.

22. Henri V: **Théorie générale de l'action de quelques diastase**. *Comptes rendus de l'Académie des sciences* 1902, **135**:916–919.

23. Brown AJ: **Enzyme action**. *J. Chem. Soc., Trans.* 1902, **81**:373.

24. Michaelis L, Menten ML: **Die kinetik der invertinwirkung**. *Biochemische Zeitschrift* 1913, **49**:333–369.

25. Briggs GE, Haldane JB.: **A note on the kinetics of enzyme action**. *Biochemical journal* 1925, **19**:338.

26. Bohr C: **Theoretische behandlung der quantitativen verhältnisse bei der sauerstoff aufnahme des hämoglobins**. *Zentralblatt Physiol* 1903, **17**:682–688.

27. Hill AV: **The possible effects of the aggregation of the molecules of hæmoglobin on its dissociation curves**. *Journal of Physiology* 1910, **40**:iv–vii.

28. Monod J, Wyman J, Changeux JP: **On the Nature of Allosteric Transitions: A Plausible Model**. *J. Mol. Biol* 1965, **12**:88–118.

29. Koshland DE, Némethy G, Filmer D: **Comparison of experimental binding data and theoretical models in proteins containing subunits**. *Biochemistry* 1966, **5**:365–385.

30. Koch AL, Putnam FW, Evans EA: **The purine metabolism of Escherichia coli**. *Journal of Biological Chemistry* 1952, **197**:105.

31. Abelson PH, Bolton ET, Aldous E: **Utilization of carbon dioxide in the synthesis of proteins by Escherichia coli. II**. *J. Biol. Chem* 1952, **198**:173–178.

32. Gots JS, Chu EC: **Studies on purine metabolism in bacteria. I. The role of p-aminobenzoic acid**. *J. Bacteriol* 1952, **64**:537–546.

33. Adelberg EA, Umbarger HE: **Isoleucine and valine metabolism in Escherichia coli. V. alpha-Ketoisovaleric acid accumulation**. *J. Biol. Chem* 1953, **205**:475–482.

34. Umbarger HE: **Evidence for a Negative-Feedback Mechanism in the Biosynthesis of Isoleucine**. *Science* 1956, **123**:848.

35. Higgins J: **Dynamics and Control in Cellular Systems**. In *Control of Energy Metabolism*. New York: Academic Press; 1965, **1**:13–46.

36. Kacser H, Burns JA: **The control of flux**. *Symposia of the Society for Experimental Biology* 1973, **27**:65–104.

37. Heinrich R, Rapoport TA: **A linear steady-state treatment of enzymatic chains. General properties, control and effector strength**. *Eur. J. Biochem* 1974, **42**:89–95.

38. *Systems Biology*. Berlin/Heidelberg: Springer-Verlag; 2005, **13**.

39. Burns J, Cornishbowden A, Groen A, Heinrich R, Kacser H, Porteous J, Rapoport S, Rapoport T, Stucki J, Tager J: **Control analysis of metabolic systems**. *Trends in Biochemical Sciences* 1985, **10**:16–16.

40. Schuster S: **The definitions of metabolic control analysis revisited**. *Biosystems* 1992, **27**:1–15.

41. Fell D: *Understanding the control of metabolism*. 1st ed. Brookfield VT: Portland Press; 1997.

42. Kohn M, Chiang E: **Metabolic network sensitivity analysis**. *Journal of Theoretical Biology* 1982, **98**:109–126.

43. Giersch C: **Control analysis of metabolic networks. 1. Homogeneous functions and the summation theorems for control coefficients**. *Eur J Biochem* 1988, **174**:509–513.

44. Kacser H: **The control of enzyme systems in vivo: elasticity analysis of the steady state**. *Biochem. Soc. Trans* 1983, **11**:35–40.

45. Fell DA, Sauro HM: **Metabolic control and its analysis. Additional relationships between elasticities and control coefficients**. *Eur. J. Biochem* 1985, **148**:555–561.

46. Hofmeyr JH, Kacser H, van der Merwe KJ: **Metabolic control analysis of moiety-conserved cycles**. *Eur. J. Biochem* 1986, **155**:631–641.

47. Sauro HM: **Moiety-conserved cycles and metabolic control analysis: problems in sequestration and metabolic channelling**. *BioSystems* 1994, **33**:55–67.

48. Sauro HM, Small JR, Fell DA: **Metabolic control and its analysis. Extensions to the theory and matrix method**. *Eur. J. Biochem* 1987, **165**:215–221.

49. Small JR, Fell DA: **The matrix method of metabolic control analysis: its validity for complex pathway structures**. *J. Theor. Biol* 1989, **136**:181–197.

50. Reder C: **Metabolic control theory: A structural approach**. *Journal of Theoretical Biology* 1988, **135**:175–201.

51. Flint HJ, Tateson RW, Barthelmess IB, Porteous DJ, Donachie WD, Kacser H: **Control of the flux in the arginine pathway of Neurospora crassa. Modulations of enzyme activity and concentration**. *Biochem. J* 1981, **200**:231–246.

52. Middleton RJ, Kacser H: **Enzyme variation, metabolic flux and fitness: alcohol dehydrogenase in Drosophila melanogaster**. *Genetics* 1983, **105**:633–650.

53. Dykhuizen DE, Dean AM, Hartl DL: **Metabolic flux and fitness**. *Genetics* 1987, **115**:25–31.

54. Kruckeberg AL, Neuhaus HE, Feil R, Gottlieb LD, Stitt M: **Decreased-activity mutants of phosphoglucose isomerase in the cytosol and chloroplast of Clarkia xantiana. Impact on mass-action ratios and fluxes to sucrose and starch, and estimation of Flux Control Coefficients and Elasticity Coefficients**. *Biochem. J* 1989, **261**:457–467.

55. Ekkehard Neuhaus H, Stitt M: **Control analysis of photosynthate partitioning**. *Planta* 1990, **182**:445–454.

56. Heinisch J: **Isolation and characterization of the two structural genes coding for phosphofructokinase in yeast**. *Mol. Gen. Genet* 1986, **202**:75–82.

57. Davies SE, Brindle KM: **Effects of overexpression of phosphofructokinase on glycolysis in the yeast Saccharomyces cerevisiae**. *Biochemistry* 1992, **31**:4729–4735.

58. Ruyter GJ, Postma PW, van Dam K: **Control of glucose metabolism by enzyme IIGlc of the phosphoenolpyruvate-dependent phosphotransferase system in Escherichia coli**. *J. Bacteriol* 1991, **173**:6184–6191.

59. Quick WP, Schurr U, Scheibe R, Schulze E-D, Rodermel SR, Bogorad L, Stitt M: **Decreased ribulose-1,5-bisphosphate carboxylase-oxygenase in transgenic tobacco transformed with “antisense” rbcS**. *Planta* 1991, **183**.

60. Stitt M, Quick WP, Schurr U, Schulze E-D, Rodermel SR, Bogorad L: **Decreased ribulose-1,5-bisphosphate carboxylase-oxygenase in transgenic tobacco transformed with “antisense” rbcS**. *Planta* 1991, **183**.

61. Salter M, Knowles RG, Pogson CI: **Quantification of the importance of individual steps in the control of aromatic amino acid metabolism**. *Biochem. J* 1986, **234**:635–647.

62. Torres NV, Mateo F, Meléndez-Hevia E, Kacser H: **Kinetics of metabolic pathways. A system in vitro to study the control of flux**. *Biochem. J* 1986, **234**:169–174.

63. Rognstad R: **Rate-limiting steps in metabolic pathways**. *J. Biol. Chem* 1979, **254**:1875–1878.

64. Fell DA: **Metabolic control analysis: a survey of its theoretical and experimental development**. *Biochem. J* 1992, **286 (Pt 2)**:313–330.

65. Cornish-Bowden A: **Metabolic Control Analysis in Theory and Practice**. *Advances in Molecular and Cell Biology* 1995, **11**:21–64.

66. Atkinson DE: **Limitation of Metabolite Concentrations and the Conservation of Solvent Capacity in the Living Cell**. *Current topics in cellular regulation* 1969, **1**:29–42.

67. Fulton AB: **How crowded is the cytoplasm?** *Cell* 1982, **30**:345–347.

68. Laurent TC: **The interaction between polysaccharides and other macromolecules. 5. The solubility of proteins in the presence of dextran**. *Biochemical Journal* 1963, **89**:253.

69. Clegg JS: **Properties and metabolism of the aqueous cytoplasm and its boundaries**. *Am. J. Physiol* 1984, **246**:R133–151.

70. Kempner ES, Miller JH: **The molecular biology of Euglena gracilis. IV. Cellular stratification by centrifuging**. *Exp. Cell Res* 1968, **51**:141–149.

71. Schnell S: **Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws**. *Progress in Biophysics and Molecular Biology* 2004, **85**:235–260.

72. Laurent TC: **Enzyme Reactions in Polymer Media**. *Eur J Biochem* 1971, **21**:498–506.

73. Savageau MA: **Biochemical systems analysis. I. Some mathematical properties of the rate law for the component enzymatic reactions**. *J. Theor. Biol* 1969, **25**:365–369.

74. Kopelman R: **Fractal Reaction Kinetics**. *Science* 1988, **241**:1620–1626.

75. Vazquez A, de Menezes MA, Barabási A-L, Oltvai ZN: **Impact of Limited Solvent Capacity on Metabolic Rate, Enzyme Activities, and Metabolite Concentrations of S. cerevisiae Glycolysis**. *PLoS Comput Biol* 2008, **4**:e1000195.

76. **GNN - Quick Guide to Sequenced Genomes** [http://www.genomenewsnetwork.org/resources/sequenced_genomes/genome_guide_p1.shtml].

77. Gutenkunst RN, Waterfall JJ, Casey FP, Brown KS, Myers CR, Sethna JP: **Universally Sloppy Parameter Sensitivities in Systems Biology Models**. *PLoS Comput Biol* 2007, **3**:e189.

78. Selekov EE: **Analysis of Stoichiometric Regulation of Energy Metabolism**. In Berlin: Akademie Verlag; 1975:17–19.

79. Shapiro HM: **Input-output models of biological systems: Formulation and**

applicability. *Computers and Biomedical Research* 1969, **2**:430–445.

80. Clarke BL: **Complete set of steady states for the general stoichiometric dynamical system**. *J. Chem. Phys.* 1981, **75**:4970.

81. Palsson BO: *Systems Biology: Properties of Reconstructed Networks*. 1st edition. Cambridge University Press; 2006.

82. Hofmeyr JH: **Steady-state modelling of metabolic pathways: a guide for the prospective simulator**. *Comput. Appl. Biosci* 1986, **2**:5–11.

83. Clarke BL: **Stoichiometric network analysis**. *Cell Biochemistry and Biophysics* 1988, **12**:237–253.

84. Papoutsakis ET: **Equations and calculations for fermentations of butyric acid bacteria**. *Biotechnol. Bioeng* 1984, **26**:174–187.

85. Papoutsakis ET, Meyer CL: **Equations and calculations of product yields and preferred pathways for butanediol and mixed-acid fermentations**. *Biotechnol. Bioeng* 1985, **27**:50–66.

86. Sonnleitner B, Käppeli O: **Growth of Saccharomyces cerevisiae is controlled by its limited respiratory capacity: Formulation and verification of a hypothesis**. *Biotechnol. Bioeng.* 1986, **28**:927–937.

87. Tsai SP, Lee YH: **Application of metabolic pathway stoichiometry to statistical analysis of bioreactor measurement data**. *Biotechnol. Bioeng.* 1988, **32**:713–715.

88. Vallino JJ, Stephanopoulos G: **Metabolic flux distributions in Corynebacterium glutamicum during growth and lysine overproduction**. *Biotechnol. Bioeng.* 1993, **41**:633–646.

89. Watson MR: **Metabolic maps for the Apple II**. *Biochemical Society Transactions* 1984, **12**:1093–1094.

90. Dantzig G: *Notes on linear programming -- Part III: Computational algorithm of the revised simplex method*. Rand Corporation; 1953.

91. Schuetz R, Kuepfer L, Sauer U: **Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli**. *Mol Syst Biol* 2007, **3**:119.

92. Watson MR: **A discrete model of bacterial metabolism**. *Bioinformatics* 1986, **2**:23–27.

93. Atkinson D: **Adenine nucleotides as universal stoichiometric metabolic coupling agents**. *Advances in Enzyme Regulation* 1971, **9**:207–219.

94. Fell DA, Small JR: **Fat synthesis in adipose tissue. An examination of stoichiometric constraints.** *Biochemical Journal* 1986, **238**:781.

95. Majewski RA, Domach MM: **Simple constrained-optimization view of acetate overflow**

in E. coli. *Biotechnol. Bioeng* 1990, **35**:732–738.

96. Savinell J, Palsson B: **Network analysis of intermediary metabolism using linear optimization. I. Development of mathematical formalism**. *Journal of Theoretical Biology* 1992, **154**:421–454.

97. Savinell J, Palsson B: **Network analysis of intermediary metabolism using linear optimization. II. Interpretation of hybridoma cell metabolism**. *Journal of Theoretical Biology* 1992, **154**:455–473.

98. Varma A, Palsson BO: **Metabolic Capabilities of Escherichia coli: I. Synthesis of Biosynthetic Precursors and Cofactors**. *Journal of Theoretical Biology* 1993, **165**:477–502.

99. Varma A, Boesch BW, Palsson BO: **Biochemical Production Capabilities of Escherichia coli**. *Biotechnol. Bioeng.* 1993, **42**:59–73.

100. Varma A, Palsson BO: **Metabolic Capabilities of Escherichia coli II. Optimal Growth Patterns**. *Journal of Theoretical Biology* 1993, **165**:503–522.

101. Steels EL, Learmonth RP, Watson K: **Stress tolerance and membrane lipid unsaturation in Saccharomyces cerevisiae grown aerobically or anaerobically**. *Microbiology* 1994, **140**:569.

102. Varma A, Palsson BO: **Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110.** *Applied and Environmental Microbiology* 1994, **60**:3724.

103. Varma A, Palsson BO: **Parametric sensitivity of stoichiometric flux balance models applied to wild-type Escherichia coli metabolism**. *Biotechnol. Bioeng* 1995, **45**:69–79.

104. Varma A, Palsson BO: **Metabolic Flux Balancing: Basic Concepts, Scientific and Practical Use**. *Nat Biotechnol* 1994, **12**:994–998.

105. van Gulik WM, Heijnen JJ: **A metabolic network stoichiometry analysis of microbial growth and product formation**. *Biotechnol. Bioeng.* 1995, **48**:681–698.

106. Nissen TL, Schulze U, Nielsen J, Villadsen J: **Flux Distributions in Anaerobic, Glucose-Limited Continuous Cultures of Saccharomyces Cerevisiae**. *Microbiology* 1997, **143**:203–218.

107. Pramanik J, Keasling JD: **Stoichiometric model of Escherichia coli metabolism: Incorporation of growth-rate dependent biomass composition and mechanistic energy requirements**. *Biotechnol. Bioeng.* 1997, **56**:398–421.

108. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM: **Whole-genome random sequencing and assembly of Haemophilus influenzae Rd**. *Science* 1995, **269**:496–512.

109. Price N, Papin JA, Schilling CH, Palsson BO: **Genome-scale microbial in silico models: the constraints-based approach**. *Trends in Biotechnology* 2003, **21**:162–169.

110. Covert MW, Schilling CH, Famili I, Edwards JS, Goryanin II, Selkov E, Palsson BO: **Metabolic modeling of microbial strains in silico**. *Trends in Biochemical Sciences* 2001, **26**:179–186.

111. Edwards JS, Covert M, Palsson BO: **Metabolic modelling of microbes: the flux-balance approach**. *Environ. Microbiol* 2002, **4**:133–140.

112. Thiele I, Palsson BO: **A protocol for generating a high-quality genome-scale metabolic reconstruction**. *Nat Protoc* 2010, **5**:93–121.

113. Munro S: **Systems Biology of Thermotoga neapolitana for Biological Hydrogen Production**. 2010.

114. Edwards JS, Palsson BO: **Systems properties of the Haemophilus influenzae Rd metabolic genotype**. *J. Biol. Chem* 1999, **274**:17410–17416.

115. Schilling CH, Palsson BO: **Assessment of the metabolic capabilities of Haemophilus influenzae Rd through a genome-scale pathway analysis**. *J. Theor. Biol* 2000, **203**:249–283.

116. Edwards JS, Palsson BO: **The Escherichia coli MG1655 in silico metabolic genotype: its definition, characteristics, and capabilities**. *Proc. Natl. Acad. Sci. U.S.A* 2000, **97**:5528–5533.

117. Förster J, Famili I, Fu P, Palsson B, Nielsen J: **Genome-Scale Reconstruction of the Saccharomyces cerevisiae Metabolic Network**. *Genome Research* 2003, **13**:244–253.

118. Duarte NC, Herrgård MJ, Palsson BO: **Reconstruction and validation of Saccharomyces cerevisiae iND750, a fully compartmentalized genome-scale metabolic model**. *Genome Research* 2004, **14**:1298.

119. Famili I, Förster J, Nielsen J, Palsson BO: **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* 2003, **100**:13134–13139.

120. Förster J, Famili I, Palsson BO, Nielsen J: **Large-scale evaluation of in silico gene deletions in Saccharomyces cerevisiae**. *Omics A Journal of Integrative Biology* 2003, **7**:193–202.

121. Chechik G, Oh E, Rando O, Weissman J, Regev A, Koller D: **Activity motifs reveal principles of timing in transcriptional control of the yeast metabolic network**. *Nat Biotechnol* 2008, **26**:1251–1259.

122. Deutscher D, Meilijson I, Schuster S, Ruppin E: **Can single knockouts accurately single out gene functions?** *BMC Syst Biol* 2008, **2**:50.

123. Cimini D, Patil KR, Schiraldi C, Nielsen J: **Global transcriptional response of Saccharomyces cerevisiae to the deletion of SDH3**. *BMC Syst Biol* 2009, **3**:17.

124. Oberhardt MA, Palsson BØ, Papin JA: **Applications of genome-scale metabolic reconstructions**. *Mol Syst Biol* 2009, **5**.

125. Poolman MG, Miguet L, Sweetlove LJ, Fell DA: **A Genome-Scale Metabolic Model of Arabidopsis and Some of Its Properties**. *Plant Physiology* 2009, **151**:1570–1581.

126. Grafahrend-Belau E, Schreiber F, Koschutzki D, Junker BH: **Flux Balance Analysis of Barley Seeds: A Computational Approach to Study Systemic Properties of Central Metabolism**. *Plant Physiology* 2008, **149**:585–598.

127. Oura E: **The effect of aeration on the growth energetics and biochemical composition of baker's yeast, with an appendix: Reactions leading to the formation of yeast cell material from glucose and ethanol.** 1972.

128. Vaughan-Martini A, Martini A: **A taxonomic key for the genus Saccharomyces**. *Systematic and Applied Microbiology* 1993, **16**:113–119.

129. Nurminen T, Konttinen K, Suomalainen H: **Neutral lipids in the cells and cell envelope fractions of aerobic baker's yeast and anaerobic brewer's yeast**. *Chemistry and Physics of Lipids* 1975, **14**:15–32.

130. Hunter K, Rose AH: **Lipid composition of Saccharomyces cerevisiae as influenced by growth temperature**. *Biochim. Biophys. Acta* 1972, **260**:639–653.

131. Kaneko H, Hosohara M, Tanaka M, Itoh T: **Lipid composition of 30 species of yeast**. *Lipids* 1976, **11**:837–844.

132. Schulze U: **Anaerobic physiology of Saccharomyces cerevisiae.** 1995.

133. Kuepfer L, Sauer U, Blank LM: **Metabolic functions of duplicate genes in Saccharomyces cerevisiae**. *Genome Res* 2005, **15**:1421–1430.

134. Mo ML, Palsson BØ, Herrgård MJ: **Connecting extracellular metabolomic measurements to intracellular flux states in yeast**. *BMC Syst Biol* 2009, **3**:37.

135. Herrgard MJ, Swainston N, Dobson P, Dunn WB, Arga KY, Arvas M, Buthgen N, Borger S, Costenoble R, Heinemann M, Hucka M, Le Novere N, Li P, Liebermeister W, Mo ML, Oliveira AP, Petranovic D, Pettifer S, Simeonidis E, Smallbone K, Spasie I, Weichart D, Brent R, Broomhead DS, Westerhoff HV, Kurdar B, Penttila M, Klipp E, Palsson BO, Sauer U, Oliver SG, Mendes P, Nielsen J, Kell DB: **A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology**. *Nat Biotech* 2008, **26**:1155–1160.

136. Nookaew I, Jewett MC, Meechai A, Thammarongtham C, Laoteng K, Cheevadhanarak S, Nielsen J, Bhumiratana S: **The genome-scale metabolic model iIN 800 of Saccharomyces cerevisiae and its validation: a scaffold to query lipid metabolism**. *BMC Syst Biol* 2008, **2**:71.

137. **Yeast consensus metabolic network v2** [http://www.comp-sys-bio.org/yeastnet/v2/index.html].

138. **Yeast consensus metabolic network v4** [http://www.comp-sys-bio.org/yeastnet/].

139. Xia J-M, Yuan Y-J: **Comparative Lipidomics of Four Strains of Saccharomyces cerevisiae Reveals Different Responses to Furfural, Phenol, and Acetic Acid**. *J. Agric. Food Chem.* 2009, **57**:99–108.

140. Vanrolleghem PA, de Jong-Gubbels P, van Gulik WM, Pronk JT, van Dijken JP, Heijnen S: **Validation of a Metabolic Network for Saccharomyces cerevisiae Using Mixed Substrate Studies**. *Biotechnology Progress* 1996, **12**:434–448.

141. Gombert AK, Moreira dos Santos M, Christensen B, Nielsen J: **Network Identification and Flux Quantification in the Central Metabolism of Saccharomyces cerevisiae under Different Conditions of Glucose Repression**. *Journal of Bacteriology* 2001, **183**:1441–1451.

142. Carlson R, Fell D, Srienc F: **Metabolic pathway analysis of a recombinant yeast for rational strain development**. *Biotechnol. Bioeng.* 2002, **79**:121–134.

143. Herrgård MJ: **Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in Saccharomyces cerevisiae**. *Genome Research* 2006, **16**:627–635.

144. Hjersted JL, Henson MA, Mahadevan R: **Genome-scale analysis of Saccharomyces cerevisiae metabolism and ethanol production in fed-batch culture**. *Biotechnol. Bioeng.* 2007, **97**:1190–1204.

145. Zomorrodi AR, Maranas CD: **Improving the iMM 904 S. cerevisiae metabolic model using essentiality and synthetic lethality data**. *BMC Syst Biol* 2010, **4**:178.

146. Heavner BD, Smallbone K, Barker B, Mendes P, Walker LP: **Yeast 5 - an expanded reconstruction of the Saccharomyces Cerevisiae metabolic network**. *BMC Systems Biology* 2012, **6**:55.

147. Bro C, Regenberg B, Förster J, Nielsen J: **In silico aided metabolic engineering of Saccharomyces cerevisiae for improved bioethanol production**. *Metabolic Engineering* 2006, **8**:102–111.

148. Kennedy CJ, Boyle PM, Waks Z, Silver PA: **Systems-Level Engineering of Nonfermentative Metabolism in Yeast**. *Genetics* 2009, **183**:385–397.

149. Asadollahi MA, Maury J, Patil KR, Schalk M, Clark A, Nielsen J: **Enhancing sesquiterpene production in Saccharomyces cerevisiae through in silico driven metabolic engineering**. *Metabolic Engineering* 2009, **11**:328–334.

150. Soh KC, Miskovic L, Hatzimanikatis V: **From network models to network responses: integration of thermodynamic and kinetic properties of yeast genome-scale metabolic networks**. *FEMS Yeast Research* 2012, **12**:129–143.

151. Price ND, Reed JL, Palsson BO: **Genome-scale models of microbial cells: evaluating the consequences of constraints**. *Nature Reviews Microbiology* 2004, **2**:886–897.

152. Kumar VS, Maranas CD: **GrowMatch: an automated method for reconciling in**

silico/in vivo growth predictions. *PLoS Comput Biol* 2009, **5**.

153. Kumar S, Dasika MS, Maranas CD: **Optimization based automated curation of metabolic reconstructions**. *BMC Bioinformatics* 2007, **8**:212.

154. Szappanos B, Kovács K, Szamecz B, Honti F, Costanzo M, Baryshnikova A, Gelius-Dietrich G, Lercher MJ, Jelasity M, Myers CL, Andrews BJ, Boone C, Oliver SG, Pál C, Papp B: **An integrated approach to characterize genetic interaction networks in yeast metabolism**. *Nature Genetics* 2011, **43**:656–662.

155. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A, Cuellar AA, Dronov S, Gilles ED, Ginkel M, Gor V, Goryanin II, Hedley WJ, Hodgman TC, Hofmeyr JH, Hunter PJ, Juty NS, Kasberger JL, Kremling A, Kummer U, Le Novère N, Loew LM, Lucio D, Mendes P, Minch E, Mjolsness ED, Nakayama Y, Nelson MR, Nielsen PF, Sakurada T, Schaff JC, Shapiro BE, Shimizu, TS, Spence HD, Stelling J, Takahashi K, Tomita M, Wagner J, Wang J: **The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models**. *Bioinformatics* 2003, **19**:524–531.

156. **The 1st Workshop on Software Platforms for Systems Biology** [http://sbml.org/Events/Workshops/The_1st_Workshop_on_Software_Platforms_for_Systems_Biology].

157. **Extensible Markup Language (XML) 1.0** [http://www.w3.org/TR/1998/REC-xml-19980210].

158. Hucka M, Finney A, Bornstein BJ, Keating SM, Shapiro BE, Matthews J, Kovitz BL, Schilstra MJ, Funahashi A, Doyle JC, Kitano H: **Evolving a lingua franca and associated software infrastructure for computational systems biology: the Systems Biology Markup Language (SBML) project**. *Syst. Biol.* 2004, **1**:41.

159. Bornstein BJ, Keating SM, Jouraku A, Hucka M: **LibSBML: an API Library for SBML**. *Bioinformatics* 2008, **24**:880–881.

160. Keating SM, Bornstein BJ, Finney A, Hucka M: **SBMLToolbox: an SBML toolbox for MATLAB users**. *Bioinformatics* 2006, **22**:1275.

161. Becker SA, Feist AM, Mo ML, Hannum G, Palsson BØ, Herrgard MJ: **Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox**. *Nat Protoc* 2007, **2**:727–738.

162. Makhorin A: *GLPK (GNU Linear Programming Kit)*. Free Software Foundation; 2000.

CHAPTER 3 – EVALUATING SPHINGOLIPID BIOCHEMISTRY IN THE CONSENSUS RECONSTRUCTION OF YEAST METABOLISM[1]

Abstract:

Reconstructed metabolic networks are the basis for genome scale models of cellular metabolism. Such models have demonstrated good predictive accuracy when analyzed with the optimization-based approach of flux balance analysis, so are of particular interest for informing the rational design of metabolic engineering strategies for industrial biotechnology applications.

As a preliminary step towards applying a metabolic model of the industrially important yeast *Saccharomyces cerevisiae*, we validated the underlying reconstruction of yeast metabolism in the Yeast Consensus Reconstruction by examining the model's representation of yeast sphingolipid metabolism. Sphingolipids, which comprise 30% of the plasma membrane lipids in yeast, incorporate very long chain fatty acids of lengths up to C-26. Such long-chain hydrocarbons are of particular interest for the production of liquid transportation fuels.

We began with literature review and consultation with experts in yeast sphingolipid metabolism to compile a list a list of 243 reactions associated with yeast sphingolipid metabolism. We found that the Yeast Reconstruction contains only 41 reactions in this pathway, and that the reconstruction includes inaccuracies.

This documentation of the incomplete reconstruction of sphingolipid metabolism in the Yeast Consensus Metabolic Reconstruction demonstrates opportunity for improving computational reconstruction of established biochemical knowledge. The extent of information

[1] This chapter has previously been published as: Heavner BD, Henry SA, Walker LP: **Evaluating Sphingolipid Biochemistry in the Consensus Reconstruction of Yeast Metabolism**. *Industrial Biotechnology* 2012, **8**: 72-78. doi:10.1089/ind.2012.0002.

missing from the Yeast Reconstruction suggests that incorporating knowledge from the primary literature is an underemphasized tool for refining genome scale metabolic models. Our expert-knowledge driven and pathway-centric approach to reconstruction curation provides a rational approach for community participation in the ongoing effort to reconstruct the yeast biochemical reaction network.

Introduction:

Application of constraint-based analysis and simulation methods to computational models derived from reconstructed metabolic networks[1] provides a theoretical framework that has shown great potential for facilitating rational metabolic engineering.[2, 3] Successful application of such methods for strain engineering requires both that the metabolic network model is capable of accurate predictions of phenotypes resulting from genetic modifications and that the biochemical network reconstruction, the structured knowledge base from which a metabolic model is derived,[4] contains detail of sufficient scope and accuracy to describe the metabolic process of interest. Although algorithms have been developed for adding reactions to metabolic models in order to improve their ability to predict gene essentiality[5–7], evaluation of the details, scope, and biochemical accuracy of a metabolic reconstruction requires careful manual curation of the reconstructed metabolic network relying upon relevant literature.

Motivated by increasing interest in the microbial production of long-chain aliphatic hydrocarbons for advanced cellulosic biofuels, we undertook an evaluation of the recent update to the consensus metabolic reconstruction of the yeast *Saccharomyces cerevisiae*[8] ("Yeast v4.05"), with particular focus on the biosynthesis of complex sphingolipids. Sphingolipids comprise about 7-8% of the plasma membrane mass, and 30% of the plasma membrane lipids in *Saccharomyces cerevisiae*, though the exact lipid composition is dependent upon environmental

conditions.[9] Accurate and complete computational reconstruction of sphingolipid metabolism is a worthwhile goal because sphingolipids incorporate industrially-useful very long chain fatty acids (VLCFAs), [10] and also because the fluxes through reactions of the sphingolipid biosynthetic pathway are a focus of active research. Yeast mutants have been observed which appear to prioritize inositol flux to the biosynthesis of sphingolipids at the expense of phosphatidylinositol (PI), an observation which links sphingolipid biosynthesis with the metabolism of storage lipids, phospholipids, and sterols.[11] Sphingolipids play important roles in signalling and responding to environmental stress (including heat, osmotic, and low-pH stresses),[9, 12, 13] and thus are key players in the dynamic process by which yeast cells restructure their plasma membrane in response to effects such as ethanol stress.[14] Metabolic precursors of complex sphingolipids have been implicated in both cell growth arrest, and cellular proliferation and survival, giving rise to the concept of a "sphingolipid rheostat" which plays an important role in determining whether cells survive or die.[15] Compounds involved in sphingolipid metabolism have been found to affect yeast metabolism through signal cascades resulting in transcriptional regulatory events,[16] but the mechanisms influencing flux through sphingolipid pathways are not well understood. Significantly, recent research provides evidence that sphingolipid synthesis is at least partially regulated through substrate availability due to changes in PI turnover, rather than solely through transcriptional regulation.[17]

Computational models are an important tool in efforts to understand and manipulate sphingolipid metabolism. Dynamic models have been successfully applied to investigate sphingolipid regulation,[18] and a small-scale stoichiometrically constrained model has been developed to suggest possible targets for antifungal drugs.[19] In this work, we set out to examine how well sphingolipid metabolism has been incorporated in a previously published

genome-scale model of yeast metabolism.

Though there are "a large array of theoretical lipid species" in *S. cerevisiae*,[20] we chose to limit our evaluation to the most abundant sphingolipids produced by wild-type *S. cerevisiae* in rich media under normal laboratory conditions - those consisting of a C-18 sphingoid base amide-linked to C-24 or C-26 VLCFAs.[21] Since the sphingoid base may be hydroxylated to form phytosphingosine and the VLCFA may include 0, 1, or 2 hydroxyl groups,[22] our list of reactions included the hydroxylation reactions and substrates required to form the various hydroxylation permutations, called ceramide -1, -2', -2, -3, and -4, with the corresponding downstream complex sphingolipids of the –A, -B', -B, -C, and –D series (Figure 3.1).

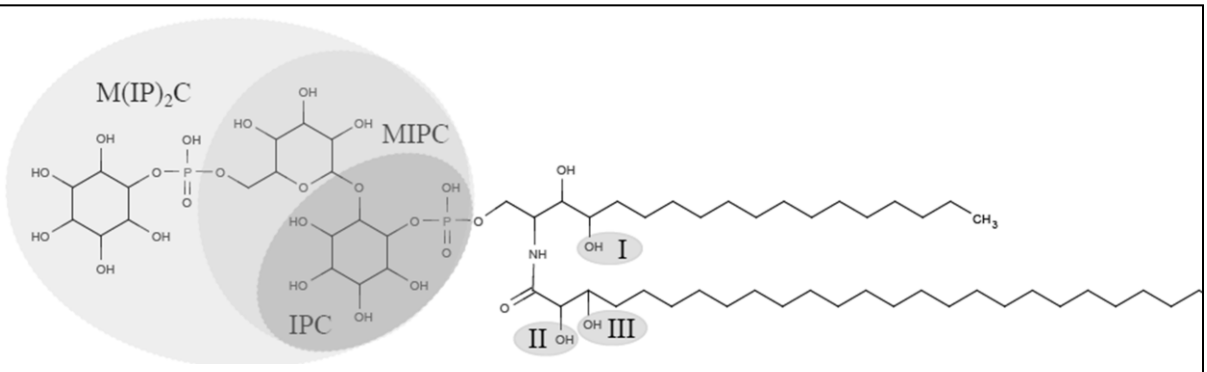


Figure 3.1: Yeast complex sphingolipids

Yeast complex sphingolipids of the –A, -B, -B', -C, or –D families are derivatives of ceramides -1, -2, -2', -3, and -4, which are distinguished by hydroxylation at the sites labeled I, II, and III [22]. Yeast complex sphingolipids include inositol phosphorylceramides (IPC), mannose inositol phosphoceramides (MIPC), and inositol phosphomannosylinositol phosphoceramides ($M(IP)_2C$). Though the specific location of site III is not settled in the literature, we chose to use that indicated in [10].

The biosynthetic pathway for complex sphingolipids has been the subject of recent

review articles.[10, 23] Following the example of these articles, the scope of our validation effort begins with the condensation of palmitoyl-CoA with serine, and ends with the breakdown of complex sphingolipids by Isc1p. Since they are shared with other portions of lipid metabolism, the reactions of the fatty acid elongation cycle are outside the scope of this evaluation.

Following a review of the literature and online resources, we generated a list of 243 reactions and 23 genes involved in the biosynthesis and catabolism of complex sphingolipids in yeast. This list of reactions, with references to supporting literature, is available in Appendix 1 as Supplementary Table 3.1. We found that Yeast v4.05 includes only 41 reactions involved in sphingolipid metabolism (Appendix 1, Supplementary Table 3.2), a reconstruction that differs from the network documented in the literature. This observation supports the conclusion that even "well curated metabolic reconstructions" [24] benefit from manual review, particularly in important metabolic pathways falling outside central carbon metabolism. Continued curation, validation, and refinement of reconstructions remains necessary to yield models that are better for informing metabolic engineering strategies. From our detailed notes of the differences between the biochemical literature and the Yeast consensus reconstruction, we have generated suggested changes to the Yeast Reconstruction, which have been submitted to the model's curators to improve this community resource.

Materials and Methods:

We examined the reconstruction of sphingolipid metabolism in Yeast v4.05[8] by comparing reactions and gene annotation included in the model with information in relevant biochemical literature. The list of 243 reactions generated from our literature review, with references to supporting literature, is available in Appendix 1 as Supplementary Table 3.1. We began by listing 23 genes associated with sphingolipid metabolism (excluding the fatty acid

elongation cycle) in a recent review article[10] (Appendix 1 Supplementary Table 3.2), along with the loci of these genes, as listed in the Saccharomyces Genome Database (SGD)[25]. After reviewing the description of these gene functions in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database,[26] SGD, the Universal Protein Resource Knowledge Base (UniProtKB),[27] and the MetaCyc and BioCyc databases,[28] we conducted a more detailed review of the primary literature to generate a list of biochemical reactions that we could compare to the reactions included in Yeast v4.05. Since there is wide variation in the nomenclature used to describe chemical intermediates and products in sphingolipid metabolism, we relied upon the Nature Lipidomics Gateway (LIPID MAPS)[29] and the Chemical Entities of Biological Interest (ChEBI)[30] databases to resolve different compound names. In cases where nomenclature varied between primary sources but the compounds of interest were not included in LIPID MAPS or ChEBI, we chose the nomenclature that was most frequently used in the literature we reviewed.

We downloaded the SBML[31] distribution of Yeast v4.05[8] from the sourceforge code repository[32] and imported the model to MATLAB® version 2010a (MathWorks, Natick, MA), on a desktop PC running Windows XP (Microsoft, Redmond WA) using the libSBML[33] interface of the "readcbmodel" function of version 2.0 of the COBRA Toolbox.[34]

Genes are associated with reactions in Yeast v4.05 through reaction annotations indicating the open reading frame locus, which corresponds to the gene which codes for the protein catalyzing the reaction. We confirmed that none of the genes we selected from the literature were included in the list of "poorly characterized genes" distributed as supplemental data to the paper describing the Yeast metabolic model.[8] Following confirmation, COBRA toolbox functions and our own MATLAB® scripts were used to retrieve information for each

reaction in the Yeast model which is annotated with a gene that was identified as relating to sphingolipid metabolism. Model information – including reaction substrates, products, directionality, compartmentalization, and gene association – was compared to the reaction list that had been generated from the literature review.

Results:

We found that Yeast v4.05 contains 17% of the biochemical and transport reactions relating to sphingolipid metabolism and incorporates 87% of the genes identified from our review of relevant biochemical literature (Figure 3.2).

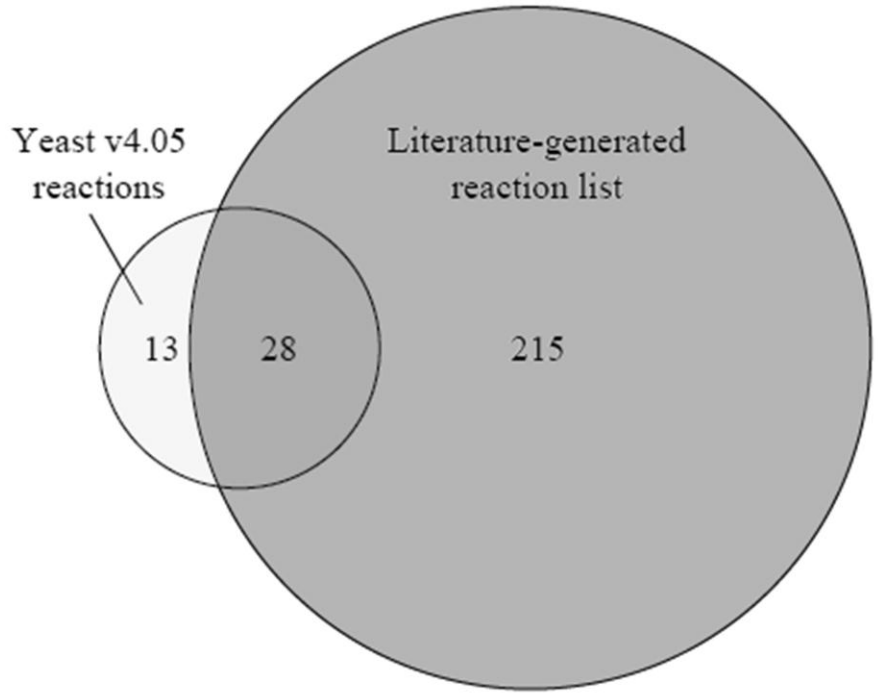


Figure 3.2: Sphingolipid reaction coverage of Yeast v4.05

Literature review generated a list of 243 reactions involved in yeast sphingolipid metabolism, including hypothetical intercompartmental transport reactions necessary for modeling. Yeast v4.05 contains 41 reactions in the sphingolipid pathway. Of these 41 reactions, 28 were present on our list (including six reactions that were lumped reactions in Yeast v4.05).

The Yeast v4.05 reconstruction of sphingolipid metabolism differs from reactions described in the literature (Appendix 1 Supplementary Table 3.1), primarily in differentiation between ceramide and sphingolipid species. The differences between our literature-based reconstruction and the Yeast v4.05 reconstruction are summarized in Table 3.1.

	Literature-based reconstruction	Yeast v4.05
Reactions	243	41
Unique Chemical Species	747	138
Complex Sphingolipids	30	4
Genes	23	20
Orphan subnetworks	0	1

Table 3.1: Literature-based reconstruction of sphingolipid metabolism differs from the Yeast v4.05 reconstruction

Our literature-based reconstruction of sphingolipid metabolism includes more reactions, chemical species, complex sphingolipids, and genes than the Yeast reconstruction. Additionally, it does not include an orphan subnetwork involving the generic species "inositol phosphosphingolipid" which is included in the Yeast reconstruction.

Our review of literature found that the products of 23 genes involved in the yeast sphingolipid metabolic pathway (Appendix 1 Supplementary Table 3.3) catalyze 118 reactions. This list of reactions was supplemented with an additional 125 hypothetical intercompartmental transport reactions which were necessary for modeling purposes. These hypothetical reactions enable flux balance analysis of compartmentalized metabolic network models, even though the mechanism of such intercompartmental transport may be unknown. The list of reactions is included with references in Appendix 1 as Supplementary Table 3.1.

We found that Yeast v4.05 includes 41 reactions either involved in sphingolipid metabolism or annotated as being catalyzed by gene products that are associated with

sphingolipid metabolism (Appendix 1 Supplementary Table 3.2). This sum includes 5 hypothetical intercompartmental transport reactions needed for modeling. Yeast v4.05 does not contain reactions annotated with three of the 23 genes associated with sphingolipid metabolism: *TSC3*, *LIP1*, and *CSH1* (open reading frames *YBR058C-A*, *YMR298W*, and *YBR161W*, respectively) (Appendix 1 Supplementary Table 3.3).

The Yeast v4.05 network (a network diagram is available in Appendix 1 as Supplemental Figure 3.1) includes incomplete and inconsistent differentiation between ceramide and sphingolipid species which vary by hydroxylation[22] or length of the VLCFA moiety.[20] Though Yeast v4.05 partially differentiates ceramides by hydroxylation through inclusion of ceramides-1, -2, -2', and -3, and includes ceramides with both C-24 and C-26 VLCFAs, it has less complete differentiation of the inositol phosphoceramides (IPC) (Yeast v4.05 only contains IPC-B, and IPC-C), and does not differentiate between differing VLCFAs for any of the complex sphingolipids. Further, Yeast v4.05 does not differentiate among the mannose inositol phosphoceramides (MIPC) or the inositol phosphomannosylinositol phosphoceramides ($M(IP)_2C$). Ceramide-4 and the derivative complex sphingolipids of the –D family are missing from this model, and the model incorrectly includes reactions to produce IPC-B from ceramide-2' instead of ceramide-2.

If the hydroxylation and VLCFA length of ceramide and ceramide species was intentionally simplified in Yeast v4.05 for modeling purposes, it was done so inconsistently. However, the increased model complexity that would be added due to consistent differentiation of ceramide and sphingolipid species is justified because the differentiation has been observed *in vivo* following disruption of genes which are included in Yeast v4.05, such as SCS7[20] and CSG1/2.[35] Varying hydroxylation has been shown to affect membrane properties[36], and

metabolic enzymes such as CSG1 have been shown to have different affinities for sphingolipids of different families[35].

Since Yeast v4.05 contains more ceramides than sphingolipids, the reactions responsible for the synthesis of complex sphingolipids are lumped together. Though there are two ceramide-2' species which differ by VLCFAs, Yeast 4.05 includes reactions (reaction indexes r_0621 or r_0622) to convert them both to a single IPC species, incorrectly called IPC-B (metabolite index s_0828) instead of IPC-B'. Yeast v4.05 does not include a reaction to produce IPC-B from either ceramide-2 species, but does include reactions to produce a single IPC-C species from either of the two ceramide-3 species (reaction indexes r_0623 and r_0624). The model does not include IPC-A, -B', or D. The model includes two reactions to produce a single MIPC species (metabolite index s_1013) from either IPC-B or IPC-C (reaction indexes r_0723 and r_0724). The single $M(IP)_2C$ species (metabolite index s_0825) in Yeast v4.05 is produced from this generic MIPC. Through this inconsistent lumping of the later stages of the biosynthetic pathway, biologically meaningful distinctions are lost, and inaccuracies are introduced to Yeast v4.05 (such as the production of IPC-B from ceramide-2').

Although Yeast v4.05 includes complex sphingolipid degradation reactions correctly annotated with *ISC1* [21], these reactions (reaction indexes r_0625 and r_0626) do not use any of the Yeast v4.05 IPC, MIPC, or $M(IP)_2C$ species as a substrate. Instead, they use an additional generic species, "inositol phosphosphingolipid", which exists in the mitochondrion and endoplasmic reticulum compartments of Yeast v4.05 (metabolite ids s_0827 and s_0826, respectively). Reactions r_0625 and r_0626 are the only reactions the "inositol phosphosphingolipid" metabolite participates in, and it is otherwise unconnected to the reconstructed network, forming an orphan subnetwork. Thus, Yeast v4.05 does not include any

effective pathway for the degradation of complex sphingolipids.

The complex sphingolipid synthetic reactions in Yeast v4.05's "Golgi apparatus" compartment are incorrectly modeled as reversible reactions. Thus, a practical consequence of the use of general species for IPC-B, IPC-C, MIPC, and $M(IP)_2C$ species is that Yeast v4.05 includes an interconversion pathway for ceramides or sphingolipids with differing VLCFA lengths via the reactions catalyzed by Aur1p, the Csg2p /Sur1p complex, and Ipt1p. The biochemical details of such interconversion in yeast are poorly understood at this time,[37] but there is little evidence suggesting that these enzymes catalyze the reverse reactions included in Yeast v4.05 *in vivo.*

A second consequence of the incomplete differentiation between ceramide and sphingolipid species in Yeast v4.05 is that the constraint based optimization problem solved in flux balance analysis does not yield a solution for maximizing biomass production under anaerobic conditions. Specifically, the model incorrectly includes only two paths to the required biomass component $M(IP)_2C$, and both require oxygen as cofactors. One path in the Yeast v4.05 model proceeds via the production of ceramide-2' by hydroxylation of ceramide-1 (reaction indexes r_0287 or r_0289), and the other path proceeds via the production of ceramide-3 by hydroxylation of ceramide-2 (reaction indexes r_0294 or r_0295). One way to circumvent this problem with Yeast v4.05 is to remove complex sphingolipids from the biomass definition. The model's incorrect oxygen requirement for complex sphingolipid biosynthesis can also be corrected by expanding the biochemistry reflected in the yeast metabolic network reconstruction to include the reactions to synthesize the unhydroxylated complex sphingolipids derived from ceramide-1. However, neither approach is sufficient to enable flux balance analysis prediction of anaerobic growth with the model due to errors in the reconstruction which lay outside the scope

of this curation effort.

To resolve differences between the sphingolipid pathway described in yeast biochemical literature and Yeast v4.05, we have generated suggested changes to the Yeast Reconstruction (a network diagram including our suggested changes is available in Appendix 1 as Supplemental Figure 3.2), and have submitted a list of proposed changes to the maintainers of the consensus reconstruction for inclusion in future versions of the Yeast model. The proposed changes include: 221 new reactions; 70 new chemical species; annotation changes for five reactions; the addition of three new genes; the modification of 19 existing reactions to improve cofactor or stoichiometric errors; 95 new intercompartmental transport reactions; and 62 new "IS A" reactions to create the generic species that are included in the biomass definition used for flux balance analysis of the Yeast model.

Additionally, we have submitted corrections to other pathway or biochemical databases where we have noted omissions or errors. Thus, 63 new compounds were submitted to ChEBI; the UniProtKB/Swiss-Prot description of *SCS7* function was updated; the Nature LipidMaps entry for ceramide-3 was corrected; and corrections were submitted to KEGG to assign yeast genes to reactions that were previously annotated only with Enzyme Commission numbers.

Discussion:

Through collaboration with subject area experts and reviewing biochemical literature, we have found that the Yeast Reconstruction does not accurately describe yeast sphingolipid metabolism. We have compiled a list of suggested changes to the reconstruction, which will enable the construction of genome scale metabolic models which better reconstruct this important portion of yeast lipid metabolism.

Incorporation of the suggested changes will result in a yeast metabolic reconstruction that is more complete because of broader inclusion of yeast metabolic genes, more consistent in

differentiating between sphingolipid and ceramide species, and more biochemically accurate in its description of the biosynthesis and degradation of complex sphingolipids. Further, this curation effort will improve the predictive accuracy of the Consensus Model: expanding the number of complex sphingolipids from 4 to 30 enables more accurate biomass demand functions, and careful examination of the three additional genes to be added to the sphingolipid pathway refines the list of essential genes us as the basis for phenotype prediction metrics: TSC3 is an essential gene only at elevated temperatures,[38] deletion of CSH1 has been found to increase fitness in minimal media,[39] and LIP1 deletion has been found to be cause reduced growth rate[40] (though it was scored as essential in multiple large-scale knockout surveys[41, 42] which have been used as a basis for evaluating yeast metabolic models[43]).

Though the process of curating the reconstruction will continue, these improvements to sphingolipid metabolism are an important step towards improving the computational reconstruction of yeast lipid metabolism, a portion of metabolism which has proved challenging to previous reconstructions.[43, 44, 8] Improving the reconstruction of sphingolipid metabolism is a prerequisite to applying computer models derived from the reconstruction to research questions such as the interplay between fluxes of PI and complex sphingolipids, the consequences of genetic engineering on lipid metabolism, and mechanisms of stress response. As stoichiometrically constrained models expand to address signalling and regulatory networks,[45] accurate reconstruction of sphingolipid metabolism will be necessary to model the signalling role of pathway intermediates.

Our suggested refinements to the consensus reconstruction, though limited in scope to sphingolipid metabolism, demonstrate that examining known biochemistry, whether through manual literature review or collaboration with experts in yeast metabolism, remains a valuable

technique for improving the completeness and accuracy of the computational representation of the yeast molecular interaction network as the modeling community works to develop a "highly curated organism-specific knowledge base."[46]

The benefits of improving the reconstruction extend beyond efforts to apply yeast models for engineering applications. Recently developed computational approaches which enable high-throughput generation of draft metabolic models depend upon a database of well-curated reactions for filling gaps in draft reconstructions.[47] The documented limitations of current databases,[48] as well as errors encountered in the present study, demonstrate that the databases currently used for gap filling may not be sufficient for complete and accurate high-throughput model building from draft metabolic reconstructions.

The expansion of the reconstruction of yeast sphingolipid metabolism which would result if the suggested changes are incorporated in the Yeast Reconstruction creates new modeling opportunities, but also raises new challenges to existing model evaluation approaches, such as flux balance analysis. On one hand, researchers can begin to develop and test more detailed hypotheses about complex sphingolipid metabolism in yeast – research that has begun with smaller scope reconstructions[19] but has not been expanded to the genome scale. On the other hand, the biochemical constraints which determine factors such as the relative abundance of chemical species synthesized through parallel metabolic pathways (such as IPC-B and IPC-C) are currently unknown. Without imposing constraints (either on specific reactions or in defining the objective function), flux balance analysis using a biomass objective function would not be expected to correctly predict the mixture of complex ceramides that has been observed in wild-type yeast of the S288C background.[20, 49] Expanding the biochemistry included in the Yeast Reconstruction would facilitate development of such modeling efforts, which is not currently

possible because the reconstruction does not differentiate among the great diversity of complex sphingolipids.[50]

It has been suggested that “a draft model reconstruction followed by even a detailed manual curation may not be a sufficient strategy to bring a eukaryotic genome-scale model to the same quality level as a microbial one.”[51] However, the current work demonstrates that even after 8 years of effort by a variety of independent groups, the process of manual curation of the yeast biochemical network reconstruction is unfinished, and so a definitive answer to that question is not at hand. There remains a large and growing corpus of biochemical knowledge which should be added to the Yeast Reconstruction. Consequently, iterative manual curation remains an essential and unfinished community project for the computational reconstruction of the yeast biochemical reaction network that is necessary for bottom-up mechanistic models of yeast metabolism.

REFERENCES

1. Feist AM, Herrgaard MJ, Thiele I, Reed JL, Palsson BO: **Reconstruction of biochemical networks in microorganisms**. *Nat Rev Microbiol* 2008, **7**:129–143.

2. Milne CB, Kim P-J, Eddy JA, Price ND: **Accomplishments in genome-scale in silico modeling for industrial and medical biotechnology**. *Biotechnol J* 2009, **4**:1653–1670.

3. Oberhardt MA, Chavali AK, Papin JA: **Flux balance analysis: interrogating genome-scale metabolic networks**. *Methods in Molecular Biology, Systems Biology* 2009, **500**:61–80.

4. Thiele I, Palsson BO: **A protocol for generating a high-quality genome-scale metabolic reconstruction**. *Nat Protoc* 2010, **5**:93–121.

5. Burgard AP, Pharkya P, Maranas CD: **Optknock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization**. *Biotechnol Bioeng* 2003, **84**:647–657.

6. Kumar S, Dasika MS, Maranas CD: **Optimization based automated curation of metabolic reconstructions**. *BMC Bioinformatics* 2007, **8**:212.

7. Kumar VS, Maranas CD: **GrowMatch: an automated method for reconciling in silico/in vivo growth predictions**. *PLoS Comput Biol* 2009, **5**.

8. Dobson PD, Jameson D, Simeonidis E, Lanthaler K, Pir P, Lu C, Swainston N, Dunn WB, Fisher P, Hull D, Brown M, Oshota O, Stanford NJ, Kell DB, King RD, Oliver SG, Stevens RD, Mendes P: **Further developments towards a genome-scale metabolic model of yeast**. *BMC Syst Biol* 2010, **4**:145.

9. Patton JL, Lester RL: **The phosphoinositol sphingolipids of Saccharomyces cerevisiae are highly localized in the plasma membrane**. *J Bacteriol* 1991, **173**:3101–3108.

10. Dickson RC: **Thematic Review Series: Sphingolipids. New insights into sphingolipid metabolism and function in budding yeast**. *J Lipid Res* 2008, **49**:909–921.

11. Gaspar ML, Hofbauer HF, Kohlwein SD, Henry SA: **Coordination of Storage Lipid Synthesis and Membrane Biogenesis: Evidence for Cross-Talk Between Triacylglycerol Metabolism and Physphatidylinositol Synthesis**. *J Biol Chem* 2011, **286**:1696–1708.

12. Patton JL, Srinivasan B, Dickson RC, Lester RL: **Phenotypes of sphingolipid-dependent strains of Saccharomyces cerevisiae**. *J Bacteriol* 1992, **174**:7180–7184.

13. Dickson RC, Nagiec EE, Skrzypek M, Tillman P, Wells GB, Lester RL: **Sphingolipids are potential heat stress signals in Saccharomyces**. *J Biol Chem* 1997, **272**:30196–30200.

14. Alexandre H, Rousseaux I, Charpentier C: **Ethanol adaptation mechanisms in Saccharomyces cerevisiae**. *Biotechnol Appl Biochem* 1994, **20 (Pt 2)**:173–183.

15. Spiegel S: **Sphingosine 1-Phosphate, a Key Cell Signaling Molecule**. *J Biol Chem* 2002,

277:25851–25854.

16. Jesch SA, Gaspar ML, Stefan CJ, Aregullin MA, Henry SA: **Interruption of Inositol Sphingolipid Synthesis Triggers Stt4p-dependent Protein Kinase C Signaling**. *J Biol Chem* 2010, **285**:41947–41960.

17. Brice SE, Alford CW, Cowart LA: **Modulation of Sphingolipid Metabolism by the Phosphatidylinositol-4-phosphate Phosphatase Sac1p through Regulation of Phosphatidylinositol in Saccharomyces cerevisiae**. *J Biol Chem* 2009, **284**:7588–7596.

18. Alvarez-Vasquez F: **Integration of kinetic information on yeast sphingolipid metabolism in dynamical pathway models**. *J Theor Biol* 2004, **226**:265–291.

19. Kavun Ozbayraktar FB, Ulgen KO: **Stoichiometric network reconstruction and analysis of yeast sphingolipid metabolism incorporating different states of hydroxylation**. *BioSystems* 2011, **104**:63–75.

20. Guan XL, Wenk MR: **Mass spectrometry-based profiling of phospholipids and sphingolipids in extracts fromSaccharomyces cerevisiae**. *Yeast* 2006, **23**:465–477.

21. Kitagaki H, Cowart L, Matmati N, Vaenadeavalos S, Novgorodov S, Zeidan Y, Bielawski J, Obeid L, Hannun Y: **Isc1 regulates sphingolipid metabolism in yeast mitochondria**. *Biochim Biophys Acta Biomembranes* 2007, **1768**:2849–2861.

22. Haak D, Gable K, Beeler T, Dunn T: **Hydroxylation of Saccharomyces cerevisiae Ceramides Requires Sur2p and Scs7p**. *J Biol Chem* 1997, **272**:29704–29710.

23. Cowart LA, Obeid LM: **Yeast sphingolipids: Recent developments in understanding biosynthesis, regulation, and function**. *Biochim Biophys Acta Molecular and Cell Biology of Lipids* 2007, **1771**:421–431.

24. Becker SA, Price ND, Palsson BO: **Metabolite coupling in genome-scale metabolic networks**. *BMC Bioinformatics* 2006, **7**:111.

25. **Saccharomyces Genome Database** [http://www.yeastgenome.org/].

26. Kanehisa M, Goto S: **KEGG: kyoto encyclopedia of genes and genomes**. *Nucleic Acids Res* 2000, **28**:27–30.

27. The UniProt Consortium: **The Universal Protein Resource (UniProt) in 2010**. *Nucleic Acids Res* 2009, **38**:D142–D148.

28. Caspi R, Altman T, Dale JM, Dreher K, Fulcher CA, Gilham F, Kaipa P, Karthikeyan AS, Kothari A, Krummenacker M, Latendresse M, Mueller LA, Paley S, Popescu L, Pujar A, Shearer AG, Zhang P, Karp PD: **The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases**. *Nucleic Acids Res* 2010, **38**:D473.

29. Sud M, Fahy E, Cotter D, Brown A, Dennis EA, Glass CK, Merrill AH, Murphy RC, Raetz

CRH, Russell DW, Subramaniam S: **LMSD: LIPID MAPS structure database**. *Nucleic Acids Res* 2007, **35**:D527–D532.

30. Degtyarenko K, de Matos P, Ennis M, Hastings J, Zbinden M, McNaught A, Alcantara R, Darsow M, Guedj M, Ashburner M: **ChEBI: a database and ontology for chemical entities of biological interest**. *Nucleic Acids Res* 2007, **36**:D344–D350.

31. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A, Cuellar AA, Dronov S, Gilles ED, Ginkel M, Gor V, Goryanin II, Hedley WJ, Hodgman TC, Hofmeyr JH, Hunter PJ, Juty NS, Kasberger JL, Kremling A, Kummer U, Le Novère N, Loew LM, Lucio D, Mendes P, Minch E, Mjolsness ED, Nakayama Y, Nelson MR, Nielsen PF, Sakurada T, Schaff JC, Shapiro BE, Shimizu, TS, Spence HD, Stelling J, Takahashi K, Tomita M, Wagner J, Wang J: **The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models**. *Bioinformatics* 2003, **19**:524–531.

32. **yeastnet** [http://yeast.sf.net/].

33. Bornstein BJ, Keating SM, Jouraku A, Hucka M: **LibSBML: an API Library for SBML**. *Bioinformatics* 2008, **24**:880–881.

34. Schellenberger J, Que R, Fleming RMT, Thiele I, Orth JD, Feist AM, Zielinski DC, Bordbar A, Lewis NE, Rahmanian S, Kang J, Hyduke DR, Palsson BØ: **Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0**. *Nat Protoc* 2011, **6**:1290–1307.

35. Uemura S, Kihara A, Inokuchi J, Igarashi Y: **Csg1p and Newly Identified Csh1p Function in Mannosylinositol Phosphorylceramide Synthesis by Interacting with Csg2p**. *J Biol Chem* 2003, **278**:45049–45055.

36. Idkowiak-Baldys J, Grilley MM, Takemoto JY: **Sphingolipid C4 hydroxylation influences properties of yeast detergent-insoluble glycolipid-enriched membranes**. *FEBS Letters* 2004, **569**:272–276.

37. Dickson RC: **Roles for sphingolipids in Saccharomyces cerevisiae**. *Adv Exp Med Biol* 2010, **688**:217–231.

38. Kastenmayer JP, Ni L, Chu A, Kitchen LE, Au W-C, Yang H, Carter CD, Wheeler D, Davis RW, Boeke JD, Snyder MA, Basrai MA: **Functional genomics of genes with small open reading frames (sORFs) in S. cerevisiae**. *Genome Res.* 2006, **16**:365–373.

39. Breslow DK, Cameron DM, Collins SR, Schuldiner M, Stewart-Ornstein J, Newman HW, Braun S, Madhani HD, Krogan NJ, Weissman JS: **A comprehensive strategy enabling high-resolution functional analysis of the yeast genome**. *Nat Methods* 2008, **5**:711–718.

40. Vallée B, Riezman H: **Lip1p: a novel subunit of acyl-CoA ceramide synthase**. *EMBO J* 2005, **24**:730–741.

41. Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connelly C, Davis K, Dietrich F, Dow SW, El Bakkoury M, Foury F, Friend SH, Gentalen E, Giaever G, Hegemann JH, Jones T, Laub M, Liao H, Liebundguth N, Lockhart DJ, Lucau-Danila A, Lussier M, M'Rabet N, Menard P, Mittmann M, Pai C, Rebischung C, Revuelta JL, Riles L, Roberts CJ, Ross-MacDonald P, Scherens B, Snyder M, Sookhai-Mahadeo S, Storms RK, Véronneau S, Voet M, Volckaert G, Ward TR, Wysocki R, Yen GS, Yu K, Zimmermann K, Philippsen P, Johnston M, Davis RW: **Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis**. *Science* 1999, **285**:901–906.

42. Giaever G, Chu AM, Ni L, Connelly C, Riles L, Véronneau S, Dow S, Lucau-Danila A, Anderson K, André B, Arkin AP, Astromoff A, El-Bakkoury M, Bangham R, Benito R, Brachat S, Campanaro S, Curtiss M, Davis K, Deutschbauer A, Entian K-D, Flaherty P, Foury F, Garfinkel DJ, Gerstein M, Gotte D, Güldener U, Hegemann JH, Hempel S, Herman Z, Jaramillo DF, Kelly DE, Kelly SL, Kötter P, LaBonte D, Lamb DC, Lan N, Liang H, Liao H, Liu L, Luo C, Lussier M, Mao R, Menard P, Ooi SL, Revuelta JL, Roberts CJ, Rose M, Ross-Macdonald P, Scherens B, Schimmack G, Shafer B, Shoemaker DD, Sookhai-Mahadeo S, Storms RK, Strathern JN, Valle G, Voet M, Volckaert G, Wang C, Ward TR, Wilhelmy J, Winzeler EA, Yang Y, Yen G, Youngman E, Yu K, Bussey H, Boeke JD, Snyder M, Philippsen P, Davis RW, Johnston M: **Functional profiling of the Saccharomyces cerevisiae genome**. *Nature* 2002, **418**:387–391.

43. Duarte NC, Herrgård MJ, Palsson BO: **Reconstruction and validation of Saccharomyces cerevisiae iND750, a fully compartmentalized genome-scale metabolic model**. *Genome Research* 2004, **14**:1298.

44. Nookaew I, Jewett MC, Meechai A, Thammarongtham C, Laoteng K, Cheevadhanarak S, Nielsen J, Bhumiratana S: **The genome-scale metabolic model iIN 800 of Saccharomyces cerevisiae and its validation: a scaffold to query lipid metabolism**. *BMC Syst Biol* 2008, **2**:71.

45. Covert MW, Xiao N, Chen TJ, Karr JR: **Integrating metabolic, transcriptional regulatory and signal transduction models in Escherichia coli**. *Bioinformatics* 2008, **24**:2044.

46. Herrgard MJ, Swainston N, Dobson P, Dunn WB, Arga KY, Arvas M, Buthgen N, Borger S, Costenoble R, Heinemann M, Hucka M, Le Novere N, Li P, Liebermeister W, Mo ML, Oliveira AP, Petranovic D, Pettifer S, Simeonidis E, Smallbone K, Spasie I, Weichart D, Brent R, Broomhead DS, Westerhoff HV, Kurdar B, Penttila M, Klipp E, Palsson BO, Sauer U, Oliver SG, Mendes P, Nielsen J, Kell DB: **A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology**. *Nat Biotech* 2008, **26**:1155–1160.

47. Henry CS, DeJongh M, Best AA, Frybarger PM, Linsay B, Stevens RL: **High-throughput generation, optimization and analysis of genome-scale metabolic models**. *Nat Biotechnol* 2010, **28**:977–982.

48. Poolman MG, Bonde BK, Gevorgyan A, Patel HH, Fell DA: **Challenges to be faced in the reconstruction of metabolic networks from public databases**. *IEE Proc Syst Biol* 2006, **153**:379.

49. Ejsing CS, Sampaio JL, Surendranath V, Duchoslav E, Ekroos K, Klemm RW, Simons K, Shevchenko A: **Global analysis of the yeast lipidome by quantitative shotgun mass spectrometry**. *PNAS* 2009, **106**:2136.

50. Hannun YA, Obeid LM: **Many ceramides**. *J Biol Chem* 2011, **286**:27855–27862.

51. Zomorrodi AR, Maranas CD: **Improving the iMM 904 S. cerevisiae metabolic model using essentiality and synthetic lethality data**. *BMC Syst Biol* 2010, **4**:178.

CHAPTER 4 - YEAST 5 – AN EXPANDED RECONSTRUCTION OF THE SACCHAROMYCES CEREVISIAE METABOLIC NETWORK[1]

Abstract

Background: Efforts to improve the computational reconstruction of the *Saccharomyces cerevisiae* biochemical reaction network and to refine the stoichiometrically constrained metabolic models that can be derived from such a reconstruction have continued since the first stoichiometrically constrained yeast genome scale metabolic model was published in 2003. Continuing this ongoing process, we have constructed an update to the Yeast Consensus Reconstruction, Yeast 5. The Yeast Consensus Reconstruction is a product of efforts to forge a community-based reconstruction emphasizing standards compliance and biochemical accuracy via evidence-based selection of reactions. It draws upon models published by a variety of independent research groups as well as information obtained from biochemical databases and primary literature.

Results: Yeast 5 refines the biochemical reactions included in the reconstruction, particularly reactions involved in sphingolipid metabolism; updates gene-reaction annotations; and emphasizes the distinction between reconstruction and stoichiometrically constrained model. Although it was not a primary goal, this update also improves the accuracy of model prediction of viability and auxotrophy phenotypes and increases the number of epistatic interactions. This update maintains an emphasis on standards compliance, unambiguous metabolite naming, and computer-readable annotations available through a structured document format. Additionally, we have developed MATLAB scripts to evaluate the model's predictive accuracy and to demonstrate basic model applications, such as simulating aerobic and anaerobic growth. These scripts, which provide an independent tool for evaluating the performance of various

[1] This chapter has previously been published as: Heavner BD, Smallbone K, Barker B, Mendes P, Walker LP: **Yeast 5 – an Expanded Reconstruction of the *Saccharomyces cerevisiae* Metabolic Network**. *BMC Systems Biology* 2012, 6: 55.

stoichiometrically constrained yeast metabolic models using flux balance analysis, are included as Appendix 2.

Conclusions: Yeast 5 expands and refines the computational reconstruction of yeast metabolism and improves the predictive accuracy of a stoichiometrically constrained yeast metabolic model. It differs from previous reconstructions and models by emphasizing the distinction between the yeast metabolic reconstruction and the stoichiometrically constrained model, and makes both available at http://yeast.sf.net/ as separate systems biology markup language (SBML) files. Through this separation, we intend to make the modeling process more accessible, explicit, transparent, and reproducible.

Background

Efforts to improve the computational reconstruction of the *Saccharomyces cerevisiae* biochemical reaction network and to refine the metabolic models that can be derived from such a reconstruction have continued since the first yeast genome scale metabolic model was published [1]. The distinction between reconstruction (termed GEnome scale Network REconstructions (GENREs) [2]) and derived models (termed GEnome scale Models (GEMs) [3]) remains important to differentiate between the established biochemical knowledge included in a GENRE and the modeling assumptions required for analysis or simulation with a GEM. A GENRE serves as a structured knowledge base of established biochemical facts, while a GEM is a model which supplements the established biochemical information with additional (potentially hypothetical) information to enable computational simulation and analysis. Examples of widely used yeast GENREs include the Kyoto Encylopedia of Genes and Genomes, KEGG [4]), and the Yeast Biochemical Pathway Database, YeastCyc [5]. The history of yeast GEMs has recently been reviewed [6].

Though a GEM may be considered finished when it is sufficient for a particular modeling

application, the effort to build a complete and accurate GENRE is ongoing as biochemical research continues (even information that is fundamental to the construction of a GENRE, such as genome annotation, is considered to be a working hypothesis and subject to ongoing revision [7]). Reflecting the ongoing process of yeast GEM and GENRE improvement [6], we have constructed an update to the Yeast Consensus Reconstruction [8]. The Yeast Consensus Reconstruction is a product of efforts to forge a community-based reconstruction emphasizing standards compliance and biochemical accuracy via evidence-based selection of reactions. It draws upon models published by a variety of independent research groups [1, 9–12], as well as information obtained from biochemical databases and primary literature. Thus, the Yeast Consensus Reconstruction serves as an example of the community-based approach which has given rise to the concept of a "reconstruction annotation jamboree" [13]. Though there remain many challenging problems to implementing and maintaining community-based science [14], the jamboree approach to network reconstruction and model building has also been successfully applied to build a consensus reconstruction and model of *Salmonella Typhimurium* LT2 [15].

The Yeast Consensus Reconstruction has benefited from the continued involvement of the broader research community. Previous updates to the Yeast Consensus Reconstruction [16] have focused on filling gaps in the metabolic reconstruction to improve network connectivity in a graph-theoretical sense, expanding the reconstruction of portions of metabolism that had not been included in previous reconstructions, and enabling Flux Balance Analysis (FBA) [17] by adding the necessary (but hypothetical) transport reactions and sink reactions (such as the biomass reaction). These low-confidence reactions in the Yeast Consensus Reconstruction are annotated with use of specialized Systems Biology Ontology (SBO) terms [18], an approach designed to facilitate differentiation between the higher-confidence reactions which form the Yeast GENRE and the lower confidence reactions required to evaluate a GEM with FBA. Enabling FBA of the consensus reconstruction has resulted in increased interest in applying the model to guide bioengineering efforts [19, 20]. In turn, this increased interest has stimulated

community participation, which has highlighted opportunities for further improving the Consensus Yeast Reconstruction GENRE and the derived GEM.

Therefore, we decided to undertake an update to the Yeast Consensus Reconstruction to refine the biochemical reactions included in the GENRE, particularly reactions involved in sphingolipid metabolism [20, 21]; to review gene-reaction annotation; to emphasize and clarify the distinction between GENRE and GEM; to facilitate application of the GEM for bioengineering applications; and to solicit and facilitate further collaboration among researchers who wish to further improve the yeast GENRE and GEM. Although it was not a primary goal, this update also improves the accuracy of GEM phenotype predictions due to the incorporation of reaction constraints from previous models and relevant literature. We endeavored to conduct this update while maintaining an emphasis on standards compliance, unambiguous metabolite naming, and computer-readable annotations available through a structured document format. The metabolites included in Yeast 5 are unambiguously annotated with their identifiers in the Chemical Entitites of Biological Interest (ChEBI) database [22], and reactions are annotated with the PubMed ID of primary literature evidence justifying the reaction's inclusion in the reconstruction.

We have incorporated the results of these efforts to the consensus reconstruction to produce Yeast 5. Yeast 5 expands and refines the yeast GENRE and improves the predictive accuracy of the yeast GEM. Further, it differs from previous reconstructions and models by emphasizing the distinction between the yeast GENRE and GEM, and makes both available as separate systems biology markup language (SBML) files [23]. Through this separation of GENRE and GEM, we intend to make the modeling process more explicit, transparent, and reproducible. Both files are available from YeastNet (http://yeast.sf.net/). In addition to the GEM and GENRE SBML files, we have developed MATLAB scripts to evaluate the model's predictive accuracy and to demonstrate basic model applications, such as simulating aerobic and anaerobic metabolism with Yeast 5. These scripts, which provide an independent tool for

evaluating the performance of various yeast GEMs are included in Appendix 2.

Results

Improvements to facilitate community use and collaboration

As a consensus reconstruction and model, Yeast 5 will depend upon community use and suggested modifications for future improvement. The Yeast 5 GENRE may serve as a resource for construction of new models, of both genome and smaller scale. Thus, in addition to the emphasis on standards compliance, the Yeast 5 SBML files include coding conventions to facilitate use by popular software. Recognizing that MATLAB is a commonly used platform for systems and computational biology [24], we include scripts demonstrating use of both the GENRE and GEM with the SBML Toolbox [25] and the COBRA Toolbox [26]. These MATLAB scripts, testYeastModel.m, modelToReconstruction.m, and fluxDistribution.m, are available in Appendix 2. Additionally, as discussed in Materials and Methods, Yeast 5 includes conventions for exchange reactions and boundary species which are used in the MATLAB-compatible COBRA Toolbox, although such conventions are not currently included in SBML specifications [23].

Yeast GENRE Changes

The Yeast 5 GENRE is an evidence-driven biochemical knowledge-base. It does not include the low-confidence or hypothetical reactions and metabolites required to conduct FBA, nor does it include constraints on reaction reversibility which may be added in the course of model-building. Since it does not include compounds such as “biomass”, reactions such as “growth”, or hypothetical intercompartmental transport reactions, the Yeast 5 GENRE is more specific than Yeast 4, which did not differentiate between GENRE and GEM. It contains 1418 metabolites which participate in 2110 reactions, catalyzed by 918 verified *Saccharomyces cerevisiae* genes. In comparison, Yeast 4 includes 1481 metabolites, 2030 reactions, and 924

genes. The Yeast 5 GENRE does not include genes annotated in the Saccharomyces Genome Database (SGD) as “dubious” or “uncharacterized”, while Yeast 4 includes 4 such genes (YFR055W, YML082W, YPL275W, and YPL276W). Yeast 4 included reactions annotated with 29 open reading frames which are not included in the Yeast 5 GENRE. However, the Yeast 5 GENRE includes 23 open reading frames which are not included in Yeast 4. The 29 open reading frames which are present in Yeast 4 but not Yeast 5 are: YAL014C, YAL030W, YAR042W, YCR073W-A, YDL019C, YDR313C, YDR331W, YDR468C, YEL011W, YEL013W, YER093C, YFR055W, YGR199W, YHR005C, YHR073W, YMR068W, YNL006W, YNR034W, YOL078W, YPL145C, YPL275W, YPL276W, YOR237W, YIL105C, YJL058C, YJR160C, YKL203C, YML082W, and YNL047C. These open reading frames were removed from reaction annotations because of inadequate literature evidence supporting the Yeast 4 annotation. The 23 open reading frames included in the Yeast 5 GENRE but not in Yeast 4 are: YBR001C, YBR058C-A, YBR161W, YBR199W, YDR196C, YDR367W, YGR138C, YGR277C, YMR241W, YMR278W, YMR298W, YPL023C, YPL053C, YPL189W, YPR156C, YOR175C, YGL084C, YIL083C, YJL200C, YKL088W, YKL132C, YML056C, and YNL029C. Additional annotation information about these ORFs is provided in Appendix 2 as Supplemental Table 4.5 (ORFs present in Yeast 4 but not in Yeast 5) and Supplemental Table 4.6 (ORFs present in Yeast 5 but not in Yeast 4).

In addition to being more specific, the Yeast 5 GENRE is also more complete than Yeast 4. Sphingolipid metabolism has been acknowledged to be a poorly reconstructed portion of the yeast metabolic network since the first yeast GEM, iFF708 [1]. Yeast 5 incorporates suggested literature-referenced refinements to sphingolipid metabolism [20, 21]. Thus the Yeast 5 GENRE contains the most complete reconstruction to date of the broad suite of complex sphingolipids that has been observed in yeast [27].

Yeast GEM Changes

The Yeast 5 GEM is a model, and thus includes biomass demand functions and low-confidence reactions such as intercompartmental transport reactions which enhance network connectivity and enable FBA. It also includes reaction directionality constraints to improve the accuracy of model phenotype predictions and exchange reactions which allow model users to simulate a growth medium. The Yeast 5 GEM includes more reactions and metabolites than previously published yeast GEMs (Table 1), though it includes 6 fewer genes than Yeast 4. The Yeast 5 GEM includes reactions annotated with 918 different open reading frames, accounting for 18.5% of the 4949 verified open reading frames included in the Saccharomyces Genome Database [28] as of October 12, 2011. The Yeast 5 GEM has 326 more directionally constrained reactions than Yeast 4 (69% of all reactions in the Yeast 5 GEM are constrained, compared to 55.6% in Yeast 4). The majority of these new constraints are applied to reactions involved in cofactor utilization or production, with a particular emphasis on reactions involving ATP/ADP. Reactions involving NAD(P)/H are constrained where literature evidence supports irreversible reactions *in vivo.*

Since the Yeast GEM includes both general classes of compounds (e.g."fatty acid") and specific members of these classes (e.g. "octanoate"), we use non-reversible encapsulating reactions called "isa" reactions to provide pathways from specific to generic compounds (e.g., octanoate "isa" fatty acid). The use of "isa" reactions is discussed further in the Discussion section. To accommodate the more specific biochemistry included in the Yeast 5 GENRE, the Yeast 5 GEM includes 261 "isa" reactions, compared to 162 in Yeast 4. Additionally, the Yeast 5 GEM includes 2 different lipid pseudoreactions, which create the "lipid" portion of biomass (details of the Yeast 5 GEM biomass definition are included in Appendix 2 as Supplemental Table 4.3). As described in the "simulating yeast growth" discussion, including two differing biomass definitions enables simulation of anaerobic yeast metabolism despite the incomplete reconstruction of yeast lipid biochemistry in the Yeast 5 GENRE.

Table 1: Comparison of yeast metabolic models

	Yeast 5	Yeast 4[a]	iMM904bs[b]	iND750[c]
Model description				
Number of metabolites	1655	1481	1228	1061
Number of reactions	2110	2030	1575	1266
Number of genes	918	924	904	750
Number of dubious genes[d]	0	4	17	17
Blocked reactions[e]	38%	26%	31%	41%
Viability analysis				
Sensitivity[f]	97%	95%	93%	96%
Specificity[g]	47%	44%	57%	43%
Positive predictive value[h]	86%	85%	89%	87%
Negative predictive value[i]	84%	73%	69%	77%
Geometric mean[j]	46%	42%	53%	41%
Auxotrophy analysis				
Auxotroph-inducing genes included[k]	70	73	73	69
Correct auxotroph predictions	73%	66%	69%	58%
Incorrectly predicted as viable in minimal media	24%	30%	27%	39%
Incorrectly predicted as inviable in supplemented media	3%	4%	4%	3%
Epistatic interaction analysis[l]				
Epistatic interations (% of pairwise genes)	16%	-	15%	21%
Total number of epistatic interactions	65,730	-	63,176	57,808
Average Additional Interactions per additional gene[m]	196.46	-	34.63	-

[a]yeast 4: [16]
[b]iMM904bs: [49]
[c]iND750: [9]
[d]dubious genes are ORFs annotated as "dubious" (809 ORFs) or "uncharacterized" (857 ORFs) in SGD
[e]blocked reactions cannot carry flux
[f]True positive / (true positive + false negative)
[g]True negative / (true negative + false positive)
[h]True positive / (true positive + false positive)
[i]True negative / (true negative + false negative)
[j]see Kuepfer et al. 2005 for discussion of applying geometric mean
[k]see Appendix 2 Supplemental Tables 4.1 and 4.2 for list of essential genes and auxotroph-inducing genes
[l]At 50%x50% flux reductions. See Appendix 2 Supplemental Table 4.4 for other restriction levels.
[m]number of interactions/number of genes, compared to iND750 as base case

Yeast GEM Performance

Though improving gene essentiality predictions was not a primary objective of updating Yeast 4, simulations using Yeast 5 have increased agreement with a list of essential genes and genes which cause auxotrophies; more realistic prediction of internal fluxes; and increased number of genetic interactions. Additionally, simulations with the Yeast 5 GEM predict auxotrophies resulting from gene deletions better than other recent yeast GEMs, and predict gene essentiality with accuracy comparable to other recent yeast GEMs (Table 1).

Gene essentiality predictions

Since the phenotype resulting from a gene mutation is dependent upon media and environmental conditions as well as changes to the metabolic network, the use of gene essentiality predictions as a metric for model evaluation requires careful definition of both simulation assumptions and of the data set used for comparison between simulation and observation. This is particularly important if such a metric is to be used for comparison of different GEMs. We document our approach to simulating gene essentiality in Materials and Methods, and in the testYeastmodel.m MATLAB script included in Appendix 2l.

The Yeast 5 GEM includes reactions annotated with 918 genes. 144 of these genes are included in a list of essential genes we compiled from the Saccharomyces Genome Deletion project [29] and annotation included in Saccharomyces Genome Database [28] (Appendix 2 Supplemental Tables 4.1 and 4.2). An additional 70 genes are included in a list of genes causing auxotrophies when deleted (the construction of these gene lists is described in Materials and Methods). The remaining 704 genes in the model are not on the compiled lists of essential or auxotroph-inducing genes, and are therefore considered.

The results of single gene deletion simulations conducted via FBA of the Yeast 5 GEM using a simulated glucose-limited defined media are summarized in Table 2. The model

predicted that biomass could be produced in 684 of the 704 cases in which genes annotated as inessential or non-auxotrophic were deleted (true positive results) and that biomass could not be produced in 20 cases where these inessential or non-auxotrophic genes were deleted (false negative results). Thus, the model has a 97.2% sensitivity for this list of essential genes. If the model simulation predicted that biomass could be produced following a gene deletion, the deleted gene was not listed as essential or auxotroph-inducing in 86% of the cases (an 86% positive predictive value).

Table 2: Single-gene Deletion Results (918 genes)

684 (75%) True Positives (model simulation predicts growth when inessential genes are deleted)	**20** (2%) False Negatives (model simulation predicts no growth when inessential genes are deleted)
113 (12%) False Positives (model simulation predicts growth when essential genes are deleted)	**101** (11%) True Negatives (model simulation predicts no growth when essential genes are deleted)

The model predicted that biomass could not be produced following deletion of 101 of the of the 214 genes included on the essential or auxotrophy-inducing gene lists (true negative results), but that biomass could still be produced following deletion of 113 genes included on those lists (false positive results). Thus, the model has a 47% specificity for this list of essential genes. If the model simulation predicted that biomass could not be produced following a gene deletion, the deleted gene was listed as essential or auxotroph-inducing in 83.5% of the cases (an 83.5% negative predictive value).

Comparing Yeast 5 GEM knockout simulations with our list of essential and auxotroph-

inducing genes yields a geometric mean overall predictive accuracy (as suggested by [10] of 45.88%, an improvement over Yeast 4's 41.61% geometric mean accuracy for this gene list. These results, including comparison of the simulations using other recently published GEMs and the same essential and auxotroph-inducing gene lists, are summarized in Table 1.

Auxotroph-inducing mutations

To extend our analysis of Yeast 5 GEM simulation capabilities, we conducted additional FBA growth simulation focusing on the 70 genes included in both the Yeast 5 GEM and the list of genes whose mutation or deletion causes auxotrophies. In 51 single-gene deletion simulations, the model predicted that biomass could not be produced in minimal media but could be produced in a supplemented media, the expected behavior for an auxotroph mutant (see Materials and Methods for more information about our approach to simulated media). In 17 cases, model simulation predicted that biomass could be produced in minimal media, and thus did not accurately predict the auxotrophic phenotype. In 2 cases, model simulation predicted that biomass could not be produced in either minimal or maximal media, and so the model incorrectly predicted that the gene deletion could not be saved by media supplementation.

Auxotroph phenotypes have not previously been a metric used to evaluate yeast GEMs. We found that simulation with the Yeast 5 GEM had better agreement with observed auxotrophic phenotypes than simulations conducted with other previously published GEMs (Table 1).

Flux predictions

Optimal solutions found when conducting FBA of the Yeast 5 GEM to maximize biomass flux include fluxes through key internal reactions, a flux distribution which better matches fluxes observed *in vivo* [30] than solutions found when applying FBA to Yeast 4 (Table 3). Specifically, applying FBA to the Yeast 5 GEM in a simulated glucose-limited aerobic environment predicts that the reactions of glycolysis and the TCA cycle have fluxes, and that

ethanol is not produced. When the model constraints are adjusted to simulate an anaerobic environment, FBA predicts fluxes through the reactions of glycolysis, but not the TCA cycle, and ethanol is produced. Thus, simulations with the Yeast 5 GEM reflect the shift from respiratory to fermentative metabolism which is observed in oxygen-limited yeast cultures. Simulations using the Yeast 4 GEM do not reflect this same behavior.

Table 3: Yeast 5 and Yeast 4 simulated flux predictions

	yeast 5	yeast 4
Sample FBA flux predictions		
glucose-limited, aerobic growth rate	0.09	0.17
glucose-limited, anaerobic growth rate	0.02	0
aerobic flux through glycolysis[a]	1.33	0.89
anaerobic flux through glycolysis	1.85	-
aerobic flux through TCA cycle[b]	1.06	0.01
anaerobic flux through TCA cycle	0	-
aerobic ethanol production	0	0
anaerobic ethanol production	1.74	-

Fluxes are normalized to the glucose uptake flux, which is set to 1 mmol/g dry weight/h

[a]glycolysis flux measured through pyruvate kinase

[b]TCA cycle flux measured through malate dehydrogenase

Increased number of genetic interactions

Recognizing recent efforts to investigate system-level organization of cellular metabolism via the phenotypic effects of multiple gene deletions using yeast GEMs [31–35], we compared the number of epistatic interactions predicted by growth simulations using the Yeast 5 GEM with the number of interactions predicted by simulations using the iMM904 [12] and iND750 [9] models (Table 1). When reaction fluxes were restricted to 50% of wild-type in a pairwise manner, we found that a lower percentage of genes included in the Yeast 5 and iMM904 GEMs were predicted to exhibit epistatic interactions in FBA simulations, but the expanded size of these models led to an increased number of total epistatic interactions compared to the iND750 GEM. If the number of interactions are averaged over the number of genes in each model, the Yeast 5 GEM adds an average of 196.5 new epistatic interactions per additional gene,

and iMM904 adds an average of 34.6 new interactions per gene. The number of interactions predicted using each of these models using varying levels of flux restriction for each gene and reaction pair are provided in Appendix 2 as Supplemental Table 4.4.

Limitations

Research to expand our understanding of yeast metabolism and biochemistry is ongoing, and the process of integrating established biochemical knowledge into computational reconstructions lags research advancements. Thus, the Yeast 5 GENRE is not a complete reconstruction of the yeast biochemical network, and though it offers improvements over earlier models, simulations using the Yeast 5 GEM do not fully reflect observed biological phenomena. GENRE and GEM limitations suggest opportunities for future efforts to improve computational reconstruction of established biochemistry and can highlight portions of metabolism that are ripe for further research [36].

A limitation of the Yeast 5 GENRE which suggests future opportunities for improving the Yeast Reconstruction is that due to the lack of information about enzyme specificity or the metabolic significance of variation among similar chemical species, the Yeast 5 GENRE uses general classes of chemical compounds rather than the enumeration of many similar compounds. For example, Yeast 5 generalizes the many possible triglyceride compounds which may by incorporating fatty acyl moieties of varying length [27] into a single model species, called "triglyceride". Yeast 5 also includes similar generalized species for other compounds, particularly those involved in lipid and sterol metabolism. Though this approach is also followed by previous yeast GENREs and other metabolic pathway tools, expansion of such general species by differentiating among biochemically relevant species has been shown to be a successful approach to expanding computational reconstructions of metabolic networks [20]. The appropriate level of detail or generalization for metabolic (or biochemical) network reconstruction depends upon the intended use of a GEM, and would be expected to change in the

future as our knowledge of enzyme specificity and the metabolic relevance of differences among similar chemical compounds advances.

As with other reconstructed metabolic networks, the continued existence of blocked pathways (Table 1) highlights that our knowledge of yeast metabolism is incomplete. Such blocked reactions are an important tool for documenting portions of metabolism that would benefit from further research [36]. The Yeast 5 GENRE remains limited by knowledge gaps. Where our knowledge of intercompartmental transport of metabolites is limited, such gaps pose a particular challenge to FBA. Thus, the Yeast 5 GEM includes hypothetical transport reactions to better connect portions of the metabolic network that are unconnected in a graph-theoretical sense. Such hypothetical transport reactions are annotated with SBO term [18] SBO:397 ("omitted process").

Optimal solutions found when conducting FBA on the Yeast 5 GEM may include fluxes that differ from those observed in yeast: we have found optimal solutions in which mitochondrial coenzyme A is synthesized *in situ* rather than transported from the cytoplasm, and model growth simulations incorrectly predict that yeast is not a pantothenate auxotroph or a nicotinic acid auxotroph in anaerobic conditions. Additionally, although the model predicts that biomass can only be produced anaerobically if the biomass definition is modified (see Materials and Methods), the reason that simulated anaerobic biomass production using an unmodified biomass definition is blocked is not because of the biological requirement of yeast fatty acid desaturate for oxygen. Instead, simulated anaerobic biomass production is blocked due to other, as yet unidentified limitations in the reconstruction of phospholipid and sterol biosynthesis. Refining the solution space to more closely match observed biological phenomena through improved reconstruction or expanded constraints remains an ongoing research effort for reconstructed metabolic networks.

Due to varying interpretations of experimental evidence, uncertainty of metabolic

mechanism, and varying approaches taken as modelers work to reconstruct different portions of metabolism, it is likely that the Yeast 5 consensus reconstruction has additional limitations which will be discovered as it is used. These limitations provide opportunities for continued research to improve the computational reconstruction and simulation of the yeast biochemical network. Thus, though Yeast 5 consists of a more complete reconstruction and more accurate model of yeast metabolism than previous efforts, the goal of building a complete and accurate computational reconstruction of yeast metabolism must remain an ongoing community effort.

Discussion

Yeast 5 is the most recent update to the consensus reconstruction of the yeast metabolic network. It consists of a genome-scale reconstruction (GENRE), a genome scale model (GEM), and MATLAB scripts designed to facilitate evaluation of yeast GEMs and to demonstrate simulation and analysis using the COBRA and SBML toolboxes. This update improves the consensus reconstruction's coverage of established biochemical knowledge, and improves the predictive ability of simulations using the yeast GEM. The included scripts lower the barriers for the research community to use the model and to contribute to the collaborative effort to improve the computational reconstruction of yeast metabolism.

Models and reconstructions

We emphasize the distinction between a reconstruction, or GENRE, and a model, or GEM, to more clearly delineate the established biochemical knowledge of a reconstruction from the assumptions and hypotheses that are required for modeling and simulation. This distinction makes the modeling process more transparent and reproducible, essential attributes for community-based scientific efforts such as the consensus reconstruction of yeast metabolism.

Evaluating yeast reconstructions and models

As evidence of improvements to the consensus yeast reconstruction, we have presented a

comparison of viability, auxotrophy predictions, and genetic interaction effects, produced using FBA of the Yeast 5 GEM and other metabolic models (Table 1). However, we emphasize that such metrics of phenotype predictive ability must be evaluated with great care. This need for careful use of such metrics has been discussed previously [37], but we have found that this point deserves additional emphasis. Specifically, the two goals of expanding the reconstruction of metabolic networks and improving model predictions of mutant viability may be contradictory in some situations. The Yeast 5 GEM sensitivity, specificity, positive predictive value, negative predictive value, and geometric mean could all be improved by the reduction of false positive predictions - simulations which predict that biomass can be produced although reactions annotated as being catalyzed by "essential genes" have been blocked. The number of false positive predictions could be reduced by expanding the biomass function to require products of reactions annotated with essential genes, by removing parallel pathways to force fluxes through reactions annotated with essential genes, or by removing metabolites and reactions which create dead-end pathways which include reactions annotated with essential genes (a method used to improve lethality prediction metrics when the iLL672 model was derived from iFF708 [10]). However, while such techniques improve a model's ability to predict single-gene mutant viability, they also reduce the scope of a GENRE as a structured knowledge base of established biochemical facts.

Expanding the reconstruction of a metabolic network would be expected to increase both dead-end pathways and network redundancy. Dead ends would be introduced through the inclusion of established knowledge regarding pathways that are not fully elucidated. In such pathways, the production of intermediates may be established, but their fate is not yet known. In such cases, expanding the reconstruction of established knowledge would not be expected to improve a model's ability to predict single-gene mutant viability, and so such metrics would not reflect the expanded scope of the reconstruction. Another example of network expansion which may not be reflected in single-gene deletion metrics is expanding the reconstruction's coverage

of isoenzymes. Network redundancy increases through expanded inclusion of isoenzymes, which introduce parallel metabolic paths for the production of chemical intermediates or products. In the absence of regulatory constraints (which are beyond the scope of a metabolic reconstruction), these parallel pathways would increase the rate of false positive prediction since a metabolic pathway from substrate to product would exist in the reconstruction, even if a given isoenzyme were individually insufficient to support growth *in vivo*. Thus, a metabolic model based upon a reconstruction with improved coverage of established biochemistry or isoenzymes would make less accurate predictions of individual gene essentiality than a model based upon a less complete reconstruction.

A second problem with metrics based upon lists of essential genes is that gene essentiality is dependent upon strain, media, and environmental conditions (for example, [38] identify mutants which are inositol auxotrophs only at elevated temperatures). Though a general definition of "essential" could imply "in complex media at 30º C", the difficulties of computationally reconstructing complex media and the lack of integration of temperature effects on metabolic networks means that there remains an element of subjectivity in defining a list of essential genes. Researchers have previously used different data sets to define gene essentiality for model analysis [9, 12]. Thus, if model predictivity metrics are to be used, care must be taken to ensure a common list of essential genes when evaluating different models by such metrics. Yeast 5 includes the MATLAB script testYeastModel.m to document the list of genes we considered essential for our comparison of yeast GEMs.

That essentiality metrics must be used with care and considered in context is not to say that such metrics are without value, however. Model simulation results that differ from *in vivo* experiments can guide efforts to improve computational reconstruction or to highlight the need for additional biochemical investigation of metabolic dead ends. Simulations resulting in false negative results, in which the model predicts that biomass cannot be formed, but *in vivo* experiments have observed growth, suggest that the reconstruction is incomplete or the model

has limitations such as incorrect biomass definition or missing simulated media components. Indeed, the use of gap filling algorithms to improve phenotype predictive metrics for metabolic models by adding hypothesized gene functions or reactions is considered standard practice for GEM development [39].

Like auxotroph phenotype predictions, predicted epistatic interactions have not previously been used as a metric for comparing yeast GEMs. And like other metrics based upon phenotypic prediction, the use of epistatic interactions must also be qualified. Specifically, expansion in the number of reactions in the model is a likely contributor to the amount of simulated epistasis. The number of genetic interactions can also be increased by pleiotropy, or the number of reactions associated to a particular gene. Yeast 5 has an 8.43% increase in the mean number of reactions per gene compared to iMM904. The most prominent example is ISC1, a gene important in sphingolipid metabolism. ISC1 is included in annotation for 60 reactions in Yeast 5, but only 18 in iMM904. ISC1 accounts for a 1% increase in pleiotropy by itself. FOX2, a multi-function enzyme involved in beta-oxidation, ranked highest for pleiotropy in iMM904 with 23 reactions. It is also associated with 23 reactions in Yeast 5. Additional trends in positive or negative epistasis across different types of mutations can also be observed for these models (Appendix 2 Supplemental Table 4.4; Xu & Barker in preparation).

Generic demand reactions in Yeast 5 - towards a functional biomass definition

The yeast consensus reconstruction draws upon data sources with varying levels of compound specificity. For example, the KEGG database includes a general representation of yeast sphingolipid metabolism, while recent suggestions for changes to Yeast 4 include more specific chemical species [20]. Thus, some Yeast 5 reactions which are derived from KEGG use generic species as substrates or products, while reactions derived from other sources use more specific species. To accommodate the formation of generic chemical species for reactions which consume them while preserving biochemical accuracy in reactions that have more specific

biochemistry, the Yeast GEM includes “isa” reactions. Examples of generic species produced by “isa” reactions include “complex sphingolipid”, “fatty acid”, and “acyl-CoA”.

Where “isa” reactions are reversible, they can introduce unrealistic interconversion of metabolites. For example, since octanoate is a fatty acid and hexadecanoate is a fatty acid, reversible “isa” reactions in a model would create a nonrealistic pathway by which octanoate could be converted to hexadecanoate via the “isa” reaction, instead of through biochemical pathways which have been documented *in vivo*. In order to prevent such non-realistic interconversion fluxes, Yeast 5 “isa” reactions are not reversible. Thus, more general compounds can only serve as sinks of more specific compounds in the Yeast 5 GEM, and not as sources.

An unanticipated result of this approach to varying levels of biochemical specificity in the model is that “isa” reactions effectively embed the logical OR into the model. Thus, where the biomass definition includes “lipid” as a required component, this objective function can be satisfied by any of the compounds that can be converted to “lipid” via an “isa” reaction (Figure 1). Thus, although “biomass” must be defined if maximizing biomass production is the objective function for FBA, “biomass” does not need to be a specifically determined compound with a fixed stoichiometry for FBA to be successfully applied to stoichiometrically constrained metabolic reconstructions.

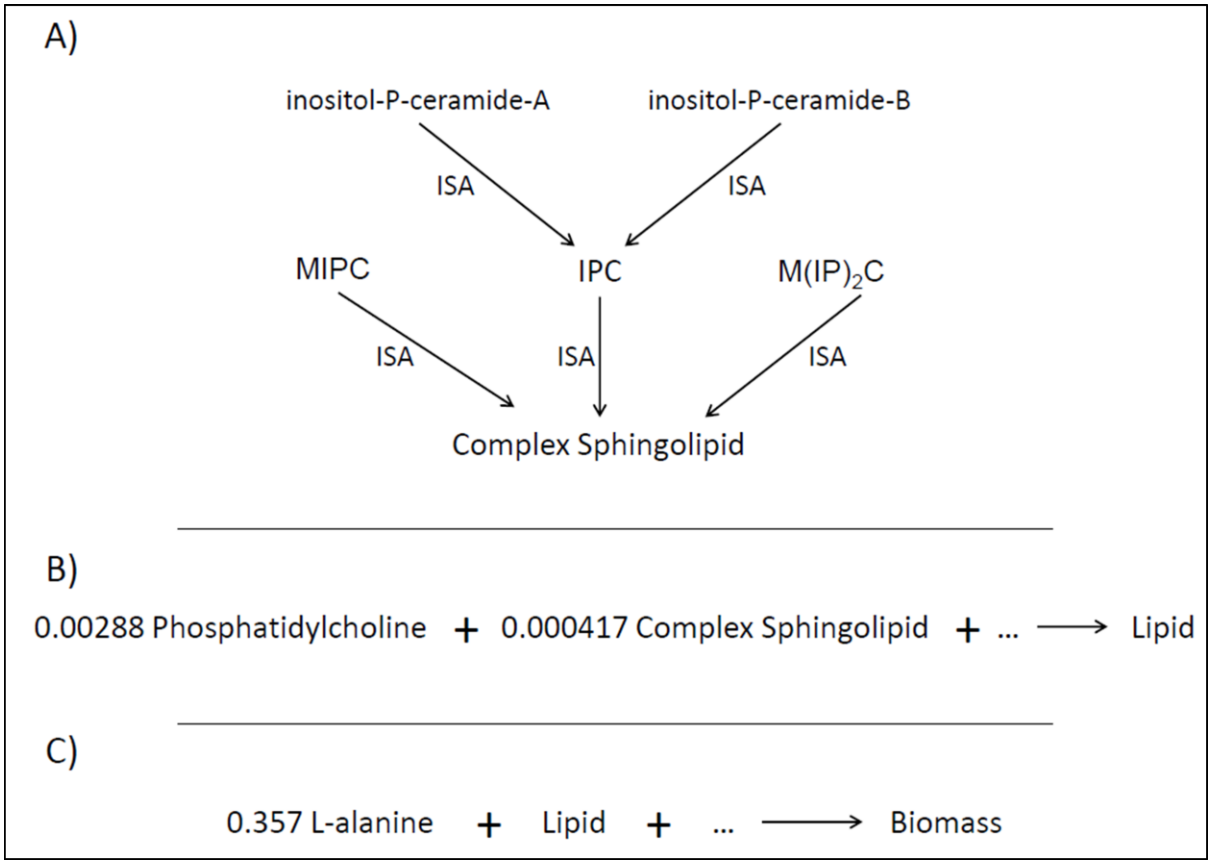


Figure 1: Using "isa" reactions

Yeast 5 uses "isa" reactions to encapsulate specific chemical species within more general classes. For example, **A)** the specific species inositol-P-ceramide-A "isa" inositol phosphoceramide (IPC). In turn, IPC "isa" complex sphingolipid. **B)** Complex sphingolipids participate in the stoichiometrically constrained reaction which produces the species "lipid". **C)** The lipid species is a component of biomass. This hierarchical model structure embeds logic in the biomass definition: biomass consists of L-alanine AND phosphatidylcholine AND (inositol-P-ceramide-A OR Inositol-P-ceramide-B OR any of the 88 other complex sphingolipids included in the reconstruction). A model user is free to constrain the fluxes which produce specific complex sphingolipids to model an observed lipid composition, or may leave the model unconstrained if the more general biomass definition is sufficient for their needs.

Continuing efforts to reconstruct yeast metabolism - an invitation for continued community involvement

Computational reconstruction and modeling of yeast metabolism is an ongoing project. Suggestions for improving the yeast consensus reconstruction or derived models should be submitted to network.reconstruction@manchester.ac.uk. Metabolites and enzymes should be unambiguously identified, using existing model or database (ChEBI or UniProt) identifiers. New reactions should be supplied with primary evidence for their mechanism and catalysis, via PubMed identifiers. Reactions without evidence should have clear reasons for their proposed addition.

We also invite researchers to submit models derived from the yeast consensus reconstruction for hosting at http://yeast.sf.net/. Assumptions and constraints should be documented, for example with code documenting how to build the GEM from the Yeast GENRE. Such models may be submitted for publication separately from updates to the Yeast GENRE, and may follow the iNNXXX naming convention which has been previously used for identifying GEMs [40].

Conclusions

The Yeast 5 expansion of the Yeast Consensus Reconstruction refines the computational reconstruction of yeast metabolism and improves the predictive accuracy of a stoichiometrically constrained yeast metabolic model. It refines the biochemical reactions included in the reconstruction, particularly reactions involved in sphingolipid metabolism; updates gene-reaction annotations; and emphasizes the distinction between reconstruction (GENRE) and stoichiometrically constrained model (GEM). This update also improves the accuracy of model prediction of viability and auxotrophy phenotypes and increases the number of epistatic interactions. Yeast 5 differs from previous reconstructions and models by emphasizing the distinction between the yeast metabolic reconstruction and the stoichiometrically constrained

model, and makes both available at http://yeast.sf.net/ as separate systems biology markup language (SBML) files. Through this separation, we intend to make the modeling process more accessible, explicit, transparent, and reproducible. The Yeast Consensus Reconstruction remains a community-based resource which emphases standards compliance and biochemical accuracy via evidence-based selection of reactions.

Though Yeast 5 consists of a more complete reconstruction and more accurate model of yeast metabolism than previous efforts, the goal of building a complete and accurate computational reconstruction of yeast metabolism must remain an ongoing community effort. As Yeast 5 limitations are identified, they provide opportunities for continued research to improve the computational reconstruction and simulation of the yeast biochemical network.

Methods

Yeast 5 scope

The scope of Yeast Consensus Reconstruction was originally determined by the data sets used for its construction: the iMM904 [12] and iLL672 [10] models, which included information from the KEGG and SGD databases, along with other sources. Subsequently, the Consensus Reconstruction was expanded [16] to include information (particularly focusing on lipid metabolism) from the iIN800 model [11]. Yeast 5 further expands the scope of reconstruction to refine details of sphingolipid metabolism [20, 21]. Though the stoichiometrically constrained approach can theoretically be expanded to include all biochemical reactions in the organism being modeled [41], the Yeast reconstruction is currently limited to the yeast metabolic network. Although Yeast 5 is not strain-specific, the auxotroph information we used for our analysis focused on auxotrophies documented in *S. cerevisiae* reference strain SC288C. Since multiple yeast genome sequences are now available, future updates to the Yeast Consensus Reconstruction may become strain specific.

Deriving reconstruction (GENRE) from model (GEM)

We must maintain a distinction between the *reconstruction* of yeast metabolism, an evidence-driven biochemical knowledge-base, and its corresponding *model*, which relies on a number of assumptions to make quantitative flux predictions [39]. We discriminate between the two through the use of the Systems Biology Ontology (SBO) [18]. Specifically, reactions marked up with specific SBO terms may be automatically removed from a model to create a reconstruction. Encapsulating "isa" reactions of the form "A isa B" are annotated with SBO:395 ("encapsulating process"). Other reactions without literature evidence, that are omitted in the reconstruction, such as biomass production and most transport reactions without an associated transporter are annotated with SBO:397 ("omitted process"). Reaction constraints (lower and upper bounds) are also removed in the reconstruction. The transformation is performed using the SBMLToolbox [25].

Yeast Model Conventions

The Yeast 5 GEM includes conventions to facilitate model analysis using the COBRA Toolbox, such as standardized representation of exchange reactions, though such conventions are not required in the SBML specification. However, we have chosen to rely upon the SBML standard for encoding reaction and metabolite annotation, rather than the nonstandard custom notes field currently used by the COBRA toolbox. This information, which includes metabolite ChEBI identifiers and literature references supporting reaction inclusion, is encoded in the model .sbml file and is available to MATLAB users via the SBML toolbox. An additional model convention is that the Yeast 5 GEM includes biomass as a species in the model to help make our approach to simulating biomass production more explicit.

Exchange reactions in the Yeast 5 GEM follow a convention of using compounds in the model as reactants, leading to exchange reactions of the form "reactant -> ", with an entry of +1 in the stoichiometric matrix. Thus, positive flux values for exchange reactions represent

compounds produced in FBA simulation, and negative flux values represent compounds consumed. Since reactions must include both substrate and product in SBML, we have followed the COBRA toolbox convention of denoting exchange reaction species which lay outside the model with the subscript "_b". These species are not loaded into the COBRA toolbox data structure, and serve only as placeholders for exchange reaction substrates.

We have chosen to include biomass as a species in the Yeast 5 GEM. Thus, when conducting FBA on the Yeast 5 GEM, the biomass exchange reaction can be selected if biomass optimization is the desired objective function. For modeling purposes, the species "biomass" is produced in the Yeast 5 GEM via a reaction which consumes 37 biomass precursor compounds to produce biomass, ADP, protons, and phosphate. The biomass precursors include water, polysaccharides, nucleotides, amino acids, riboflavin, sulfate, and the general species "lipid". The "lipid" species in the Yeast 5 GEM serves a function similar to Zanghellini et al.'s "virtual membrane particle" [42] or Nookaew et al.'s lipid species [11]. It is a lumped species which incorporates many different specific lipid compounds. We have used two different definitions of "lipid" to enable simulation of aerobic and anaerobic biomass production (see "Simulating yeast growth"). The aerobic lipid pseudoreaction consumes 15 lipids and sterols, to produce the generic "lipid" species, while the anaerobic lipid pseudoreaction omits the sterols 14-demethyllanosterol and ergosta-5,7,22,24(28)-tetraen-3beta-ol. Thus, there are two Yeast 5 GEM biomass definitions: an aerobic biomass consisting of 52 compounds, and an anaerobic biomass definition consisting of 50 compounds. The growth, biomass pseudoreaction, and lipid pseudoreactions are detailed in Appendix 2 Supplemental Table 4.3.

Constraining reactions in the Yeast 5 GEM

The Yeast 5 GEM has over 300 more directionally constrained reactions than Yeast 4. Such constraints incorporate thermodynamic information into the GEM, and often serve in part to limit Type III cycling, which arises from the linear programing approach, but are

thermodynamically infeasible [43]. It is noted that such cycling can also be eliminated by minimizing the total flux, or by applying geometric FBA [44]. New constraints were added by re-evaluating “isa” reactions and applying a heuristic approach to reactions involving energy-carrying cofactors (ATP and NAD(P)), supported by evidence from other models, literature, and pathway databases. As discussed in “Generic demand reactions in Yeast 5 - towards a functional biomass definition”, directional constraints were added to “isa” reactions in the Yeast 5 GEM to prevent unrealistic interconversion of chemically distinct metabolites via fluxes through general species. Additional constraints were added with a heuristic approach similar to [45] which focuses on reactions which may produce ATP and those that use NAD or NADP as cofactors. We directionally constrained such reactions only when such constraints were supported by constraints in the iND750 and iMM904 reconstructions and by the directionality specified in the BioCyc database [5].

Simulating yeast growth

The Yeast 5 GEM includes 170 exchange reactions, each of which defines a compound that can be included as a medium component for simulation purposes. As distributed, the simulated media is a glucose-limited minimal aerobic medium with constrained limited uptake of glucose, and unconstrained exchange of oxygen, ammonium, protons, iron(2+), phosphate, potassium, sodium, sulfate and water. To simulate anaerobic growth, the oxygen exchange reaction may be constrained to disallow oxygen uptake. Simulating anaerobic growth also requires that the simulated media be supplemented by allowing exchange of ergosterol, lanosterol, zymosterol and phosphatidate, and the biomass definition be changed by removing 14-demethyllanosterol and ergosta-5,7,22,24(28)-tetraen-3beta-ol from the “lipid” definition. These requirements reflect the observation that yeasts require sterols [46, 47] and fatty acids [48] when cultured under rigidly anaerobic conditions. However, from a modeling perspective, these requirements arise from the biomass definition, for which the biochemistry is not firmly established, and from the reconstruction of sterol metabolism, which is incomplete in the Yeast 5

GEM.

Growth simulations were performed using the Cobra toolbox [26]. We have included MATLAB scripts which demonstrate simulation of aerobic and anaerobic growth, investigation of internal fluxes, and gene essentiality tests in Appendix 2.

Gene deletion simulation

Gene deletion simulations were performed using the Cobra toolbox [26]. To compare model gene essentiality predictions with observed phenotypes, gene lists were compiled for genes considered essential, and those which have been observed to cause auxotrophies when deleted. The essential gene list was compiled beginning with 1191 unique open reading frames reported to cause inviable mutants upon deletion in the Saccharomyces Genome Deletion Project [29] and the YKOv2 supplemental data set available from http://www-sequence.stanford.edu/group/yeast_deletion_project/data_sets.html. Since this data set screened in complex media (which is only incompletely accounted for with FBA simulation due to the limited number of exchange reactions), we refined the list of "essential" genes first by removing any ORFs which have not been reported as verified in the SGD database [28], and then by supplementing it with a list of gene mutations which have been reported to cause auxotroph phenotypes. The auxotroph-inducing gene list was generated by by searching the SGD database for "inviable" and "auxotrophy" phenotypes. Since the biochemistry of temperature signaling is beyond the scope of the Yeast 5 GEM, we removed temperature-dependent inositol auxotroph mutants [38] from this list. The lists of genes we used for evaluating model essentiality predictions, as well as a MATLAB script which can be used to evaluate other yeast GEMs essentiality predictions, are included as Supplemental Information.

Simulation of genetic interactions

Genetic interactions were quantified with the nonscaled multiplicative definition of

epistasis, $\varepsilon = W_{xy} - W_x W_y$ [34]. In this definition, W_x and W_y are fitness scores for organisms with a mutation in genes x and y, respectively, and W_{xy} is the fitness of the organism with both mutations present. Any $\varepsilon \neq 0$ indicates a genetic interaction under the assumption that both of the genes independently and multiplicatively contribute to fitness. Following the example of [31], we quantified fitness as maximum biomass production rate (as determined by FBA) relative to the rate of biomass production in simulations conducted with the wild-type model. As suggested by [34], we simulated genetic perturbation by limiting flux through all reactions associated to a specific enzyme by a fixed amount of the wild-type geometric FBA flux (0%, 10%, ..., 90%). Constraining the flux to fractions of wild-type flux allows investigation of essential reactions, as well as inessential reactions. The number of epistatic interactions for each level of flux restriction is included in Supplemental Table 4.

REFERENCES

1. Förster J, Famili I, Fu P, Palsson B, Nielsen J: **Genome-Scale Reconstruction of the Saccharomyces cerevisiae Metabolic Network**. *Genome Res* 2003, **13**:244–253.

2. Price ND, Reed JL, Palsson BO: **Genome-scale models of microbial cells: evaluating the consequences of constraints**. *Nat Rev Microbiol* 2004, **2**:886–897.

3. Feist AM, Herrgaard MJ, Thiele I, Reed JL, Palsson BO: **Reconstruction of biochemical networks in microorganisms**. *Nat Rev Microbiol* 2008, **7**:129–143.

4. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M: **KEGG for representation and analysis of molecular networks involving diseases and drugs**. *Nucleic Acids Res* 2010, **38**:D355–360.

5. Karp PD, Ouzounis CA, Moore-Kochlacs C, Goldovsky L, Kaipa P, Ahrén D, Tsoka S, Darzentas N, Kunin V, López-Bigas N: **Expansion of the BioCyc collection of pathway/genome databases to 160 genomes**. *Nucleic Acids Res* 2005, **33**:6083–6089.

6. Österlund T, Nookaew I, Nielsen J: **Fifteen years of large scale metabolic modeling of yeast: Developments and impacts**. *Biotechnology Advances* 2011.

7. Fisk DG, Ball CA, Dolinski K, Engel SR, Hong EL, Issel-Tarver L, Schwartz K, Sethuraman A, Botstein D, Michael Cherry J, The Saccharomyces Genome Database Project: **Saccharomyces cerevisiae S288C genome annotation: a working hypothesis**. *Yeast* 2006, **23**:857–865.

8. Herrgård MJ, Swainston N, Dobson P, Dunn WB, Arga KY, Arvas M, Büthgen N, Borger S, Costenoble R, Heinemann M, Hucka M, Le Novère N, Li P, Liebermeister W, Mo ML, Oliveira AP, Petranovic D, Pettifer S, Simeonidis E, Smallbone K, Spasié I, Weichart D, Brent R, Broomhead DS, Westerhoff HV, Kürdar B, Penttilä M, Klipp E, Palsson BØ, Sauer U, Oliver SG, Mendes P, Nielsen J, Kell DB: **A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology**. *Nat Biotechnol* 2008, **26**:1155–1160.

9. Duarte NC, Herrgård MJ, Palsson BO: **Reconstruction and validation of Saccharomyces cerevisiae iND750, a fully compartmentalized genome-scale metabolic model**. *Genome Res* 2004, **14**:1298.

10. Kuepfer L, Sauer U, Blank LM: **Metabolic functions of duplicate genes in Saccharomyces cerevisiae**. *Genome Res* 2005, **15**:1421–1430.

11. Nookaew I, Jewett MC, Meechai A, Thammarongtham C, Laoteng K, Cheevadhanarak S, Nielsen J, Bhumiratana S: **The genome-scale metabolic model iIN 800 of Saccharomyces cerevisiae and its validation: a scaffold to query lipid metabolism**. *BMC Syst Biol* 2008, **2**:71.

12. Mo ML, Palsson BØ, Herrgård MJ: **Connecting extracellular metabolomic measurements to intracellular flux states in yeast**. *BMC Syst Biol* 2009, **3**:37.

13. Thiele I, Palsson BØ: **Reconstruction annotation jamborees: a community approach to systems biology**. *Mol Syst Biol* 2010, **6**:361.

14. Kitano H, Ghosh S, Matsuoka Y: **Social engineering for virtual "big science" in systems biology**. *Nat Chem Biol* 2011, **7**:323–326.

15. Thiele I, Hyduke DR, Steeb B, Fankam G, Allen DK, Bazzani S, Charusanti P, Chen F-C, Fleming RM, Hsiung CA, De Keersmaecker SC, Liao Y-C, Marchal K, Mo ML, Özdemir E, Raghunathan A, Reed JL, Shin S-I, Sigurbjörnsdóttir S, Steinmann J, Sudarsan S, Swainston N, Thijs IM, Zengler K, Palsson BO, Adkins JN, Bumann D: **A community effort towards a knowledge-base and mathematical model of the human pathogen Salmonella Typhimurium LT2**. *BMC Syst Biol* 2011, **5**:8.

16. Dobson PD, Jameson D, Simeonidis E, Lanthaler K, Pir P, Lu C, Swainston N, Dunn WB, Fisher P, Hull D, Brown M, Oshota O, Stanford NJ, Kell DB, King RD, Oliver SG, Stevens RD, Mendes P: **Further developments towards a genome-scale metabolic model of yeast**. *BMC Syst Biol* 2010, **4**:145.

17. Edwards JS, Covert M, Palsson BO: **Metabolic modelling of microbes: the flux-balance approach**. *Environ Microbiol* 2002, **4**:133–140.

18. Le Novère N, Courtot M, Laibe C: **Adding semantics in kinetics models of biochemical pathways**. *Proceedings of the 2nd International ESCEC Symposium on Experimental Standard Conditions on Enzyme Characterizations* 2007.

19. Cvijovic M, Bordel S, Nielsen J: **Mathematical models of cell factories: moving towards the core of industrial biotechnology**. *Microbial Biotechnology* 2011, **4**:572–584.

20. Heavner BD, Henry SA, Walker LP: **Evaluating Sphingolipid Biochemistry in the Consensus Reconstruction of Yeast Metabolism**. *Industrial Biotechnology* 2012, **8**:72–78.

21. Kavun Ozbayraktar FB, Ulgen KO: **Stoichiometric network reconstruction and analysis of yeast sphingolipid metabolism incorporating different states of hydroxylation**. *Biosystems* 2011, **104**:63–75.

22. Degtyarenko K, de Matos P, Ennis M, Hastings J, Zbinden M, McNaught A, Alcantara R, Darsow M, Guedj M, Ashburner M: **ChEBI: a database and ontology for chemical entities of biological interest**. *Nucleic Acids Res* 2007, **36**:D344–D350.

23. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, Arkin AP, Bornstein BJ, Bray D, Cornish-Bowden A, Cuellar AA, Dronov S, Gilles ED, Ginkel M, Gor V, Goryanin II, Hedley WJ, Hodgman TC, Hofmeyr JH, Hunter PJ, Juty NS, Kasberger JL, Kremling A, Kummer U, Le Novère N, Loew LM, Lucio D, Mendes P, Minch E, Mjolsness ED, Nakayama Y, Nelson MR, Nielsen PF, Sakurada T, Schaff JC, Shapiro BE, Shimizu, TS, Spence HD, Stelling J, Takahashi K, Tomita M, Wagner J, Wang J: **The systems biology markup language**

(SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 2003, **19**:524–531.

24. Hübner K, Sahle S, Kummer U: **Applications and trends in systems biology in biochemistry**. *FEBS Journal* 2011, **278**:2767–2857.

25. Keating SM, Bornstein BJ, Finney A, Hucka M: **SBMLToolbox: an SBML toolbox for MATLAB users**. *Bioinformatics* 2006, **22**:1275.

26. Schellenberger J, Que R, Fleming RMT, Thiele I, Orth JD, Feist AM, Zielinski DC, Bordbar A, Lewis NE, Rahmanian S, Kang J, Hyduke DR, Palsson BØ: **Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0**. *Nat Protoc* 2011, **6**:1290–1307.

27. Ejsing CS, Sampaio JL, Surendranath V, Duchoslav E, Ekroos K, Klemm RW, Simons K, Shevchenko A: **Global analysis of the yeast lipidome by quantitative shotgun mass spectrometry**. *PNAS* 2009, **106**:2136.

28. Engel SR, Balakrishnan R, Binkley G, Christie KR, Costanzo MC, Dwight SS, Fisk DG, Hirschman JE, Hitz BC, Hong EL, Krieger CJ, Livstone MS, Miyasato SR, Nash R, Oughtred R, Park J, Skrzypek MS, Weng S, Wong ED, Dolinski K, Botstein D, Cherry JM: **Saccharomyces Genome Database provides mutant phenotype data**. *Nucleic Acids Res* 2010, **38**:D433–436.

29. Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connelly C, Davis K, Dietrich F, Dow SW, El Bakkoury M, Foury F, Friend SH, Gentalen E, Giaever G, Hegemann JH, Jones T, Laub M, Liao H, Liebundguth N, Lockhart DJ, Lucau-Danila A, Lussier M, M'Rabet N, Menard P, Mittmann M, Pai C, Rebischung C, Revuelta JL, Riles L, Roberts CJ, Ross-MacDonald P, Scherens B, Snyder M, Sookhai-Mahadeo S, Storms RK, Véronneau S, Voet M, Volckaert G, Ward TR, Wysocki R, Yen GS, Yu K, Zimmermann K, Philippsen P, Johnston M, Davis RW: **Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis**. *Science* 1999, **285**:901–906.

30. Blank LM, Kuepfer L, Sauer U: **Large-scale 13C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast**. *Genome biology* 2005, **6**:R49.

31. Segrè D, DeLuna A, Church GM, Kishony R: **Modular epistasis in yeast metabolism**. *Nat Genet* 2004, **37**:77–83.

32. Deutscher D, Meilijson I, Kupiec M, Ruppin E: **Multiple knockout analysis of genetic robustness in the yeast metabolic network**. *Nat Genet* 2006, **38**:993–998.

33. Harrison R, Papp B, Pal C, Oliver SG, Delneri D: **Plasticity of genetic interactions in metabolic networks of yeast**. *PNAS* 2007, **104**:2307–2312.

34. He X, Qian W, Wang Z, Li Y, Zhang J: **Prevalent positive epistasis in Escherichia coli and Saccharomyces cerevisiae metabolic networks**. *Nat Genet* 2010, **42**:272–276.

35. Snitkin ES, Segrè D: **Epistatic Interaction Maps Relative to Multiple Metabolic Phenotypes**. *PLoS Genetics* 2011, **7**:e1001294.

36. Rolfsson O, Palsson BØ, Thiele I: **The human metabolic reconstruction Recon 1 directs hypotheses of novel human metabolic functions**. *BMC Syst Biol* 2011, **5**:155.

37. Snitkin E, Dudley A, Janse D, Wong K, Church G, Segrè D: **Model-driven analysis of experimentally determined growth phenotypes for 465 yeast gene deletion mutants under 16 different conditions**. *Genome Biol* 2008, **9**:R140.

38. Villa-García MJ, Choi MS, Hinz FI, Gaspar ML, Jesch SA, Henry SA: **Genome-wide screen for inositol auxotrophy in Saccharomyces cerevisiae implicates lipid metabolism in stress response signaling**. *Mol Genet Genomics* 2011, **285**:125–149.

39. Thiele I, Palsson BO: **A protocol for generating a high-quality genome-scale metabolic reconstruction**. *Nat Protoc* 2010, **5**:93–121.

40. Reed JL, Vo TD, Schilling CH, Palsson BO: **An expanded genome-scale model of Escherichia coli K-12 (iJR904 GSM/GPR)**. *Genome Biol* 2003, **4**:R54.51 – R54.12.

41. Aho T, Almusa H, Matilainen J, Larjo A, Ruusuvuori P, Aho K-L, Wilhelm T, Lähdesmäki H, Beyer A, Harju M, Chowdhury S, Leinonen K, Roos C, Yli-Harja O: **Reconstruction and Validation of RefRec: A Global Model for the Yeast Molecular Interaction Network**. *PLoS ONE* 2010, **5**:e10662.

42. Zanghellini J, Natter K, Jungreuthmayer C, Thalhammer A, Kurat CF, Gogg-Fassolter G, Kohlwein SD, von Grünberg H-H: **Quantitative modeling of triacylglycerol homeostasis in yeast - metabolic requirement for lipolysis to promote membrane lipid synthesis and cellular growth**. *FEBS Journal* 2008, **275**:5552–5563.

43. Beard DA, Liang S, Qian H: **Energy balance for analysis of complex metabolic networks**. *Biophys J* 2002, **83**:79–86.

44. Smallbone K, Simeonidis E: **Flux balance analysis: A geometric perspective**. *J Theor Biol* 2009, **258**:311–315.

45. Kümmel A, Panke S, Heinemann M: **Systematic assignment of thermodynamic constraints in metabolic network models**. *BMC Bioinformatics* 2006, **7**:512.

46. Bloch KE: **Sterol structure and membrane function**. *Crit Rev Biochem* 1983, **14**:47–92.

47. Thomas KC, Hynes SH, Ingledew WM: **Initiation of anaerobic growth of Saccharomyces cerevisiae by amino acids or nucleic acid bases: ergosterol and unsaturated fatty acids cannot replace oxygen in minimal media**. *J Ind Microbiol Biotechnol* 1998, **21**:247–253.

48. Ratledge C, Evans CT: **Lipids and their Metabolism**. In *The Yeasts*. 2nd edition. edited by Rose AH, Harrison JS San Diego, CA: Academic Press; 1989, **3**.

49. Szappanos B, Kovács K, Szamecz B, Honti F, Costanzo M, Baryshnikova A, Gelius-Dietrich G, Lercher MJ, Jelasity M, Myers CL, Andrews BJ, Boone C, Oliver SG, Pál C, Papp B: **An integrated approach to characterize genetic interaction networks in yeast metabolism**. *Nat Genet* 2011, **43**:656–662.

CHAPTER 5 - REVISITING SYSTEMS-LEVEL ENGINEERING OF NONFERMENTATIVE METABOLISM IN YEAST WITH THE YEAST 5 METABOLIC MODEL

Abstract

Background: A recent report of the successful use of the iND750 yeast metabolic model to engineer endogenous one-carbon metabolism in yeast presents an opportunity to assess the robustness of simulation predictions to varying implementations, and to evaluate the recently published version of the Yeast Consensus Reconstruction (Yeast) model by comparing simulation results using the newer model to those predicted by simulations using iND750.

Results: We were able to obtain results similar but not identical to those reported previously using the iND750 model, but not with version 5.01 of the Yeast model. Thus, we generated a literature-supported set of updates to Yeast model which enable similar predictions of formate production. We found that results of an *in silico* screen for formate-producing mutants using the Yeast model are sensitive to constraints on reactions involving NAD(P) cofactors and generated a literature-supported set of constraints on these reactions which improves screening results.

We found errors in the reconstruction of C-1 metabolism in iND750, but also found that simulations with iND750 better predict formate-producing phenotypes involving "uncharacterized" yeast open reading frames than simulations with the updated Yeast model. In contrast, we found that simulations using the Yeast model better predict auxotrophic phenotypes and predicted formate-producing mutants which were not found in a screen using iND750. The code we used for this study, including references for suggested model changes, is included in Appendix 3.

Conclusions: The computational prediction of formate-producing yeast mutants is sensitive to implementation details and reaction constraints when using either the iND750 or the Yeast model. The models we evaluated had different strengths and weaknesses: the screen using the Yeast model predicts more formate-producing strains with mutations in pathways which would be expected to affect C-1 metabolism than the screen using iND750, but iND

better predicts phenotypes involving mutations in uncharacterized or dubious ORFs. Our results suggest that comparative analysis of constraint based models is a useful tool for finding phenotypes of poorly characterized yeast genes, characterizing redox metabolism, and for improving the computational reconstruction of the yeast metabolic network.

Background

In the past sixteen years, many stoichiometrically constrained genome-scale computational models of the yeast metabolic network have been published, each intended to improve upon previous modeling efforts [1]. These models have been applied to develop metabolic engineering strategies for the production of products such as ethanol, succinate, glycerol, vanillin, and sesquirtpenes [2]. The development and application of such models has been facilitated by the development of COnstraint Based Reconstruction and Analysis (COBRA) methods and software, particularly the MATLAB COBRA Toolbox [3]. Many metabolic engineering strategies have been proposed based upon computational simulation, but few have been implemented or rigorously evaluated *in vivo* [4]. The development and characterization of a formate-producing yeast strain by Kennedy et al. [5] is a noteworthy example of successful implementation of a metabolic engineering strategy developed through a constraint-based modeling approach.

Kennedy et al.'s metabolic engineering efforts focused on improving the production of non-fermentative by-products, a class of "biologically interesting and commercially attractive small molecules". They developed and implemented a model-driven metabolic engineering strategy to increase endogenous formic acid secretion in yeast, using flux balance analysis of the iND750 model of yeast metabolism [6] to simulate growth of triple-knockout mutants, then screened the simulation results for mutants predicted to produce formate (the possible triple deletions were in addition to deletion of *FDH1* and *FDH2*, which encode yeast's formate dehydrogenase enzymes). Kennedy et al. reported 8 strains which their simulations predicted would produce formate, at efficiencies ranging from 72.3% to 1.2% (Table 5.1). Selecting the top-ranked formate producer from their *in silico* screen, the authors then successfully constructed a formate-producing yeast strain, PSY3642, by knocking out the computationally suggested *ALT2*, *FUM1*, and *ZWF1* genes, each of which catalyze

reactions that affect fluxes through central carbon metabolism (Figure 5.1). Additionally, the authors constructed and characterized a *ZWF1* knockout strain and a *FUM1 ZWF1* double knockout strain, observing that formate production increased in aerobic batch cultures with each additional gene deletion.

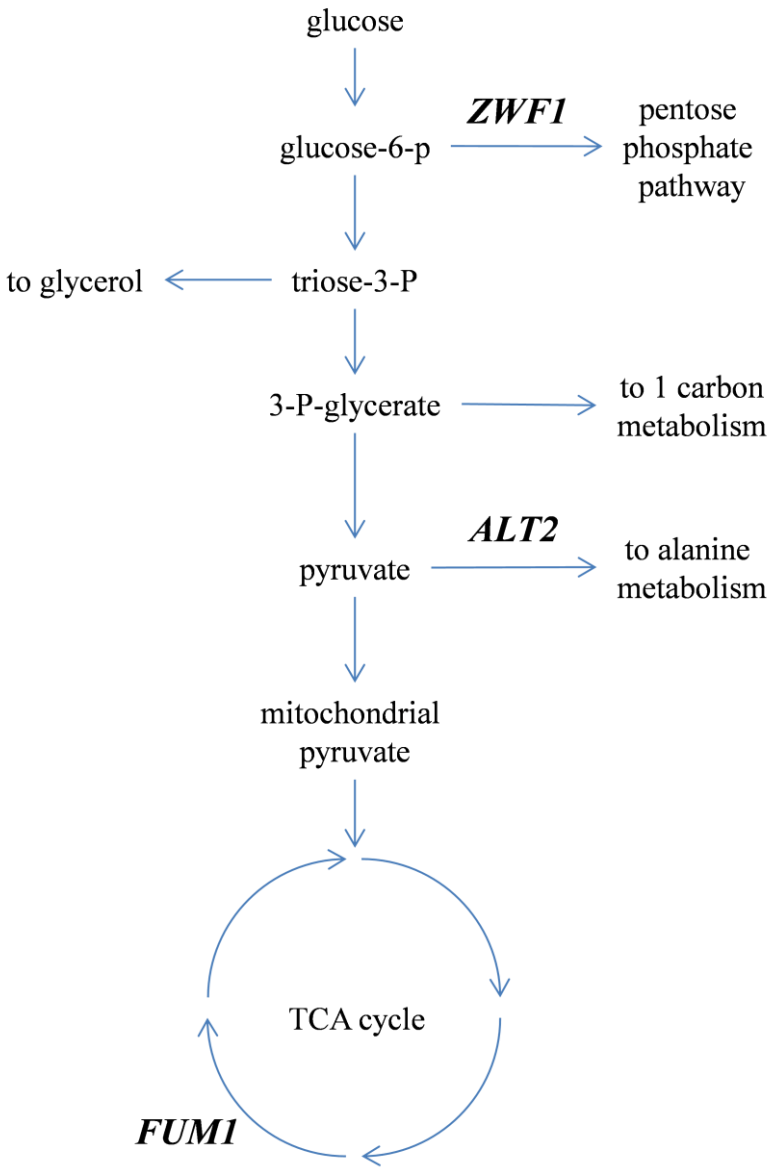
Table 5.1 – Previously reported strains predicted to produce formate by *in silico* simulation

Genotype	Predicted Efficiency (%)[a]
alt2 zwf1 fum1 fdh1 fdh2	72.3
aat2 zwf1 fum1 fdh1 fdh2	72.2
cat2 zwf1 fum1 fdh1 fdh2	72.0
cat2 rpe1 fum1 fdh1 fdh2	71.7
cat2 fbp1 fum1 fdh1 fdh2	30.5
cat2 yat2 slc1 fdh1 fdh2	2.4
cat2 yat2 cho1 fdh1 fdh2	2.3
cat2 yat2 alt2 fdh1 fdh2	1.2

[a]One hundred percent efficiency is defined as four formic acid molecules per glucose.

Kennedy et al. reported 8 strains which their *in silico* screen of simulated yeast deletion mutants predicted would produce formate [5].

Figure 5.1 – Target genes relationship to central carbon metabolism



In the iND750 model, the genes deleted in Kennedy et al.'s PSY3642 strain annotate reactions in branchpoints from central carbon metabolism. Though the Saccharomyces Genome Database annotates the open reading frame coding for *ALT2* as "uncharacterized", the iND750 model annotates a reaction involved in alanine biosynthesis with this gene.

At least nine genome scale models have been published since iND750, including *iLL672* [7], *iMH805/775* [8], *iJH732* [9], *iIN800* [10], three versions of the Yeast Consensus Reconstruction [11–13], *iMM904* [14], and *iAZ900* [15]. In this study, we replicate some of the previously reported results, and conduct a computational screen of yeast mutants following a procedure similar to that used by Kennedy et al., but implemented with the

COBRA Toolbox instead of custom code, and using the most recently published yeast metabolic model, Yeast 5 [13] along with the older iND750 model.

Results

Our simulations using iND750 replicated the computational prediction of a formate-producing phenotype for 5 of 8 previously reported cases (Table 5.2), and our *in silico* screen predicted formate-producing mutants, including the previously predicted top-performing strain, though it was not the highest predicted formate-producing mutant in our screen. We found that simulations using iND750 correctly predicted that deletion of the *ALT2* gene in addition to *FUM1* and *ZWF1* in a *fdh1 fdh2* background strain led to increased formate production, as has been observed *in vivo*.

Table 5.2 - Comparing model predictions of formate production

Genotype	Predicted Efficiency (%)[a]				
	Kennedy et al.	iND750	Yeast 5.01	Yeast 5.30	Updated Yeast 5.30
alt2 zwf1 fum1 fdh1 fdh2	72.3	72.1	0.0	69.6	71.9
aat2 zwf1 fum1 fdh1 fdh2	72.2	72.0	0.0	69.6	71.9
cat2 zwf1 fum1 fdh1 fdh2	72.0	71.8	0.0	69.6	71.9
cat2 rpe1 fum1 fdh1 fdh2	71.7	71.4	0.0	69.2	71.5
cat2 fbp1 fum1 fdh1 fdh2	30.5	30.4	0.0	69.6	2.1
cat2 yat2 slc1 fdh1 fdh2	2.4	-	0.0	0.0	0.0
cat2 yat2 cho1 fdh1 fdh2	2.3	-	no growth	no growth	no growth
cat2 yat2 alt2 fdh1 fdh2	1.2	-	0.0	0.0	0.0

[a]One hundred percent efficiency is defined as four formic acid molecules per glucose.

Growth simulations implemented with the COBRA Toolbox using iND750 agreed with 5 of the 8 reported by Kennedy et al. [5]. The version of iND750 available via the BIGG database [16] does not include reactions annotated with *YAT2*, so we could not reproduced the three lowest-ranked strains reported by Kennedy et al. Yeast 5.01 did not replicate the iND750 predictions, but Yeast 5.30 did.

Although we found broader pathway coverage in the Yeast model than iND750 and found errors in iND750's reconstruction of C-1 metabolism, we found that version 5.01 of the Yeast model did not replicate the prediction of formate-producing strains. Following two

rounds of model curation and updates, model predictions are more similar to simulations with iND750 (Table 5.2). However, simulations using the Yeast model did not predict the observed additional increase in formate flux due to additional deletion of *ALT2* (Table 5.3). This result is not surprising, since the open reading frame which encodes *ALT2*, YDR111C, is annotated as "uncharacterized" in the Saccharomyces Genome Database, and so is not included in version 5.01 of the Yeast model.

Table 5.3 – Model predictions of increasing formate production with additional gene deletions

Genotype	Predicted Efficiency (%)[a]	
	iND750	updated Yeast 5.30
fdh1 fdh2	0.0	0.0
fdh1 fdh2 zwf1	0.0	0.0
fdh1 fdh2 zwf1 fum1	71.8	71.9
fdh1 fdh2 zwf1 fum1 alt2	72.1	71.9

[a]One hundred percent efficiency is defined as four formic acid molecules per glucose.

The iND750 model accurately predicts increasing formate production with the deletion of *ALT2* in addition to *ZWF1* and *FUM1* while the Yeast model does not.

We found that the prediction of formate-producing mutants by the Yeast model was sensitive to constraints on reactions involving NAD(P) cofactors and developed a literature-supported set of constraints to these reactions which improved the performance of the Yeast model (Table 5.4). Additionally, as described in the below, we found that a screen using the Yeast model predicted formate-producing mutants that were not predicted by the iND750 screen.

Table 5.4 – Additional curation improves metrics of Yeast model predictive accuracy

	Yeast 5.30+	Yeast 5.01	Yeast 4
Model description			
Number of metabolites	1602	1655	1481
Number of reactions	2051	2110	2030
Number of genes	904	918	924
Number of dubious genes	1	0	4
Blocked reactions	43%	38%	26%
Viability analysis			
Sensitivity	96%	97%	95%
Specificity	52%	47%	44%
Positive predictive value	87%	86%	85%
Negative predictive value	78%	84%	73%
Geometric mean	50%	46%	42%
Auxotrophy analysis			
Auxotroph-inducing genes included	70	70	73
Correct auxotroph predictions	74%	73%	66%
Incorrectly predicted as viable in minimal media	21%	24%	30%
Incorrectly predicted as inviable in supplemented media	4%	3%	4%

Additional curation, which was conducted to improve the predictions of the formate-producing phenotype, also improved model prediction of gene essentiality and auxotrophy.

iND750 growth simulations with the COBRA Toolbox predict similar formate fluxes to those reported by Kennedy et al.

Since Kennedy et al. implemented their computational screen using different software than our efforts (they reported conducting flux balance analysis with "the GNU MathProg language and solved with custom-generated C code" to interface with the GNU linear programming kit), we began by simulating growth of the 8 predicted formate-producing strains reported in Kennedy et al.'s Table 5.1. The code used for these simulations, along

with other iND750 simulations reported below, is available in Appendix 3. We successfully reproduced the formate production efficiencies for five of the eight candidate high formate producing strains reported by Kennedy et al. when we simulated growth using MATLAB, the COBRA Toolbox, and the Gurobi solver with the iND750 model and the same *in silico* media composition and appropriate constraints to simulate eliminating the formate dehydrogenase reaction as reported by Kennedy et al. (Table 5.2). We could not reproduce the prediction of formate production for the remaining candidate strains because the version of iND750 that was available from the BIGG database [16] when we conducted our simulation in Spring 2012 did not include any reactions annotated with *YAT2*. Thus, we concluded that our simulation using the COBRA Toolbox and Gurobi solver provides results comparable to those produced from the Kennedy et al. implementation. Differences in simulation results may be due to differences in numerical precision and implementation or because the previous effort used a different version of the iND750 model than that currently available for download from the BIGG database. We consider our results to be similar enough to the previously reported results to enable further model and implementation comparison.

An *in silico* screen of triple mutants using the iND750 model with the COBRA toolbox includes the previously reported top-ranked mutant, but also predicts a higher formate-producing mutant

An *in silico* screen of 3898895 possible triple mutants using the iND750 model predicted 217556 triple mutants (5.58%) would produce formate. The screen found that the *alt2 fum1 zwf1* mutant, which was the top-ranked mutant reported from the Kennedy screen, would produce formate at an efficiency of 72.1%. This result is similar to the predicted 72.3% efficiency reported by Kennedy et al. However, our screen also predicted that a *put1 fum1 zwf1* mutant would produce formate at a higher 72.2% efficiency (Table 5.5).

Table 5.5 - Strains predicted to be high formate-producing mutants in screen using iND750

Rank	Genotype[a]	Predicted Efficiency (%)[b]	Kennedy et al. Rank
1	*put1 zwf1 fum1*	72.24	
2	*alt2 zwf1 fum1*	72.13	1
3	*car1 zwf1 fum1*	72.12	
4	*car2 zwf1 fum1*	72.05	
5	*ach1 zwf1 fum1*	72.01	
6	*aat2 zwf1 fum1*	71.98	2
7	*sfc1 zwf1 fum1*	71.96	
8	*dic1 zwf1 fum1*	71.92	
9	*aah1 zwf1 fum1*	71.90	
10	*urk1 zwf1 fum1*	71.89	

[a]Using a *fdh1 fdh2* background.

[b]One hundred percent efficiency is defined as four formic acid molecules per glucose.

The genotype and predicted formate-production efficiency of the top ten predicted formate-producing strains found by our *in silico* screen of 3,898,895 triple mutants, using the iND750 model of yeast metabolism. Our screen results included the top 5-ranked strains reported by Kennedy et al. [5]

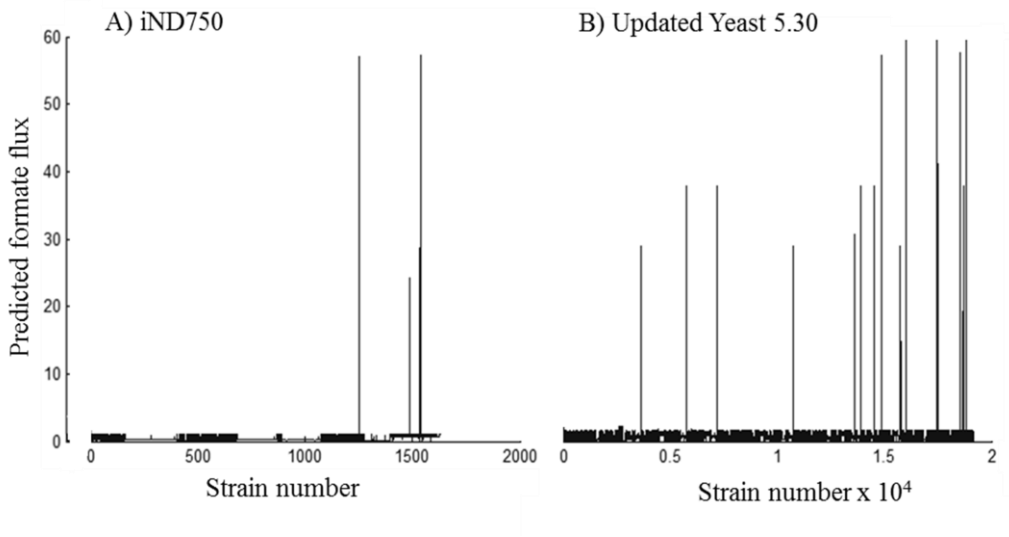
Simulations using the iND750 model predict the increasing formate produced by triple mutants over double and single mutants

When constructing the *alt2 zwf1 fum1 fdh1 fdh2* strain, Kennedy et al. also reported that total extracellular formic acid production increased from wild type to *zwf1* to *zwf1 fum1* to *zwf1 fum1 alt2*strains. Growth simulations with the iND750 model using these knockouts did not predict formate production in the *zwf1* mutant (as was observed *in vivo* by Kennedy et al.), but did predict formate production in *zwf1 fum1* and predicted increased formate production in the *zwf1 fum1 alt2* mutant (Table 5.3).

An *in silico* screen of double mutants using the iND750 model predicts three high formate-producing mutants

Since the additional formate flux predicted by the addition of a third gene deletion was relatively small, we also conducted an *in silico* screen of double mutants in a *fdh1 fdh2* background for formate-producing strains. When using the iND750 model, the screen found many mutants predicted to produce formate at low concentrations, and 3 high formate-producing yeast double mutants: *fum1 rpe1*, *fum1 fbp1*, and *fum1 zwf1* (Figure 5.2A). These mutants were predicted to produce formate with efficiencies of 71.45%, 30.38%, and 71.85%, respectively. We note that *RPE1* and *ZWF1* encode proteins in the pentose phosphate pathway, and that *FBP1*encodes Fructose-1,6-bisphosphatase, a key enzyme in the gluconeogenic pathway, which would not be expected to have a flux in the glucose-fed aerobic environment of this simulation.

Figure 5.2 - Predicted formate production in double knockout screen



Results of *in silico* screening of 41,041 double mutants in a *fdh1 fdh2* background for formate-producing strains using the iND750 model (A), and 61,776 double mutants using the updated Yeast 5.30 model (B). Simulations with the iND750 model predicted 1,635 formate-producing double mutants (3.9%), with 3 high formate-producing yeast double mutants: *fum1 rpe1*, *fum1 fbp1*, and *fum1 zwf1*. Simulations with the updated Yeast 5.30 model predicted 19,197 formate-producing mutants (31.1%), with 15 high formate-producing double mutants, which are detailed in Table 5.6.

Table 5.6 - Double deletion mutants predicted to produce formate using the updated Yeast 5.30 model

	Rank	Genotype[a]	Predicted Efficiency (%)[b]	Rank in Kennedy screen	Notes on gene deleted in addition to *FUM1*
Glyoxylate Cycle	1	*fum1 dic1*	74.39		Mitochondrial dicarboxylate carrier
	1	*fum1 mdh2*	74.39		Cytoplasmic malate dehydrogenase
TCA cycle	2	*fum1 mdh1*	74.37		Mitochondrial malate dehydrogenase
Pentose Phosphate Pathway	3	*fum1 zwf1*	71.94	1, 2, 3	Glucose-6-phosphate dehydrogenase, catalyzes the first step of the pentose phosphate pathway
	4	*fum1 rpe1*	71.50	4	D-ribulose-5-phosphate 3-epimerase, catalyzes a reaction in the non-oxidative part of the pentose-phosphate pathway
	5	*fum1 tal1*	51.53		Transaldolase, enzyme in the non-oxidative pentose phosphate pathway
Lysine Biosynthesis	6	*fum1 lys20*	47.40		Homocitrate synthase isozyme, catalyzes the first step in the lysine biosynthesis pathway
	6	*fum1 lys4*	47.40		Homoaconitase, catalyzes a step in the lysine biosynthesis pathway
	6	*fum1 lys12*	47.40		Homo-isocitrate dehydrogenase, an NAD-linked mitochondrial enzyme required for the fourth step in the biosynthesis of lysine
	6	*fum1 lys1*	47.40		Saccharopine dehydrogenase, catalyzes the final step in the lysine biosynthesis pathway
	6	*fum1 lys9*	47.40		Saccharopine dehydrogenase , catalyzes the seventh step in lysine biosynthesis pathway
Other amino acid metabolism	7	*fum1 bat1*	38.39		Mitochondrial branched-chain amino acid aminotransferase
	8	*fum1 bat2*	36.27		Cytosolic branched-chain amino acid aminotransferase
	9	*fum1 leu2*	36.05		Beta-isopropylmalate dehydrogenase, catalyzes the third step in the leucine biosynthesis pathway
	9	*fum1 leu1*	36.05		Isopropylmalate isomerase, catalyzes the second step in the leucine biosynthesis pathway

[a]Genotype in addition to *fdh1 fdh2* background required to allow formate production

[b]One hundred percent efficiency is defined as four formic acid molecules per glucose

An *in silico* screen of double deletion yeast mutants in a *fdh1 fdh2* background using the Yeast metabolic model predicted 15 high formate-producing yeast mutants. A similar screen using iND750 predicted 3 high formate-producing yeast mutants (*fum1 rpe1, fum1 fbp1,* and *fum1 zwf1*, with efficiencies of 71.45%, 30.38%, and 71.85%, respectively).

The Yeast 5 model must be modified to match *in vivo* observations and iND750 predictions.

We found that growth simulations using version 5.01 of the Yeast model did not predict formate production for any of the 8 strains reported by Kennedy et al. (Table 5.2). The code used for these simulations is available in Appendix 3. Since formate production was observed in the *alt2 zwf1 fum1* strain, this prediction is incorrect. The simulations predicted that biomass could be formed in 6 of the 8 strains, correctly predicting the auxotrophy caused by *CHO1*deletion.

To improve the Yeast model's prediction of the observed formate producing phenotype, we manually reviewed the reconstruction of C-1 metabolism by comparing it to literature describing this portion of metabolism [17–19]. During this review, we generated a list of literature-supported modifications to the Yeast model, which are documented with references in Appendix 3. These changes have been incorporated (along with changes suggested by Jouhten et al. [20]) to make version 5.30 of the Yeast model.

Growth simulations using Yeast version 5.30 predict formate production for the same 5 strains as our simulation using iND750. Additionally, Yeast version 5.30 predicts that the *CHO1* mutant will not grow in media that does not contain choline. However, as discussed below, we found that additional constraints are needed to improve the performance of a double knockout screen using Yeast 5.30. These changes, along with simulations using Yeast version 5.30, are documented in Appendix 3.

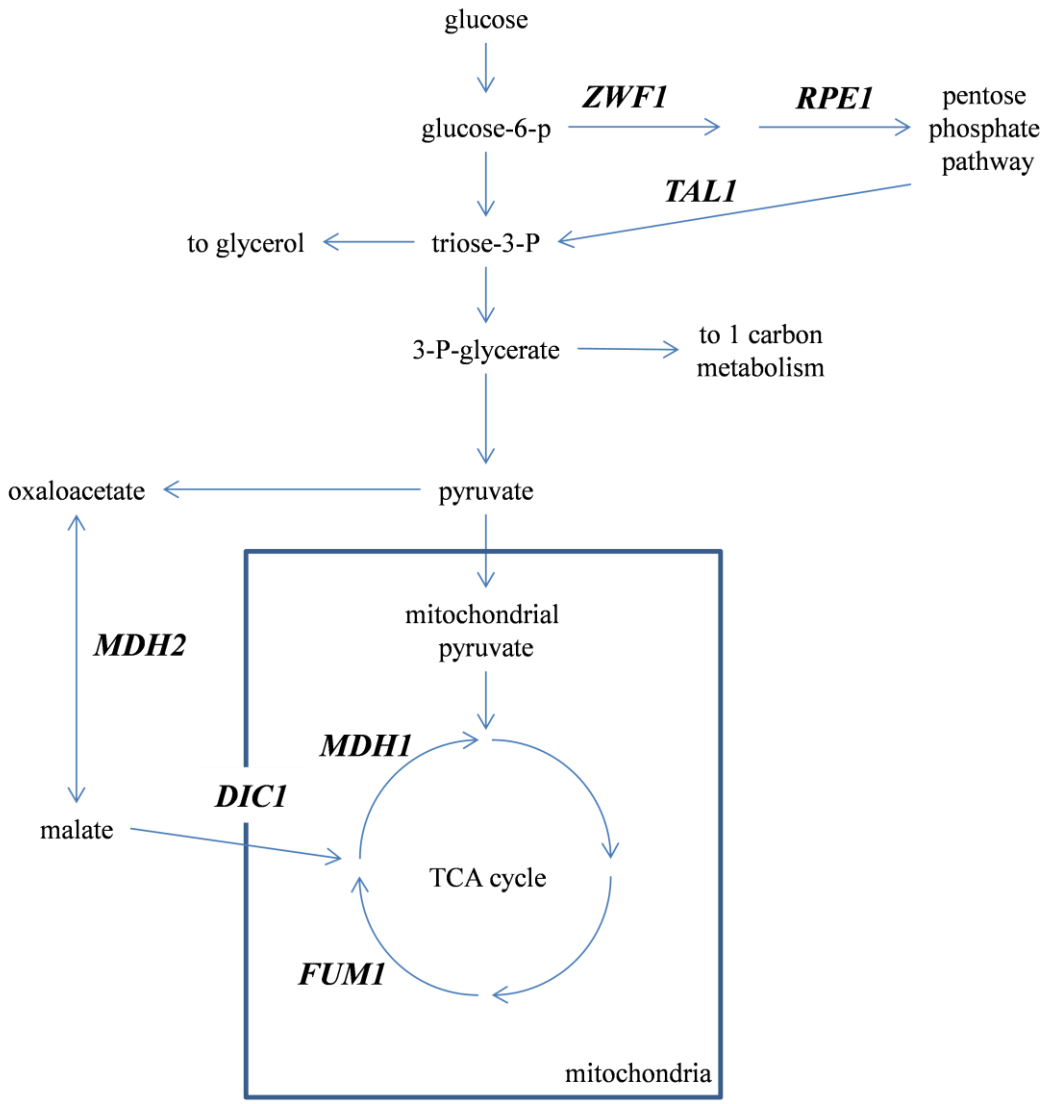
In silico double mutant screening results differ between the Yeast 5 model and the iND750 model, and are sensitive to constraints on reactions involving NAD(P) cofactors

As we did using the iND750 model, we conducted an *in silico* screen of double mutants in a *fdh1 fdh2* background for formate-producing strains using the Yeast 5.30 model. While the iND750 screen predicted 3 high formate-producing mutants, the Yeast 5.30 screen predicted thousands. When reviewing this preliminary result, we observed that the simulations predicted that all *FUM1* mutants produced formate. However, we also found that if we relaxed constraints on all cytoplasmic reactions involving NAD(P), none of these mutants were predicted to produce formate. Thus, we reviewed the literature for reactions involving these cofactors for which biochemical evidence of *in vivo* reversibility has been reported. Since efforts to incorporate thermodynamic information with genome-scale metabolic models are ongoing [4], we chose to restrict the directionality of other reactions involving these cofactors, a choice which may make the model predictions less robust in differing redox conditions. The set of reversible reactions involving NAD that we used for further simulation consists of the model reactions named (R,R)-butanediol dehydrogenase, aldehyde dehydrogenase (2-methylbutanol, NAD), aldehyde dehydrogenase (2-phenylethanol, NAD), aldehyde dehydrogenase (isoamyl alcohol, NAD), aldehyde dehydrogenase (isobutyl alcohol, NAD), aldehyde dehydrogenase (tryptophol, NAD), formaldehyde dehydrogenase, glutamate dehydrogenase (NAD), glyceraldehyde-3-phosphate dehydrogenase, malate dehydrogenase, cytoplasmic, saccharopine dehydrogenase (NAD, L-lysine forming), and NAD transport (reaction ids r_0003, r_0166, r_0169, r_0179, r_0182, r_0186, r_0443, r_0470, r_0486, r_0714, r_0988, and r_1956). The set of reversible reactions involving NADP consists of model reactions named D-arabinose 1-dehydrogenase (NADP), glutamate

dehydrogenase (NADP), methylenetetrahydrofolate dehydrogenase (NADP), saccharopine dehydrogenase (NADP, L-glutamate forming), and NADP(+) transport (reaction ids r_0321, r_0471, r_0732, r_0989, and r_1963).

During the review of cytosolic reactions involving NAD(P), we also found literature evidence for further updates to the Yeast model, which are documented with references in Appendix 3. We incorporated these updates and constraints for an Updated Yeast 5.30. This updated model predicted 15 high formate-producing strains in an *in silico* screen of double knockout mutants for (Figure 5.2B). The highest-ranked predicted formate-producing mutants included knockouts in the TCA and glyoxylate cycles, as well as the pentose phosphate pathway. Additional formate-producers were predicted from deletions in amino acid biosynthetic metabolism (Table 5.6). It remains unclear whether the simulated formate producing strains with knockouts in amino acid metabolism reflect accurate predictions of a possible link between nitrogen metabolism and C-1 metabolism or are an artifact of the complexity of reconstructing this portion of yeast biochemistry. Thus, amino acid and nitrogen metabolism is a good target for futher model curation. However, the screen of *in silico* double deletion mutants using the updated Yeast model suggests deletions in pathways which are the same branchpoints in carbon metabolism as those suggested using the iND750 model (Figure 5.3).

Figure 5.3 – Top-ranked predicted formate producing knockout target genes' relationship to central carbon metabolism



The top-ranked predicted formate-producing mutants in an *in silico* screen using the updated Yeast 5 model are involved in the glyoxylate cycle, the TCA cycle, and the pentose phosphate pathway, pathways common to genes identified in a screen using the iND750 model.

Thus, performing a screen of simulated double-knockout mutants found that our literature-supported curation efforts affected the prediction of formate producing strains. As in our previous work, we found that this approach to curation also improved viability and auxotrophy predictions (Table 5.4), metrics which have

previously been used to evaluate the performance of such models [21]. We note that growth simulations using either iND750 or the updated Yeast model do not predict the formation of glycerol, which has been suggested as a mechanism for balancing excess NADH production in *S. cerevisiae* during growth on glucose [22]. This observation supports the assertion that additional efforts are required to accurately model yeast redox metabolism [4].

Discussion

Through our work to reimplement the computational methods of Kennedy et al [5] with the iND750 model [6] and to apply them to the Yeast [13] model, we have documented errors in both models, which we describe below, and have corrected and documented using the scripts included in Appendix 3. We have further demonstrated the benefits of literature-based curation of metabolic reconstructions as a tool for model improvement, and found that a comparative approach in which different model predictions are analyzed can give further insights to the use and improvement of these and other stoichiometrically constrained metabolic models. Our work also demonstrates that constraint-based stoichiometric models are sensitive to appropriate constraints on reactions involving redox cofactors when conducting flux balance analysis. Discussion of this finding is expanded below in the section entitled "Screening results for double and triple mutants for formate-producing strains differs between models, reflects model sensitivity to redox reaction constraints".

Results of flux balance analysis of previously predicted formate-producing mutant strains differs between models, reveals model errors

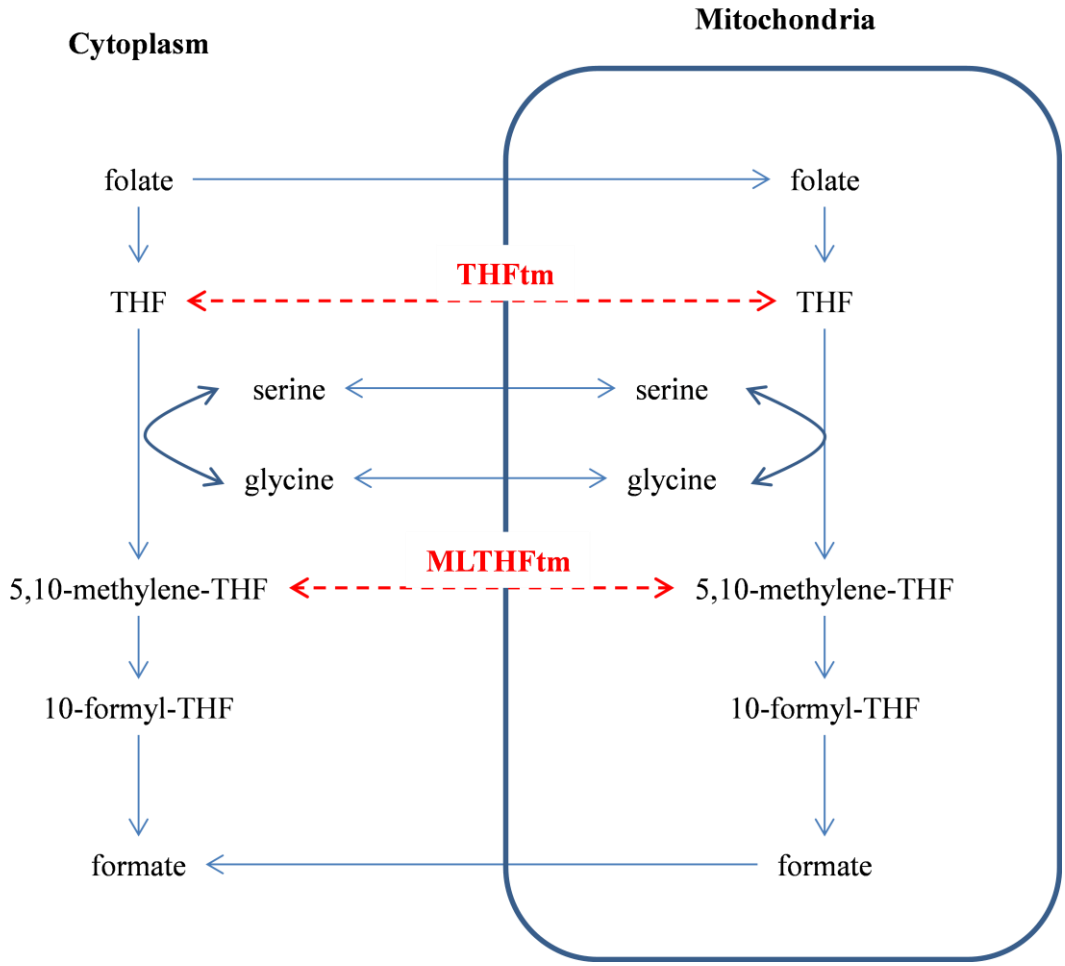
Since the version of the iND750 model that was available from the BIGG database [16] when we conducted our simulation in Spring 2012 did not include any reactions annotated with *YAT2*, we were unable to simulate growth of the *yat2* yeast

strains reported by Kennedy et al. Further, our simulations using the updated Yeast model (which does include reactions annotated with *YAT2*) did not predict formate production for these strains, and predicted that the 7th-ranked strain (a *cho1*mutant) would not be viable in choline-free media (Table 5.1). *CHO1*mutants have been reported to be choline auxotrophs [23].

For the five wild type *YAT2* strains which were previously reported as predicting formate production, our simulations using both the iND750 and the updated Yeast model predicted different formate production efficiencies than reported by Kennedy et al. That the differences in predicted formate yield we found using the iND750 model were not large suggests that our imposition of constraints to simulate the media formulation and strain background was similar to the approach used in the previous work. It is likely that the differing simulation results are a result of using different linear programming software and FBA implementation than the previous effort, not a result of any significant changes to model constraints or other possible revisions to the underlying model.

Flux balance analysis of the predicted fluxes reported for the top-performing *alt2 fum1 zwf1* strain using iND750 includes solutions which predict a large flux through reactions transporting methylene-tetrahydrofolate and tetrahydrofolate between the mitochondria and the cytosol (Figure 5.4). These reactions, called MLTHFtm and THFtm in the iND750 model are not supported by the literature on compartmentalization of C1-metabolism [17–19], suggesting that these reactions were erroneously included in iND750 (they are not present in the Yeast model). Fortunately, we found that removing these reactions from iND750 did not affect the prediction of formate production for this strain, though the model modification did affect the simulated internal flux distribution.

Figure 5.4 – Errors in the iND750 reconstruction of formate metabolism



iND750 includes two transport reactions, MLTHFtm and THFtm, which are not supported by the literature on compartmentalization of C1-metabolism [17–19]. These reactions are not present in the Yeast model. The figure is adapted from Appling [17].

Additionally, we noted that simulations with iND750 predicted unusual fluxes which we cannot rule out in the absence of further experimental evidence, but seem unlikely based on previous efforts to document yeast metabolic fluxes. The 5th-ranked strain reported by Kennedy et al. includes a mutation in *FBP1*, which codes for fructose-1,6-bisphosphatase [24],a key enzyme involved in gluconeogenesis, a pathway which would not be expected to carry a flux in a glucose-rich media [25]. When we simulated growth of a *fum1* strain using iND750, the simulation predicted a

flux through the reaction catalyzed by *FBP1* which increased the flux to the pentose phosphate pathway – an observation similar to the recently reported finding that gluconeogenesis is required to redirect flux to the oxidative pentose phosphate pathway in yeast engineered for xylose-metabolism [26]. However, since fructose-1,6-bisphosphatase is transcriptionally repressed in the presence of hexoses [27], this flux distribution seems unlikely in glucose-fed yeast mutants, and does not agree with previously reported fluxes through central carbon metabolism for *fum1* mutants [25]. Simulations of the *fum1* mutant using the updated Yeast model did not predict this gluconeogenic flux.

We found that simulations using the updated Yeast 5 model predicted that the *fum1 zwf1* mutant produced formate, but did not predict additional formate production due to deletion of *ALT2, AAT2,* or *CAT2* (the top three-ranked strains reported from the Kennedy et al. screen). The Yeast model prediction of no increased flux with the additional deletion of *ALT2* differs from Kennedy et al.'s *in vivo* measurement of higher formate production in a *fum1 zwf1 alt2* mutant than a *fum1 zwf1* mutant.

Our simulations of a *fum1 zwf1*mutant using iND750 predicted fluxes through reactions annotated with *ALT2* and *AAT2,* but not *CAT2*. We investigated each of these flux predictions, finding fluxes through reactions which are not present in the Yeast model and fluxes through reactions which are mis-annotated in iND750. We also found that the predicted formate flux in the simulated *fum1 zwf1 cat2* strain was the same as a simulated *fum1 zwf1* strain.

Although the reaction catalyzed by *ALT2* in iND750 is supported by literature evidence [28], *ALT2* is annotated as an "uncharacterized ORF" by the Saccharomyces Genome Database (SGD) [29], and is therefore not included in the Yeast 5 model. Further, simulation with iND750 predicts an upstream flux through a reaction catalyzed by a second uncharacterized ORF, YML082W. We found that adding these

two reactions to the updated Yeast model was not sufficient to produce a flux through the reaction catalyzed by *ALT2*, so simulations using a modified Yeast 5 model still did not predict an increase in formate production with this third gene deletion. It is not clear which model features of iND750 lead to the flux through the reaction annotated by *ALT2*.

Our simulations of a *fum1 zwf1*mutant using iND750 predicted fluxes through two reactions annotated with *AAT2*: an aspartate transaminase reaction and a tyrosine transaminase reaction. We found that the aspartate transaminase reaction is present in the Yeast model, but did not have a flux in simulations using that model. Further, we found that the tyrosine transaminase reaction corresponded to the reaction catalyzed by *ARO8*, not *AAT2*. This annotation appears to be an annotation error in iND750. This reaction is correctly annotated in the Yeast model, which may contribute to simulations using that model not predicting a different flux for *aat2 fum1 zwf1* than for *fum1 zwf1*.

Screening results for double and triple mutants for formate-producing strains differs between models, reflects model sensitivity to redox reaction constraints

Having successfully simulated growth and the formate-producing phenotype for the predicted top-performing formate-producing strains reported by Kennedy et al., we re-implemented their *in silico* screen of triple-knockout yeast mutants for a formate-producing phenotype using iND750. Our screen implementation identified the first five strains reported from Kennedy et al.'s screen, though only Kennedy's first and second ranked mutants were among the 10 top-ranked strains in our screen (Table 5.5). The top-ranked predicted formate-producing yeast strain in our screen was a *put1 zwf1 fum1* mutant, and Kennedy et al.'s top-ranked *alt2 zwf1 fum1* strain was the

second highest ranked in our screen. Like *ALT2*, *PUT1*is implicated in glutamate metabolism in the iND750 model. Additionally, both genes are affected by redox metabolism in this model: *ALT2* catalyzes a reaction using NAD as a cofactor, and *PUT2* catalyzes a reaction using glutamate as a substrate to produce 2-oxoglutarate. In turn, 2-oxoglutarate produces glutamate in a reaction using NADPH as a cofactor. Thus, our top-ranked screen predictions are similar to those reported by Kennedy et al., though not identical. As with our efforts to reproduce their simulation of top-ranked screens, differences are likely due to software implementation details or modifications to the iND750 model that are not incorporated to the version we downloaded from the BIGG database. Like the Kennedy et al. screening results, the computationally suggested third knockouts were at distantly located positions within the metabolic network, not directly implicated in formic acid biosynthesis or C1 metabolism.

Of the top 259 predicted formate-producing strains using the iND750 model, 258 were *zwf1 fum1*mutants (the other strain also had mutants in the pentose phosphate pathway and the TCA cycle – a *rpe1 put1 fum1* genotype). In all cases, the deletion of a third gene in addition to *FUM1* and *ZWF1* (or *RPE1*) contributed a smaller increase in predicted formate flux than the deletion of a second gene in addition to *fum1*. Therefore, we also conducted *in silico* screens of double-mutants for formate producing strains using both the iND750 and the Yeast 5 models. As discussed in the Results section, our initial screen with the Yeast 5.30 model predicted that formate producing mutants were very common – in fact, all *fum1*mutants were predicted to be high formate-producers, and no flux was predicted through the oxidative pentose phosphate pathway (and so there was no additional formate production predicted when *ZWF1*was deleted). However, when we relaxed the constraints to allow all cytosolic reactions involving NAD or NADP to be reversible

(an approach suggested to avoid over-constraining the model in the absence of thermodynamic information [4]), none of these mutants were predicted to produce formate. When relaxing constraints on only the NAD reactions, all *fum1*mutants were predicted to produce formate, demonstrating that this model is sensitive to constraints on redox reactions, and suggesting that constraints on reactions involving NAD can be tuned to affect model prediction of a flux through the pentose phosphate pathway, and C-1 metabolism (which leads to formate production). In the absence of tools to thermodynamically constrain these reactions, we chose to relax the reversibility constraints only on those reactions for which we could find *in vivo* observations of reversibility reported in the literature.

We are not the first to report observations of computationally-predicted interdependencies between redox reactions, the pentose phosphate pathway, and C-1 metabolism. Such dependencies have also been noted in efforts to improve flux balance predictions using a modified version of the iND750 model [30]. Additionally, we note that these fluxes would be expected to be dependent upon media formulation: a variable flux through oxidative pentose phosphate pathway has been documented as a result of supplementing growth media with amino acids [31].

After applying our literature-supported constraints to cytosolic reactions involving NAD and NADP to the Yeast 5 model, the results of our double knockout mutant screens differed between the iND750 and updated Yeast 5 models (Figure 5.2). A screen for formate-producing double mutants using the iND750 model predicted 3.9% of the double mutants would produce formate, with three strains predicted to be unusually high formate producers: *fum1 zwf1*, *fum1 rpe1*, and *fum1 fbp1*. The screen conducted with Yeast 5 predicted 31.1% of the double mutants would produce formate, with 15 predicted high-performing double mutants (Table 5.6). The screen using Yeast 5 did not predict that the *fum1 fbp1*double mutant would be a high

formate-producing strain, and predicted high-performing strains with mutations in the TCA cycle, the pentose phosphate cycle, and some amino acid biosynthetic pathways (Figure 5.3). Testing these predictions provides an opportunity for further model validation and improvement.

A variety of techniques have been proposed to address perceived shortcomings of flux balance analysis predictions in eukaryotes. However, advanced computational techniques may be prematurely applied when the underlying reconstruction of biochemistry is incomplete or inaccurate. In the current work, we have not applied any techniques beyond standard flux balance analysis and literature review to improve the reconstruction of established yeast biochemical knowledge. We have found that problematic predictions can be corrected through literature-based curation of biochemical reconstruction, and do not necessarily imply limitations of the computational methods used. The process of reconstructing yeast metabolism is not complete, and we expect the iterative process of community model curation will continue to improve the Yeast model.

Conclusions

Our work emphasizes need for ongoing model validation efforts, which benefit from application of genome-scale metabolic models, and emphasizes a need for improved documentation of computational models, their modification, and the details of their use. Through comparative analysis, we found that flux balance analysis of the iND750 model predicted fluxes through reactions annotated with uncharacterized open reading frames that are not present in the consensus reconstruction. Kennedy et al.'s observation of increased formate production in a *fum1 zwf1 alt2* mutant relative to a *fum1 zwf1* mutant is particularly remarkable given the annotation and reconstruction errors we found in the iND750 model and the absence of this prediction in the newer Yeast model (which does not include reactions annotated with the *ALT2* gene).

Observing the differing predictions of a phenotype for this "uncharacterized ORF" required comparative analysis of metabolic models which differ in scope (such as iND750 and the Yeast model). Screening models for fluxes through hypothesized reactions or annotations is a promising technique for guiding future efforts to describe the metabolic role of other poorly characterized yeast genes. Thus, the inclusion of such hypothetical annotation can reveal novel phenotypes. However, the inclusion of *ALT2* is not documented as hypothetical in the model, which raises the question: which portions of this model are established biochemical fact, and which portions are not? Thus, we conclude that the research community would benefit from improved annotation of metabolic models.

We were unable to reproduce the prediction that deletion of *YAT2* affected formate metabolism using the iND750 model because that gene was not present in the version of iND750 that we used for our efforts. However, we do not know the full details of how the version of iND750 we used might differ from that used by Kennedy et al. Thus, we conclude that research which applies or modifies constraint-based models would also benefit from careful and extensive documentation including sharing the model used, along with software code used to modify models and to implement simulations.

We have found that the simulation results are sensitive to details of implementation and to applied constraints. Specifically, we found that we could not precisely match previously reported *in silico* screen results with a new implementation, and we found that we could tune the frequency of formate-producing double mutants by constraining the reversibility of reactions involving NAD(P) cofactors in the Yeast model. Since different results can be found with different constraints, the set of constraints used for a screen must also be well documented. If we hope to discriminate factors which are numerically significant and those which are

biologically significant, we must continue to improve the approach to modeling and reporting computational methods and results.

Phenotype predictive power remains the standard for comparing different models of yeast metabolism, but our work demonstrates that in some cases, a correct phenotypic prediction does not necessarily imply a more accurate reconstruction of a metabolic network, and that an improved biochemical reconstruction does not guarantee improved predictive power. Thus, model flux predictions merit careful review in all cases. Such review provides a useful tool for evaluating and improving yeast models – the iterative process of building an accurate model of yeast metabolism continues.

We found that simulations using the most recent model of yeast metabolism predict more formate producing double knockout mutants than simulations using the iND750 model, and that the simulations with the Yeast model do not predict a gluconeogenic flux which is predicted by iND750. Additional *in vivo* data is required to evaluate these predictions.

Methods

All simulations were conducted using version 2.0.4 of the COBRA Toolbox [3] for MATLAB version 2011b on PCs running Windows 7. We used version 4.61 of the Gurobi Optimizer to solve the linear programming problem underlying flux balance analysis. The MATLAB code we wrote for all simulations (which includes references for model modifications as comments in the source code) is available in Appendix 3.

REFERENCES

1. Österlund T, Nookaew I, Nielsen J: **Fifteen years of large scale metabolic modeling of yeast: Developments and impacts**. *Biotechnology Advances* 2011.

2. Milne CB, Kim P-J, Eddy JA, Price ND: **Accomplishments in genome-scale in silico modeling for industrial and medical biotechnology**. *Biotechnol J* 2009, **4**:1653–1670.

3. Schellenberger J, Que R, Fleming RMT, Thiele I, Orth JD, Feist AM, Zielinski DC, Bordbar A, Lewis NE, Rahmanian S, Kang J, Hyduke DR, Palsson BØ: **Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0**. *Nat Protoc* 2011, **6**:1290–1307.

4. Soh KC, Miskovic L, Hatzimanikatis V: **From network models to network responses: integration of thermodynamic and kinetic properties of yeast genome-scale metabolic networks**. *FEMS Yeast Research* 2012, **12**:129–143.

5. Kennedy CJ, Boyle PM, Waks Z, Silver PA: **Systems-Level Engineering of Nonfermentative Metabolism in Yeast**. *Genetics* 2009, **183**:385–397.

6. Duarte NC, Herrgård MJ, Palsson BO: **Reconstruction and validation of Saccharomyces cerevisiae iND750, a fully compartmentalized genome-scale metabolic model**. *Genome Research* 2004, **14**:1298.

7. Kuepfer L, Sauer U, Blank LM: **Metabolic functions of duplicate genes in Saccharomyces cerevisiae**. *Genome Res* 2005, **15**:1421–1430.

8. Herrgård MJ: **Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in Saccharomyces cerevisiae**. *Genome Research* 2006, **16**:627–635.

9. Hjersted JL, Henson MA: **Steady-state and dynamic flux balance analysis of ethanol production by Saccharomyces cerevisiae**. *IET Syst. Biol.* 2009, **3**:167.

10. Nookaew I, Jewett MC, Meechai A, Thammarongtham C, Laoteng K, Cheevadhanarak S, Nielsen J, Bhumiratana S: **The genome-scale metabolic model iIN 800 of Saccharomyces cerevisiae and its validation: a scaffold to query lipid metabolism**. *BMC Syst Biol* 2008, **2**:71.

11. Herrgard MJ, Swainston N, Dobson P, Dunn WB, Arga KY, Arvas M, Buthgen N, Borger S, Costenoble R, Heinemann M, Hucka M, Le Novere N, Li P, Liebermeister W, Mo ML, Oliveira AP, Petranovic D, Pettifer S, Simeonidis E, Smallbone K, Spasie I, Weichart D, Brent R, Broomhead DS, Westerhoff HV, Kurdar B, Penttila M, Klipp E, Palsson BO, Sauer U, Oliver SG, Mendes P, Nielsen J, Kell DB: **A consensus**

yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nat Biotech* 2008, **26**:1155–1160.

12. Dobson PD, Jameson D, Simeonidis E, Lanthaler K, Pir P, Lu C, Swainston N, Dunn WB, Fisher P, Hull D, Brown M, Oshota O, Stanford NJ, Kell DB, King RD, Oliver SG, Stevens RD, Mendes P: **Further developments towards a genome-scale metabolic model of yeast**. *BMC Syst Biol* 2010, **4**:145.

13. Heavner BD, Smallbone K, Barker B, Mendes P, Walker LP: **Yeast 5 – an expanded reconstruction of the Saccharomyces Cerevisiae metabolic network**. *BMC Systems Biology* 2012, **In press**.

14. Mo ML, Palsson BØ, Herrgård MJ: **Connecting extracellular metabolomic measurements to intracellular flux states in yeast**. *BMC Syst Biol* 2009, **3**:37.

15. Zomorrodi AR, Maranas CD: **Improving the iMM 904 S. cerevisiae metabolic model using essentiality and synthetic lethality data**. *BMC Syst Biol* 2010, **4**:178.

16. Schellenberger J, Park J, Conrad T, Palsson B: **BiGG: a Biochemical Genetic and Genomic knowledgebase of large scale metabolic reconstructions**. *Bmc Bioinformatics* 2010, **11**:213.

17. Appling DR: **Compartmentation of folate-mediated one-carbon metabolism in eukaryotes**. *FASEB J.* 1991, **5**:2645–2651.

18. Piper MD, Hong S-P, Ball GE, Dawes IW: **Regulation of the Balance of One-carbon Metabolism in Saccharomyces cerevisiae**. *Journal of Biological Chemistry* 2000, **275**:30987–30995.

19. West MG, Horne DW, Appling DR: **Metabolic role of cytoplasmic isozymes of 5,10-methylenetetrahydrofolate dehydrogenase in Saccharomyces cerevisiae**. *Biochemistry* 1996, **35**:3122–3132.

20. Jouhten P, Rintala E, Huuskonen A, Tamminen A, Toivari M, Wiebe M, Ruohonen L, Penttilä M, Maaheimo H: **Oxygen dependence of metabolic fluxes and energy generation of Saccharomyces cerevisiae CEN.PK113-1A**. *BMC Systems Biology* 2008, **2**:60.

21. Heavner BD, Smallbone K, Barker B, Mendes P, Walker LP: **Yeast 5 - an expanded reconstruction of the Saccharomyces Cerevisiae metabolic network**. *BMC Systems Biology* 2012, **6**:55.

22. Blank LM, Ebert BE, Buehler K, Bühler B: **Redox Biocatalysis and Metabolism: Molecular Mechanisms and Metabolic Network Analysis**. *Antioxidants & Redox Signaling* 2010, **13**:349–394.

23. Atkinson KD, Jensen B, Kolat AI, Storm EM, Henry SA, Fogel S: **Yeast mutants**

auxotrophic for choline or ethanolamine. *J. Bacteriol.* 1980, **141**:558–564.

24. Entian KD, Vogel RF, Rose M, Hofmann L, Mecke D: **Isolation and primary structure of the gene encoding fructose-1,6-bisphosphatase from Saccharomyces cerevisiae**. *FEBS Lett.* 1988, **236**:195–200.

25. Blank LM, Kuepfer L, Sauer U: **Large-scale 13C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast**. *Genome biology* 2005, **6**:R49.

26. Hector RE, Mertens JA, Bowman MJ, Nichols NN, Cotta MA, Hughes SR: **Saccharomyces cerevisiae engineered for xylose metabolism requires gluconeogenesis and the oxidative branch of the pentose phosphate pathway for aerobic xylose assimilation**. *Yeast* 2011, **28**:645–660.

27. Yin Z, Smith RJ, Brown AJ: **Multiple signalling pathways trigger the exquisite sensitivity of yeast gluconeogenic mRNAs to glucose**. *Mol. Microbiol.* 1996, **20**:751–764.

28. García-Campusano F, Anaya V-H, Robledo-Arratia L, Quezada H, Hernández H, Riego L, González A: **ALT1-encoded alanine aminotransferase plays a central role in the metabolism of alanine in Saccharomyces cerevisiae**. *Canadian Journal of Microbiology* 2009, **55**:368–374.

29. Engel SR, Balakrishnan R, Binkley G, Christie KR, Costanzo MC, Dwight SS, Fisk DG, Hirschman JE, Hitz BC, Hong EL, Krieger CJ, Livstone MS, Miyasato SR, Nash R, Oughtred R, Park J, Skrzypek MS, Weng S, Wong ED, Dolinski K, Botstein D, Cherry JM: **Saccharomyces Genome Database provides mutant phenotype data**. *Nucleic Acids Res* 2010, **38**:D433–436.

30. van Berlo RJP, de Ridder D, Daran J-M, Daran-Lapujade PAS, Teusink B, Reinders MJT: **Predicting Metabolic Fluxes Using Gene Expression Differences As Constraints**. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 2011, **8**:206–216.

31. Clasquin MF, Melamud E, Singer A, Gooding JR, Xu X, Dong A, Cui H, Campagna SR, Savchenko A, Yakunin AF, Rabinowitz JD, Caudy AA: **Riboneogenesis in Yeast**. *Cell* 2011, **145**:969–980.

CHAPTER 6 – CONCLUSIONS AND FUTURE RESEARCH

The original objective of this research was to apply a stoichiometrically constrained model of yeast metabolism to develop a metabolic engineering strategy which would be applied to improve the stress tolerance of industrially useful organisms. However, during model validation, I found that existing models do not fully and accurately reconstruct established knowledge of the biochemistry of the yeast metabolic network, and so I expanded the research objective to:

1. Conduct literature-based validation of the reconstruction of sphingolipid metabolism included in the Consensus Reconstruction of Yeast Metabolism
2. Update the existing model by incorporate any changes required to correct errors in the reconstruction
3. Evaluate the impact of these changes on the predictive accuracy of the model
4. Apply the updated model to develop metabolic engineering strategies for strain improvement
5. Evaluate any differences in strategies suggested by the updated model with predictions of previously published models

I found that such models can be improved by careful review of pertinent literature, incorporation of corrections, and comparative simulation efforts. I found that yeast metabolic models may be applied to inform metabolic engineering strategies, but that such models are not yet sufficiently accurate or predictive to execute such strategies without further expert review and analysis. Through the process of improving yeast models, I have demonstrated that such models can be used to generate new hypotheses, and are therefore a valuable tool for advancing

our understanding of metabolism.

Manual review of the biochemical reactions included in the Consensus Reconstruction of Yeast Metabolism [1, 2] found that this model did not accurately reconstruct the metabolic pathways involved in the biosynthesis of complex sphingolipids (Chapter 3, [3]). The model was both incomplete, containing only 28 of 243 biochemical reactions I found in the pathway, and inaccurate – it contained an additional 13 reactions which were not supported by the literature. I found that that correcting these and other errors in the biochemistry improved the model: compared to version 4 of the Yeast Consensus Reconstruction, Yeast 5better predicts gene essentiality,knockout-induced auxotrophies, and other phenotypes (Chapter 4, [4]). I made particular efforts to document my approach to improving the model and conducting simulations by including the MATLAB scripts I used as supplemental information to published manuscripts, and as Appendices to this dissertation.

Next, I worked to return to my original objective of applying a stoichiometrically constrained model of yeast metabolism to develop metabolic engineering strategies. I wrote software to re-implement a previously reported *in silico* screen of simulated mutants for a formate-producing phenotype, and compared the predictions of the updated model to the older iND750 model (Chapter 5). I found that Yeast 5 did not accurately predict phenotypes which had been predicted by iND750 and observed *in vivo.* By comparing predicted fluxes between the models, and conducting additional literature-based curation, I found errors in both iND750 and Yeast 5, and made additional corrections to Yeast 5. Additionally, I found a set of literature-supported constraints to reaction reversibility which improved flux predictions in the Yeast model. Following these corrections, screens of simulated knockout strains using the updated Yeast model made different predictions of knockouts that would lead to a formate-producing

phenotype than earlier yeast metabolic models. Thus, I found that a comparative approach of evaluating phenotype and metabolic flux predictions using different metabolic models is a valuable tool for model refinement and improvement (Chapter 5).

Thus, the primary conclusion drawn from this work is that literature-based curation of stoichiometrically constrained models of yeast metabolism can improve model accuracy and predictive ability. In turn, improved models can be applied to develop metabolic engineering objectives.

Metabolic reconstructions contain reactions which are descriptive of established biochemistry, but also include hypothetical reactions intended to improve model predictive accuracy or to enable mathematical analysis. The distinction between established knowledge and hypothesis in such models is not always clear. Improved documentation would make such models better scientific tools. Similarly, though algorithms and proofs of concept are reported in the literature for a variety of techniques for computational analysis of such models, these methods may be implemented in different ways, yielding differing results. If metabolic models are to be used for hypothesis generation, the methods must be better documented.

Despite the drawbacks of limited documentation, mass balance-based stoichiometrically constrained metabolic models have a great deal of still-unrealized potential.

Future research

There are many opportunities to apply the application of stoichiometrically constrained models to future research, and many opportunities to refine the scope and techniques of this modeling approach. For example, if a metabolic reconstruction accurately reflects established biochemical knowledge, then any blocked reactions in the model – those reactions which cannot carry a flux because they are not connected (in a graph-theoretical sense) to the rest of the metabolic network – reflect gaps in our knowledge of metabolism. As this work demonstrates, the process of improving the reconstruction of established biochemistry is an iterative effort to

improve the model.

The differing predictions of phenotypes predicted in simulations involving the uncharacterized open reading frame *ALT2* highlights the ongoing need for additional characterization of yeast genes. The Saccharomyces Genome Database currently includes 848 uncharacterized open reading frames and 785 dubious open reading frames. As these open reading frames are better characterized, new information must be incorporated to computational reconstructions of metabolism.

Simulations using the set of reaction constraints generated for Chapter 5 make predictions which both suggest future experimental work and provide a target for further model curation. Specifically, the double knockout mutants predicted to produce formate by the model (Table 5.4) should be constructed *in vivo* and the resulting strains should be tested to determine whether these model predictions are accurate. The prediction that knockouts in the Lysine biosynthetic pathway might affect C1 metabolism are novel, and the reconstruction of amino acid pathways is not a portion of yeast metabolism that was carefully evaluated during the course of this work. Thus, amino acid metabolism is a good target for future model curation and, if necessary, refinement. Similarly, there remains opportunity for additional literature-based model curation, of redox reactions, lipid, amino acid, and nitrogen metabolism, and Acetyl-CoA metabolism.

Application of constraint-based models to scientific and engineering research would also benefit from additional consideration of practical issues. Despite the proliferation of models of yeast metabolism, there are not well-defined metrics or tools for evaluating models. For example, there is not a set of well-defined phenotypes or simulated media that can serve as an objective standard for evaluating computational predictions. Additionally, the ongoing effort to improve metabolic reconstructions would benefit from tools which facilitated collaborative curation, documentation, and model application.

The future also holds opportunity for expanding modeling approaches such as flux balance analysis beyond metabolism. Our knowledge of the relationship between signaling, transcriptional regulation, and the metabolic network is incomplete, and the distinction between

these aspects of physiology is not clear-cut. As human understanding of these interrelated cellular processes improves, modeling techniques will need to be refined to encompass a broader range of biochemical reactions than that which is currently described in metabolic models. The updated version of the Yeast model presented in this work accounts for 904 open reading frames, which is 13.7% of those currently listed in the Saccharomyces Genome Database. As our understanding of yeast biochemistry improves, model coverage may expand to account for the entire yeast genome.

This dissertation contributes to the ongoing effort to develop systems biotechnology tools which will enable the rational design of new microbial strains, and which may enable broader industrial-scale application of biotechnology. The fields of computational biology and systems biotechnology are young, and abundant opportunity remains to develop and apply this technology to meet human needs.

REFERENCES

1. Herrgard MJ, Swainston N, Dobson P, Dunn WB, Arga KY, Arvas M, Buthgen N, Borger S, Costenoble R, Heinemann M, Hucka M, Le Novere N, Li P, Liebermeister W, Mo ML, Oliveira AP, Petranovic D, Pettifer S, Simeonidis E, Smallbone K, Spasie I, Weichart D, Brent R, Broomhead DS, Westerhoff HV, Kurdar B, Penttila M, Klipp E, Palsson BO, Sauer U, Oliver SG, Mendes P, Nielsen J, Kell DB: **A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology**. *Nat Biotech* 2008, **26**:1155–1160.

2. Dobson PD, Jameson D, Simeonidis E, Lanthaler K, Pir P, Lu C, Swainston N, Dunn WB, Fisher P, Hull D, Brown M, Oshota O, Stanford NJ, Kell DB, King RD, Oliver SG, Stevens RD, Mendes P: **Further developments towards a genome-scale metabolic model of yeast**. *BMC Syst Biol* 2010, **4**:145.

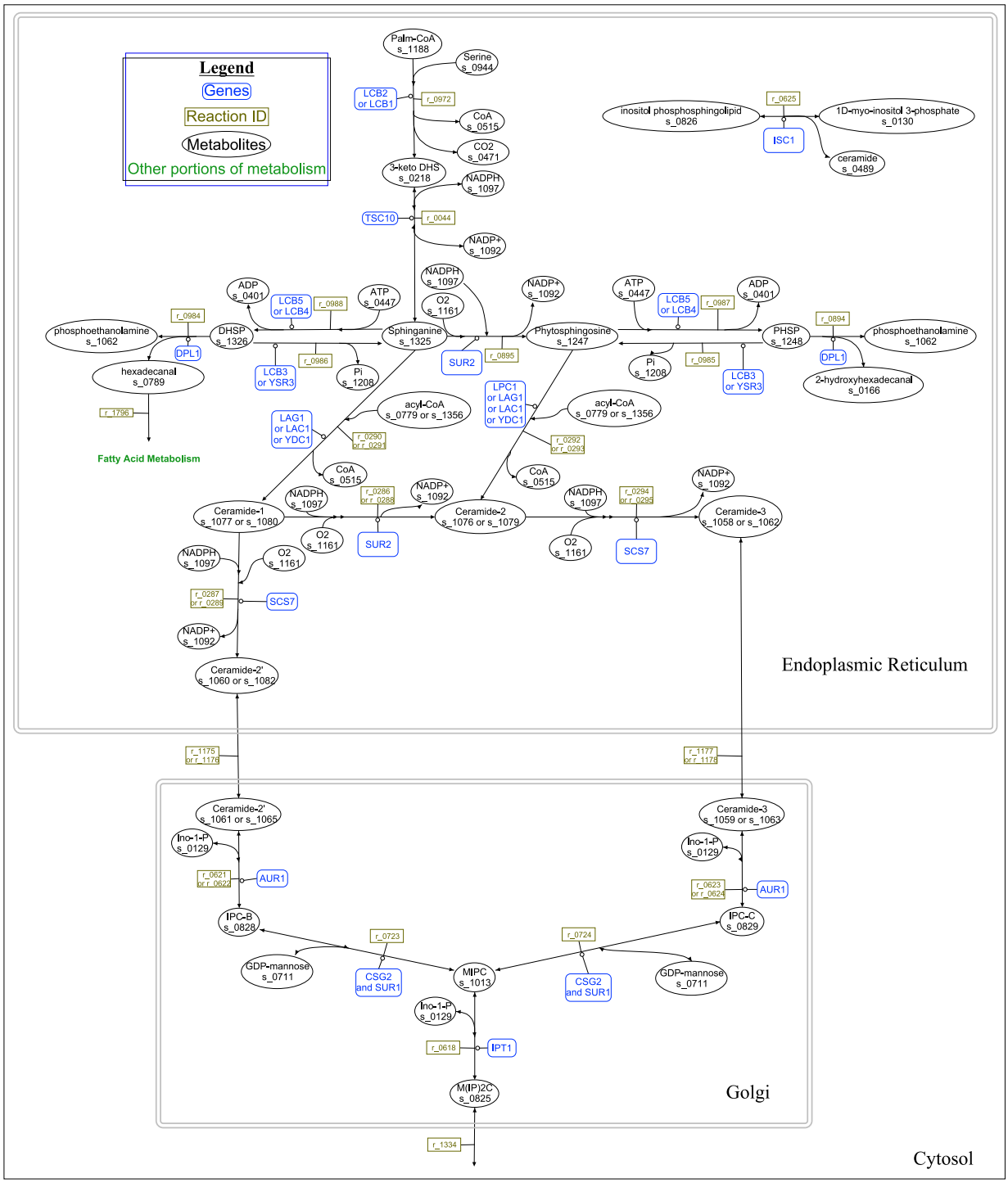
3. Heavner BD, Henry SA, Walker LP: **Evaluating Sphingolipid Biochemistry in the Consensus Reconstruction of Yeast Metabolism**. *Industrial Biotechnology* 2012, **8**:72–78.

4. Heavner BD, Smallbone K, Barker B, Mendes P, Walker LP: **Yeast 5 - an expanded reconstruction of the Saccharomyces Cerevisiae metabolic network**. *BMC Systems Biology* 2012, **6**:55.

APPENDIX 1 – SUPPLEMENTARY INFORMATION FOR CHAPTER 3

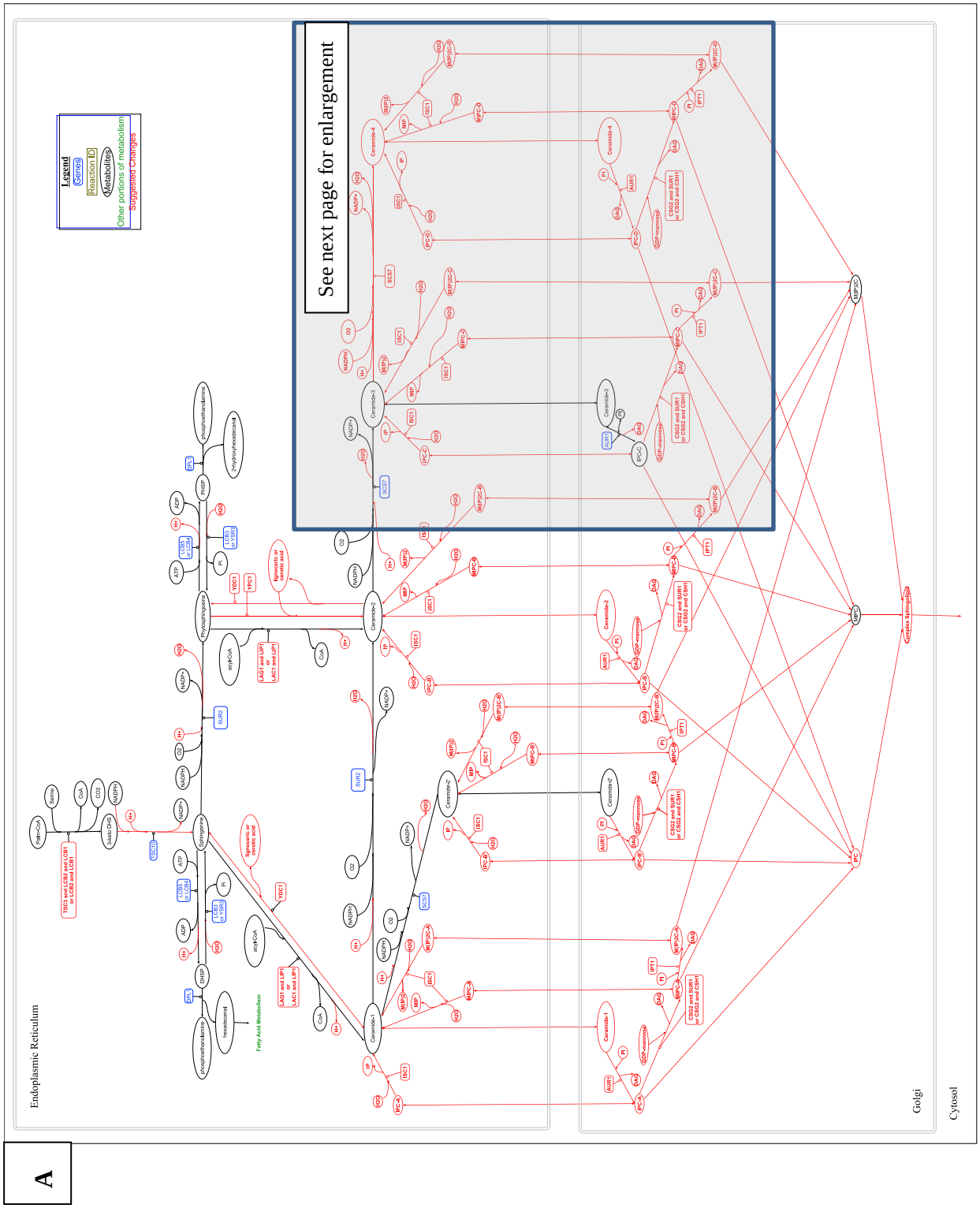
Supplementary Figure 3.1: The incomplete reconstruction of sphingolipid metabolism in Yeast v4.05

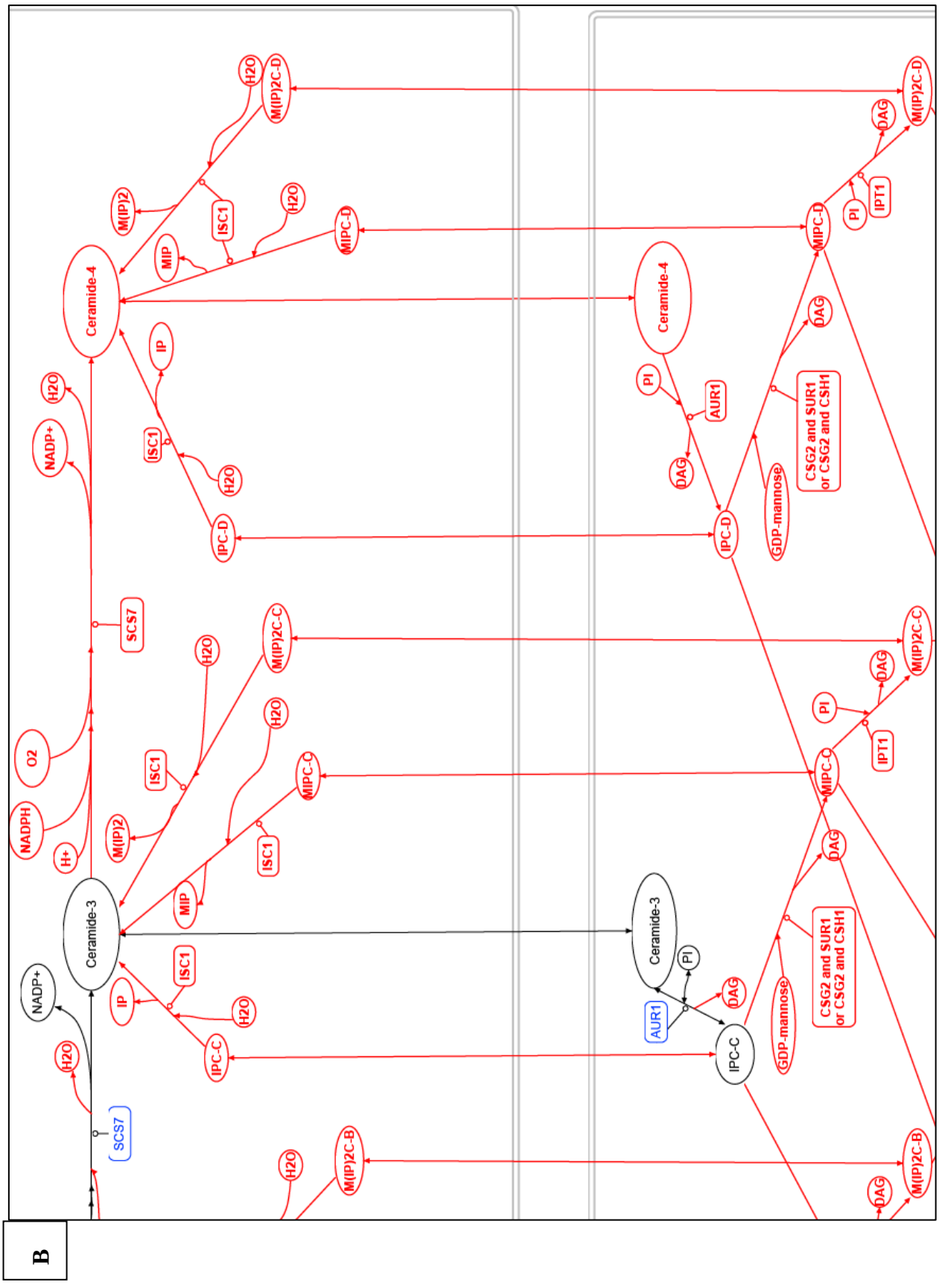
The Yeast v4.05 reconstruction of yeast sphingolipid metabolism includes incomplete and inconsistent differentiation between ceramide and sphingolipid species which vary by hydroxylation or length of the VLCFA moiety.


Legend
Genes
Reaction ID
Metabolites
Other portions of metabolism
Palm-CoA s_1188
Serine s_0944
LCB2 or LCB1
r_0972
CoA s_0515
CO2 s_0471
3-keto DHS s_0218
NADPH s_1097
TSC10
r_0044
NADP+ s_1092
inositol phosphosphingolipid s_0826
r_0625
ISC1
1D-myo-inositol 3-phosphate s_0130
ceramide s_0489
ADP s_0401
LCB5 or LCB4
r_0988
ATP s_0447
O2 s_1161
r_0984
phosphoethanolamine s_1062
DHSP s_1326
Sphinganine s_1325
Phytosphingosine s_1247
PHSP s_1248
r_0894
phosphoethanolamine s_1062
DPL1
hexadecanal s_0789
LCB3 or YSR3
r_0986
Pi s_1208
SUR2
r_0895
r_0987
r_0985
2-hydroxyhexadecanal s_0166
r_1796
Fatty Acid Metabolism
LAG1 or LAC1 or YDC1
acyl-CoA s_0779 or s_1356
LPC1 or LAG1 or LAC1 or YDC1
r_0290 or r_0291
r_0292 or r_0293
Ceramide-1 s_1077 or s_1080
r_0286 or r_0288
Ceramide-2 s_1076 or s_1079
r_0294 or r_0295
Ceramide-3 s_1058 or s_1062
SCS7
r_0287 or r_0289
Ceramide-2' s_1060 or s_1082
Endoplasmic Reticulum
r_1175 or r_1176
r_1177 or r_1178
Ceramide-2' s_1061 or s_1065
Ceramide-3 s_1059 or s_1063
Ino-1-P s_0129
r_0621 or r_0622
AUR1
r_0623 or r_0624
IPC-B s_0828
IPC-C s_0829
r_0723
r_0724
GDP-mannose s_0711
CSG2 and SUR1
MIPC s_1013
r_0618
IPT1
M(IP)2C s_0825
Golgi
r_1334
Cytosol

Supplementary Figure 3.2: Suggested modifications to the reconstruction of yeast sphingolipid metabolism

Modifications to the Yeast v4.05 reconstruction of yeast sphingolipid metabolism (red) are necessary to resolve differences between the current reconstruction and the pathway described in yeast biochemical literature. These proposed changes (which are compiled with references as Supplementary Table 3.1) have been submitted to the maintainers of Yeast Reconstruction for future expansion. The figure labeled A shows all modifications. The figure labeled B is a portion of the network, enlarged for legibility.


A
Endoplasmic Reticulum
Golgi
Cytosol
Legend
Genes
Reaction ID
Metabolites
Other portions of metabolism
Suggested Changes
See next page for enlargement
Fatty Acid Metabolism
Complex Sphingolipids


B
Ceramide-4
Ceramide-3
IPC-D
IPC-C
MIPC-D
MIPC-C
M(IP)2C-D
M(IP)2C-C
M(IP)2C-B
M(IP)2
MIP
IP
ISC1
SCS7
H2O
NADP+
NADPH
O2
H+
PI
DAG
AUR1
IPT1
GDP-mannose
CSG2 and SUR1
or CSG2 and CSH1

Supplementary Table 3.1 is the literature-based reaction list, with references.

Reactions	gene locus	Gene	References (DOI, first author)
'L-serine [endoplasmic reticulum] + palmitoyl-CoA [endoplasmic reticulum] <=> 3-ketosphinganine [endoplasmic reticulum] + carbon dioxide [endoplasmic reticulum] + coenzyme A [endoplasmic reticulum] '	'((YBR058C-A and YDR062W and YMR296C) or (YDR062W and YMR296C))'	(TSC3 and LCB2 and LCB1) or (LCB2 and LCB1)	PNAS August 16, 1994 vol. 91 no. 17 7899-7902 (Nagiec) J Bacteriol. 1991 July; 173(14): 4325-4332 (Buede) 10.1074/jbc.275.11.7597 (Gable) 10.1194/jlr.R800003-JLR200 (dickson)
'3-ketosphinganine [endoplasmic reticulum] + 2 H+ [endoplasmic reticulum] + NADPH [endoplasmic reticulum] <=> NADP(+) [endoplasmic reticulum] + sphinganine [endoplasmic reticulum] '	'YBR265W'	TSC10	10.1074/jbc.273.46.30688 (Beeler) 10.1194/jlr.R800003-JLR200 (dickson)
'sphinganine [endoplasmic reticulum] + tetracosanoyl-CoA [endoplasmic reticulum] <=> ceramide-1 (C24) [endoplasmic reticulum] + coenzyme A [endoplasmic reticulum] + H+ [endoplasmic reticulum] '	'((YHL003C and YMR298W) or (YKL008C and YMR298W))'	(LAG1 and LIP1) or (LAC1 and LIP1)	10.1093/emboj/20.11.2655 (Guillas) 10.1038/sj.emboj.7600562 (Valee) 10.1194/jlr.R800003-JLR200 (dickson)
'hexacosanoyl-CoA [endoplasmic reticulum] + sphinganine [endoplasmic reticulum] <=> ceramide-1 (C26) [endoplasmic reticulum] + coenzyme A [endoplasmic reticulum] + H+ [endoplasmic reticulum] '	'((YHL003C and YMR298W) or (YKL008C and YMR298W))'	(LAG1 and LIP1) or (LAC1 and LIP1)	10.1093/emboj/20.11.2655 (Guillas) 10.1038/sj.emboj.7600562 (Valee) 10.1194/jlr.R800003-JLR200 (dickson)
'H+ [endoplasmic reticulum] + NADPH [endoplasmic reticulum] + oxygen [endoplasmic reticulum] + sphinganine [endoplasmic reticulum] <=> H2O [endoplasmic reticulum] + NADP(+) [endoplasmic reticulum] + phytosphingosine [endoplasmic reticulum] '	'YDR297W'	SUR2	10.1194/jlr.R800003-JLR200 (dickson) 10.1074/jbc.272.47.29704 (haak)

'ceramide-1 (C24) [endoplasmic reticulum] + H+ [endoplasmic reticulum] + NADPH [endoplasmic reticulum] + oxygen [endoplasmic reticulum] <=> ceramide-2 (C24) [endoplasmic reticulum] + H2O [endoplasmic reticulum] + NADP(+) [endoplasmic reticulum] '	'YDR297W'	SUR2	10.1194/jlr.R800003-JLR200 (dickson) 10.1074/jbc.272.47.29704 (haak)
'ceramide-1 (C26) [endoplasmic reticulum] + H+ [endoplasmic reticulum] + NADPH [endoplasmic reticulum] + oxygen [endoplasmic reticulum] <=> ceramide-2 (C26) [endoplasmic reticulum] + H2O [endoplasmic reticulum] + NADP(+) [endoplasmic reticulum] '	'YDR297W'	SUR2	10.1194/jlr.R800003-JLR200 (dickson) 10.1074/jbc.272.47.29704 (haak)
'phytosphingosine [endoplasmic reticulum] + tetracosanoyl-CoA [endoplasmic reticulum] <=> ceramide-2 (C24) [endoplasmic reticulum] + coenzyme A [endoplasmic reticulum] + H+ [endoplasmic reticulum] '	'((YHL003C and YMR298W) or (YKL008C and YMR298W))'	(LAG1 and LIP1) or (LAC1 and LIP1)	10.1093/emboj/20.11.2655 (Guillas) 10.1038/sj.emboj.7600562 (Valee) 10.1194/jlr.R800003-JLR200 (dickson)
'hexacosanoyl-CoA [endoplasmic reticulum] + phytosphingosine [endoplasmic reticulum] <=> ceramide-2 (C26) [endoplasmic reticulum] + coenzyme A [endoplasmic reticulum] + H+ [endoplasmic reticulum] '	'((YHL003C and YMR298W) or (YKL008C and YMR298W))'	(LAG1 and LIP1) or (LAC1 and LIP1)	10.1093/emboj/20.11.2655 (Guillas) 10.1038/sj.emboj.7600562 (Valee) 10.1194/jlr.R800003-JLR200 (dickson)
'ATP [endoplasmic reticulum] + phytosphingosine [endoplasmic reticulum] <=> ADP [endoplasmic reticulum] + H+ [endoplasmic reticulum] + phytosphingosine 1-phosphate [endoplasmic reticulum] '	'(YLR260W or YOR171C)'	LCB5 or LCB4	10.1074/jbc.273.31.19437 (Nagiec) 10.1194/jlr.R800003-JLR200 (dickson)

'ATP [endoplasmic reticulum] + sphinganine [endoplasmic reticulum] <=> ADP [endoplasmic reticulum] + H+ [endoplasmic reticulum] + sphinganine 1-phosphate [endoplasmic reticulum] '	'(YLR260W or YOR171C)'	LCB5 or LCB4	10.1074/jbc.273.31.19437 (Nagiec) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + phytosphingosine 1-phosphate [endoplasmic reticulum] <=> phosphate [endoplasmic reticulum] + phytosphingosine [endoplasmic reticulum] '	'(YJL134W or YKR053C)'	LCB3 or YSR3	PNAS January 6, 1998 vol. 95 no. 1 150-155 (Mandala) 10.1074/jbc.272.45.28690 (Mao) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + sphinganine 1-phosphate [endoplasmic reticulum] <=> phosphate [endoplasmic reticulum] + sphinganine [endoplasmic reticulum] '	'(YJL134W or YKR053C)'	LCB3 or YSR3	PNAS January 6, 1998 vol. 95 no. 1 150-155 (Mandala) 10.1074/jbc.272.45.28690 (Mao) 10.1194/jlr.R800003-JLR200 (dickson)
phytosphingosine 1-phosphate [endoplasmic reticulum] <=> 2-hydroxyhexadecanal [endoplasmic reticulum] + O-phosphoethanolamine [endoplasmic reticulum]	YDR294C'	DPL1	10.1074/jbc.272.42.26087 (Saba) 10.1194/jlr.R800003-JLR200 (dickson)
'sphinganine 1-phosphate [endoplasmic reticulum] <=> hexadecanal [endoplasmic reticulum] + O-phosphoethanolamine [endoplasmic reticulum] '	YDR294C'	DPL1	10.1074/jbc.272.42.26087 (Saba) 10.1194/jlr.R800003-JLR200 (dickson)
'1-phosphatidyl-1D-myo-inositol [Golgi] + ceramide-2'' (C24) [Golgi] <=> diglyceride [Golgi] + inositol-P-ceramide B'' (C24) [Golgi] '	'YKL004W'	AUR1	10.1074/jbc.272.15.9809 (Nagiec) 10.1194/jlr.R800003-JLR200 (dickson)
'1-phosphatidyl-1D-myo-inositol [Golgi] + ceramide-2'' (C26) [Golgi] <=> diglyceride [Golgi] + inositol-P-ceramide B'' (C26) [Golgi] '	YKL004W'	AUR1	10.1074/jbc.272.15.9809 (Nagiec)10.1194/jlr.R800003-JLR200 (dickson)
'1-phosphatidyl-1D-myo-inositol [Golgi] + ceramide-3 (C24) [Golgi] <=> diglyceride [Golgi] + inositol-P-ceramide C (C24) [Golgi] '	'YKL004W'	AUR1	10.1074/jbc.272.15.9809 (Nagiec) 10.1194/jlr.R800003-JLR200 (dickson)

'1-phosphatidyl-1D-myo-inositol [Golgi] + ceramide-3 (C26) [Golgi] <=> diglyceride [Golgi] + inositol-P-ceramide C (C26) [Golgi] '	'YKL004W'	AUR1	10.1074/jbc.272.15.9809 (Nagiec) 10.1194/jlr.R800003-JLR200 (dickson)
'1-phosphatidyl-1D-myo-inositol [Golgi] + ceramide-1 (C24) [Golgi] <=> diglyceride [Golgi] + inositol-P-ceramide A (C24) [Golgi] '	'YKL004W'	AUR1	10.1074/jbc.272.15.9809 (Nagiec) 10.1194/jlr.R800003-JLR200 (dickson)
'1-phosphatidyl-1D-myo-inositol [Golgi] + ceramide-1 (C26) [Golgi] <=> diglyceride [Golgi] + inositol-P-ceramide A (C26) [Golgi] '	'YKL004W'	AUR1	10.1074/jbc.272.15.9809 (Nagiec) 10.1194/jlr.R800003-JLR200 (dickson)
'1-phosphatidyl-1D-myo-inositol [Golgi] + ceramide-2 (C24) [Golgi] <=> diglyceride [Golgi] + inositol-P-ceramide B (C24) [Golgi] '	'YKL004W'	AUR1	10.1074/jbc.272.15.9809 (Nagiec) 10.1194/jlr.R800003-JLR200 (dickson)
'1-phosphatidyl-1D-myo-inositol [Golgi] + ceramide-2 (C26) [Golgi] <=> diglyceride [Golgi] + inositol-P-ceramide B (C26) [Golgi] '	'YKL004W'	AUR1	10.1074/jbc.272.15.9809 (Nagiec) 10.1194/jlr.R800003-JLR200 (dickson)
'1-phosphatidyl-1D-myo-inositol [Golgi] + ceramide-4 (C24) [Golgi] <=> diglyceride [Golgi] + inositol-P-ceramide D (C24) [Golgi] '	'YKL004W'	AUR1	10.1074/jbc.272.15.9809 (Nagiec) 10.1194/jlr.R800003-JLR200 (dickson)
'1-phosphatidyl-1D-myo-inositol [Golgi] + ceramide-4 (C26) [Golgi] <=> diglyceride [Golgi] + inositol-P-ceramide D (C26) [Golgi] '	'YKL004W'	AUR1	10.1074/jbc.272.15.9809 (Nagiec) 10.1194/jlr.R800003-JLR200 (dickson)
'GDP-alpha-D-mannose [Golgi] + inositol-P-ceramide A (C24) [Golgi] <=> GDP [Golgi] + mannosylinositol phosphorylceramide A (C24) [Golgi] '	'((YBR036C and YPL057C) or (YBR036C and YBR161W))'	(CSG2 and CSG1) or (CSG2 and CSH1)	10.1074/jbc.M305498200 (Uemura) 10.1194/jlr.R800003-JLR200 (dickson)
'GDP-alpha-D-mannose [Golgi] + inositol-P-ceramide A (C26) [Golgi] <=> GDP [Golgi] + mannosylinositol phosphorylceramide A (C26) [Golgi] '	'((YBR036C and YPL057C) or (YBR036C and YBR161W))'	(CSG2 and CSG1) or (CSG2 and CSH1)	10.1074/jbc.M305498200 (Uemura) 10.1194/jlr.R800003-JLR200 (dickson)

'GDP-alpha-D-mannose [Golgi] + inositol-P-ceramide B'' (C24) [Golgi] <=> GDP [Golgi] + mannosylinositol phosphorylceramide B'' (C24) [Golgi] '	'((YBR036C and YPL057C) or (YBR036C and YBR161W))'	(CSG2 and CSG1) or (CSG2 and CSH1)	10.1074/jbc.M305498200 (Uemura) 10.1194/jlr.R800003-JLR200 (dickson)
'GDP-alpha-D-mannose [Golgi] + inositol-P-ceramide B'' (C26) [Golgi] <=> GDP [Golgi] + mannosylinositol phosphorylceramide B'' (C26) [Golgi] '	'((YBR036C and YPL057C) or (YBR036C and YBR161W))'	(CSG2 and CSG1) or (CSG2 and CSH1)	10.1074/jbc.M305498200 (Uemura) 10.1194/jlr.R800003-JLR200 (dickson)
'GDP-alpha-D-mannose [Golgi] + inositol-P-ceramide B (C24) [Golgi] <=> GDP [Golgi] + mannosylinositol phosphorylceramide B (C24) [Golgi] '	'((YBR036C and YPL057C) or (YBR036C and YBR161W))'	(CSG2 and CSG1) or (CSG2 and CSH1)	10.1074/jbc.M305498200 (Uemura) 10.1194/jlr.R800003-JLR200 (dickson)
'GDP-alpha-D-mannose [Golgi] + inositol-P-ceramide B (C26) [Golgi] <=> GDP [Golgi] + mannosylinositol phosphorylceramide B (C26) [Golgi] '	'((YBR036C and YPL057C) or (YBR036C and YBR161W))'	(CSG2 and CSG1) or (CSG2 and CSH1)	10.1074/jbc.M305498200 (Uemura) 10.1194/jlr.R800003-JLR200 (dickson)
'GDP-alpha-D-mannose [Golgi] + inositol-P-ceramide C (C24) [Golgi] <=> GDP [Golgi] + mannosylinositol phosphorylceramide C (C24) [Golgi] '	'((YBR036C and YPL057C) or (YBR036C and YBR161W))'	(CSG2 and CSG1) or (CSG2 and CSH1)	10.1074/jbc.M305498200 (Uemura) 10.1194/jlr.R800003-JLR200 (dickson)
'GDP-alpha-D-mannose [Golgi] + inositol-P-ceramide C (C26) [Golgi] <=> GDP [Golgi] + mannosylinositol phosphorylceramide C (C26) [Golgi] '	'((YBR036C and YPL057C) or (YBR036C and YBR161W))'	(CSG2 and CSG1) or (CSG2 and CSH1)	10.1074/jbc.M305498200 (Uemura) 10.1194/jlr.R800003-JLR200 (dickson)
'GDP-alpha-D-mannose [Golgi] + inositol-P-ceramide D (C24) [Golgi] <=> GDP [Golgi] + mannosylinositol phosphorylceramide D (C24) [Golgi] '	'((YBR036C and YPL057C) or (YBR036C and YBR161W))'	(CSG2 and CSG1) or (CSG2 and CSH1)	10.1074/jbc.M305498200 (Uemura) 10.1194/jlr.R800003-JLR200 (dickson)
'GDP-alpha-D-mannose [Golgi] + inositol-P-ceramide D (C26) [Golgi] <=> GDP [Golgi] + mannosylinositol phosphorylceramide D (C26)	'((YBR036C and YPL057C) or (YBR036C and YBR161W))'	(CSG2 and CSG1) or (CSG2 and CSH1)	10.1074/jbc.M305498200 (Uemura) 10.1194/jlr.R800003-JLR200 (dickson)

[Golgi] '			
'1-phosphatidyl-1D-myo-inositol [Golgi] + mannosylinositol phosphorylceramide A (C24) [Golgi] <=> diglyceride [Golgi] + inositol phosphomannosylinositol phosphoceramide A (C24) [Golgi] '	'YDR072C'	IPT1	10.1074/jbc.272.47.29620 (dickson) 10.1194/jlr.R800003-JLR200 (dickson)
'1-phosphatidyl-1D-myo-inositol [Golgi] + mannosylinositol phosphorylceramide A (C26) [Golgi] <=> diglyceride [Golgi] + inositol phosphomannosylinositol phosphoceramide A (C26) [Golgi] '	'YDR072C'	IPT1	10.1074/jbc.272.47.29620 (dickson) 10.1194/jlr.R800003-JLR200 (dickson)
'1-phosphatidyl-1D-myo-inositol [Golgi] + mannosylinositol phosphorylceramide B'' (C24) [Golgi] <=> diglyceride [Golgi] + inositol phosphomannosylinositol phosphoceramide B'' (C24) [Golgi] '	'YDR072C'	IPT1	10.1074/jbc.272.47.29620 (dickson) 10.1194/jlr.R800003-JLR200 (dickson)
'1-phosphatidyl-1D-myo-inositol [Golgi] + mannosylinositol phosphorylceramide B'' (C26) [Golgi] <=> diglyceride [Golgi] + inositol phosphomannosylinositol phosphoceramide B'' (C26) [Golgi] '	'YDR072C'	IPT1	10.1074/jbc.272.47.29620 (dickson) 10.1194/jlr.R800003-JLR200 (dickson)
'1-phosphatidyl-1D-myo-inositol [Golgi] + mannosylinositol phosphorylceramide B (C24) [Golgi] <=> diglyceride [Golgi] + inositol phosphomannosylinositol phosphoceramide B (C24)	'YDR072C'	IPT1	10.1074/jbc.272.47.29620 (dickson)10.1194/jlr.R800003-JLR200 (dickson)

[Golgi] '			
'1-phosphatidyl-1D-myo-inositol [Golgi] + mannosylinositol phosphorylceramide B (C26) [Golgi] <=> diglyceride [Golgi] + inositol phosphomannosylinositol phosphoceramide B (C26) [Golgi] '	'YDR072C'	IPT1	10.1074/jbc.272.47.29620 (dickson) 10.1194/jlr.R800003-JLR200 (dickson)
'1-phosphatidyl-1D-myo-inositol [Golgi] + mannosylinositol phosphorylceramide C (C24) [Golgi] <=> diglyceride [Golgi] + inositol phosphomannosylinositol phosphoceramide C (C24) [Golgi] '	'YDR072C'	IPT1	10.1074/jbc.272.47.29620 (dickson) 10.1194/jlr.R800003-JLR200 (dickson)
'1-phosphatidyl-1D-myo-inositol [Golgi] + mannosylinositol phosphorylceramide C (C26) [Golgi] <=> diglyceride [Golgi] + inositol phosphomannosylinositol phosphoceramide C (C26) [Golgi] '	'YDR072C'	IPT1	10.1074/jbc.272.47.29620 (dickson) 10.1194/jlr.R800003-JLR200 (dickson)
'1-phosphatidyl-1D-myo-inositol [Golgi] + mannosylinositol phosphorylceramide D (C24) [Golgi] <=> diglyceride [Golgi] + inositol phosphomannosylinositol phosphoceramide D (C24) [Golgi] '	'YDR072C'	IPT1	10.1074/jbc.272.47.29620 (dickson) 10.1194/jlr.R800003-JLR200 (dickson)
'1-phosphatidyl-1D-myo-inositol [Golgi] + mannosylinositol phosphorylceramide D (C26) [Golgi] <=> diglyceride [Golgi] + inositol	YDR072C'	IPT1	10.1074/jbc.272.47.29620 (dickson) 10.1194/jlr.R800003-JLR200 (dickson)

phosphomannosylinositol phosphoceramide D (C26) [Golgi] '			
'H2O [endoplasmic reticulum] + mannosylinositol phosphorylceramide A (C24) [endoplasmic reticulum] <=> ceramide-1 (C24) [endoplasmic reticulum] + mannose-1D-myo-inositol 1-phosphate [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + mannosylinositol phosphorylceramide A (C26) [endoplasmic reticulum] <=> ceramide-1 (C26) [endoplasmic reticulum] + mannose-1D-myo-inositol 1-phosphate [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + mannosylinositol phosphorylceramide B'' (C24) [endoplasmic reticulum] <=> ceramide-2'' (C24) [endoplasmic reticulum] + mannose-1D-myo-inositol 1-phosphate [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + mannosylinositol phosphorylceramide B'' (C26) [endoplasmic reticulum] <=> ceramide-2'' (C26) [endoplasmic reticulum] + mannose-1D-myo-inositol 1-phosphate [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + mannosylinositol phosphorylceramide B (C24) [endoplasmic reticulum] <=> ceramide-2 (C24) [endoplasmic reticulum] +	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)

mannose-1D-myo-inositol 1-phosphate [endoplasmic reticulum] '			
'H2O [endoplasmic reticulum] + mannosylinositol phosphorylceramide B (C26) [endoplasmic reticulum] <=> ceramide-2 (C26) [endoplasmic reticulum] + mannose-1D-myo-inositol 1-phosphate [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + mannosylinositol phosphorylceramide C (C24) [endoplasmic reticulum] <=> ceramide-3 (C24) [endoplasmic reticulum] + mannose-1D-myo-inositol 1-phosphate [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + mannosylinositol phosphorylceramide C (C26) [endoplasmic reticulum] <=> ceramide-3 (C26) [endoplasmic reticulum] + mannose-1D-myo-inositol 1-phosphate [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + mannosylinositol phosphorylceramide D (C24) [endoplasmic reticulum] <=> ceramide-4 (C24) [endoplasmic reticulum] + mannose-1D-myo-inositol 1-phosphate [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + mannosylinositol phosphorylceramide D (C26) [endoplasmic reticulum] <=> ceramide-4 (C26) [endoplasmic reticulum] +	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai)10.1194/jlr.R800003-JLR200 (dickson)

mannose-1D-myo-inositol 1-phosphate [endoplasmic reticulum] '			
'H2O [endoplasmic reticulum] + inositol phosphomannosylinositol phosphoceramide A (C24) [endoplasmic reticulum] <=> ceramide-1 (C24) [endoplasmic reticulum] + mannose-(1D-myo-inositol 1-phosphate)2 [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + inositol phosphomannosylinositol phosphoceramide A (C26) [endoplasmic reticulum] <=> ceramide-1 (C26) [endoplasmic reticulum] + mannose-(1D-myo-inositol 1-phosphate)2 [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + inositol phosphomannosylinositol phosphoceramide B'' (C24) [endoplasmic reticulum] <=> ceramide-2'' (C24) [endoplasmic reticulum] + mannose-(1D-myo-inositol 1-phosphate)2 [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + inositol phosphomannosylinositol phosphoceramide B'' (C26) [endoplasmic reticulum] <=> ceramide-2'' (C26) [endoplasmic reticulum] + mannose-(1D-myo-inositol 1-phosphate)2 [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)

'H2O [endoplasmic reticulum] + inositol phosphomannosylinositol phosphoceramide B (C24) [endoplasmic reticulum] <=> ceramide-2 (C24) [endoplasmic reticulum] + mannose-(1D-myo-inositol 1-phosphate)2 [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + inositol phosphomannosylinositol phosphoceramide B (C26) [endoplasmic reticulum] <=> ceramide-2 (C26) [endoplasmic reticulum] + mannose-(1D-myo-inositol 1-phosphate)2 [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + inositol phosphomannosylinositol phosphoceramide C (C24) [endoplasmic reticulum] <=> ceramide-3 (C24) [endoplasmic reticulum] + mannose-(1D-myo-inositol 1-phosphate)2 [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + inositol phosphomannosylinositol phosphoceramide C (C26) [endoplasmic reticulum] <=> ceramide-3 (C26) [endoplasmic reticulum] + mannose-(1D-myo-inositol 1-phosphate)2 [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + inositol phosphomannosylinositol phosphoceramide D (C24) [endoplasmic reticulum] <=> ceramide-4 (C24) [endoplasmic reticulum] +	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)

mannose-(1D-myo-inositol 1-phosphate)2 [endoplasmic reticulum] '			
'H2O [endoplasmic reticulum] + inositol phosphomannosylinositol phosphoceramide D (C26) [endoplasmic reticulum] <=> ceramide-4 (C26) [endoplasmic reticulum] + mannose-(1D-myo-inositol 1-phosphate)2 [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [mitochondrion] + inositol-P-ceramide A (C24) [mitochondrion] <=> 1D-myo-inositol 3-phosphate [mitochondrion] + ceramide-1 (C24) [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + inositol-P-ceramide A (C26) [mitochondrion] <=> 1D-myo-inositol 3-phosphate [mitochondrion] + ceramide-1 (C26) [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + inositol-P-ceramide B'' (C24) [mitochondrion] <=> 1D-myo-inositol 3-phosphate [mitochondrion] + ceramide-2'' (C24) [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + inositol-P-ceramide B'' (C26) [mitochondrion] <=> 1D-myo-inositol 3-phosphate [mitochondrion] + ceramide-2'' (C26) [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + inositol-P-ceramide B (C24) [mitochondrion] <=> 1D-myo-inositol 3-phosphate [mitochondrion] + ceramide-2 (C24) [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization

'H2O [mitochondrion] + inositol-P-ceramide B (C26) [mitochondrion] <=> 1D-myo-inositol 3-phosphate [mitochondrion] + ceramide-2 (C26) [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + inositol-P-ceramide C (C24) [mitochondrion] <=> 1D-myo-inositol 3-phosphate [mitochondrion] + ceramide-3 (C24) [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + inositol-P-ceramide C (C26) [mitochondrion] <=> 1D-myo-inositol 3-phosphate [mitochondrion] + ceramide-3 (C26) [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + inositol-P-ceramide D (C24) [mitochondrion] <=> 1D-myo-inositol 3-phosphate [mitochondrion] + ceramide-4 (C24) [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + inositol-P-ceramide D (C26) [mitochondrion] <=> 1D-myo-inositol 3-phosphate [mitochondrion] + ceramide-4 (C26) [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + mannosylinositol phosphorylceramide A (C24) [mitochondrion] <=> ceramide-1 (C24) [mitochondrion] + mannose-1D-myo-inositol 1-phosphate [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + mannosylinositol phosphorylceramide A (C26) [mitochondrion] <=> ceramide-1 (C26) [mitochondrion] + mannose-1D-myo-inositol 1-phosphate [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization

'H2O [mitochondrion] + mannosylinositol phosphorylceramide B'' (C24) [mitochondrion] <=> ceramide-2'' (C24) [mitochondrion] + mannose-1D-myo-inositol 1-phosphate [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + mannosylinositol phosphorylceramide B'' (C26) [mitochondrion] <=> ceramide-2'' (C26) [mitochondrion] + mannose-1D-myo-inositol 1-phosphate [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + mannosylinositol phosphorylceramide B (C24) [mitochondrion] <=> ceramide-2 (C24) [mitochondrion] + mannose-1D-myo-inositol 1-phosphate [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + mannosylinositol phosphorylceramide B (C26) [mitochondrion] <=> ceramide-2 (C26) [mitochondrion] + mannose-1D-myo-inositol 1-phosphate [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + mannosylinositol phosphorylceramide C (C24) [mitochondrion] <=> ceramide-3 (C24) [mitochondrion] + mannose-1D-myo-inositol 1-phosphate [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + mannosylinositol phosphorylceramide C (C26) [mitochondrion] <=> ceramide-3 (C26) [mitochondrion] + mannose-1D-myo-inositol 1-phosphate	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization

[mitochondrion] '			
'H2O [mitochondrion] + mannosylinositol phosphorylceramide D (C24) [mitochondrion] <=> ceramide-4 (C24) [mitochondrion] + mannose-1D-myo-inositol 1-phosphate [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + mannosylinositol phosphorylceramide D (C26) [mitochondrion] <=> ceramide-4 (C26) [mitochondrion] + mannose-1D-myo-inositol 1-phosphate [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + inositol phosphomannosylinositol phosphoceramide A (C24) [mitochondrion] <=> ceramide-1 (C24) [mitochondrion] + mannose-(1D-myo-inositol 1-phosphate)2 [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + inositol phosphomannosylinositol phosphoceramide A (C26) [mitochondrion] <=> ceramide-1 (C26) [mitochondrion] + mannose-(1D-myo-inositol 1-phosphate)2 [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + inositol phosphomannosylinositol phosphoceramide B'' (C24) [mitochondrion] <=> ceramide-2'' (C24) [mitochondrion] + mannose-(1D-myo-inositol 1-phosphate)2 [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization

'H2O [mitochondrion] + inositol phosphomannosylinositol phosphoceramide B'' (C26) [mitochondrion] <=> ceramide-2'' (C26) [mitochondrion] + mannose-(1D-myo-inositol 1-phosphate)2 [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + inositol phosphomannosylinositol phosphoceramide B (C24) [mitochondrion] <=> ceramide-2 (C24) [mitochondrion] + mannose-(1D-myo-inositol 1-phosphate)2 [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + inositol phosphomannosylinositol phosphoceramide B (C26) [mitochondrion] <=> ceramide-2 (C26) [mitochondrion] + mannose-(1D-myo-inositol 1-phosphate)2 [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + inositol phosphomannosylinositol phosphoceramide C (C24) [mitochondrion] <=> ceramide-3 (C24) [mitochondrion] + mannose-(1D-myo-inositol 1-phosphate)2 [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + inositol phosphomannosylinositol phosphoceramide C (C26) [mitochondrion] <=> ceramide-3 (C26) [mitochondrion] + mannose-(1D-myo-inositol 1-phosphate)2 [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization

'H2O [mitochondrion] + inositol phosphomannosylinositol phosphoceramide D (C24) [mitochondrion] <=> ceramide-4 (C24) [mitochondrion] + mannose-(1D-myo-inositol 1-phosphate)2 [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [mitochondrion] + inositol phosphomannosylinositol phosphoceramide D (C26) [mitochondrion] <=> ceramide-4 (C26) [mitochondrion] + mannose-(1D-myo-inositol 1-phosphate)2 [mitochondrion] '	'YER019W'	ISC1	10.1194/jlr.R800003-JLR200 (dickson) 10.1016/j.bbamem.2007.07.019 (Ktagaki) - mitochondrial localization
'H2O [endoplasmic reticulum] + inositol-P-ceramide A (C24) [endoplasmic reticulum] <=> 1D-myo-inositol 1-phosphate [endoplasmic reticulum] + ceramide-1 (C24) [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + inositol-P-ceramide A (C26) [endoplasmic reticulum] <=> 1D-myo-inositol 1-phosphate [endoplasmic reticulum] + ceramide-1 (C26) [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + inositol-P-ceramide B'' (C24) [endoplasmic reticulum] <=> 1D-myo-inositol 1-phosphate [endoplasmic reticulum] + ceramide-2'' (C24) [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + inositol-P-ceramide B'' (C26) [endoplasmic reticulum] <=> 1D-myo-inositol 1-phosphate [endoplasmic reticulum] + ceramide-2'' (C26) [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)

'H2O [endoplasmic reticulum] + inositol-P-ceramide B (C24) [endoplasmic reticulum] <=> 1D-myo-inositol 1-phosphate [endoplasmic reticulum] + ceramide-2 (C24) [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + inositol-P-ceramide B (C26) [endoplasmic reticulum] <=> 1D-myo-inositol 1-phosphate [endoplasmic reticulum] + ceramide-2 (C26) [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + inositol-P-ceramide C (C24) [endoplasmic reticulum] <=> 1D-myo-inositol 1-phosphate [endoplasmic reticulum] + ceramide-3 (C24) [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + inositol-P-ceramide C (C26) [endoplasmic reticulum] <=> 1D-myo-inositol 1-phosphate [endoplasmic reticulum] + ceramide-3 (C26) [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + inositol-P-ceramide D (C24) [endoplasmic reticulum] <=> 1D-myo-inositol 1-phosphate [endoplasmic reticulum] + ceramide-4 (C24) [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)
'H2O [endoplasmic reticulum] + inositol-P-ceramide D (C26) [endoplasmic reticulum] <=> 1D-myo-inositol 1-phosphate [endoplasmic reticulum] + ceramide-4 (C26) [endoplasmic reticulum] '	'YER019W'	ISC1	10.1074/jbc.M007721200 (Sawai) 10.1194/jlr.R800003-JLR200 (dickson)

'ceramide-1 (C24) [endoplasmic reticulum] + H+ [endoplasmic reticulum] + NADPH [endoplasmic reticulum] + oxygen [endoplasmic reticulum] <=> ceramide-2'' (C24) [endoplasmic reticulum] + H2O [endoplasmic reticulum] + NADP(+) [endoplasmic reticulum] '	'YMR272C'	SCS7	10.1194/jlr.R800003-JLR200 (dickson) 10.1074/jbc.272.47.29704 (haak)
'ceramide-1 (C26) [endoplasmic reticulum] + H+ [endoplasmic reticulum] + NADPH [endoplasmic reticulum] + oxygen [endoplasmic reticulum] <=> ceramide-2'' (C26) [endoplasmic reticulum] + H2O [endoplasmic reticulum] + NADP(+) [endoplasmic reticulum] '	'YMR272C'	SCS7	10.1194/jlr.R800003-JLR200 (dickson) 10.1074/jbc.272.47.29704 (haak)
'ceramide-2 (C24) [endoplasmic reticulum] + H+ [endoplasmic reticulum] + NADPH [endoplasmic reticulum] + oxygen [endoplasmic reticulum] <=> ceramide-3 (C24) [endoplasmic reticulum] + H2O [endoplasmic reticulum] + NADP(+) [endoplasmic reticulum] '	'YMR272C'	SCS7	10.1194/jlr.R800003-JLR200 (dickson) 10.1074/jbc.272.47.29704 (haak)
'ceramide-2 (C26) [endoplasmic reticulum] + H+ [endoplasmic reticulum] + NADPH [endoplasmic reticulum] + oxygen [endoplasmic reticulum] <=> ceramide-3 (C26) [endoplasmic reticulum] + H2O [endoplasmic reticulum] + NADP(+) [endoplasmic reticulum] '	'YMR272C'	SCS7	10.1194/jlr.R800003-JLR200 (dickson) 10.1074/jbc.272.47.29704 (haak)

'ceramide-3 (C24) [endoplasmic reticulum] + H+ [endoplasmic reticulum] + NADPH [endoplasmic reticulum] + oxygen [endoplasmic reticulum] <=> ceramide-4 (C24) [endoplasmic reticulum] + H2O [endoplasmic reticulum] + NADP(+) [endoplasmic reticulum] '	'YMR272C'	SCS7	10.1194/jlr.R800003-JLR200 (dickson) 10.1074/jbc.272.47.29704 (haak)
'ceramide-3 (C26) [endoplasmic reticulum] + H+ [endoplasmic reticulum] + NADPH [endoplasmic reticulum] + oxygen [endoplasmic reticulum] <=> ceramide-4 (C26) [endoplasmic reticulum] + H2O [endoplasmic reticulum] + NADP(+) [endoplasmic reticulum] '	'YMR272C'	SCS7	10.1194/jlr.R800003-JLR200 (dickson) 10.1074/jbc.272.47.29704 (haak)
'lignoceric acid [endoplasmic reticulum] + sphinganine [endoplasmic reticulum] <=> ceramide-1 (C24) [endoplasmic reticulum] '	'YPL087W'	YDC1	10.1074/jbc.M003683200 (Mao) 10.1194/jlr.R800003-JLR200 (dickson)
'cerotic acid [endoplasmic reticulum] + sphinganine [endoplasmic reticulum] <=> ceramide-1 (C26) [endoplasmic reticulum] '	'YPL087W'	YDC1	10.1074/jbc.M003683200 (Mao) 10.1194/jlr.R800003-JLR200 (dickson)
'ceramide-2 (C24) [endoplasmic reticulum] -> lignoceric acid [endoplasmic reticulum] + phytosphingosine [endoplasmic reticulum] '	'YPL087W'	YDC1	10.1074/jbc.M003683200 (Mao) 10.1194/jlr.R800003-JLR200 (dickson)
'ceramide-2 (C26) [endoplasmic reticulum] -> cerotic acid [endoplasmic reticulum] + phytosphingosine [endoplasmic reticulum] '	'YPL087W'	YDC1	10.1074/jbc.M003683200 (Mao) 10.1194/jlr.R800003-JLR200 (dickson)
'lignoceric acid [endoplasmic reticulum] + phytosphingosine [endoplasmic reticulum] -> ceramide-2 (C24) [endoplasmic reticulum] '	'YBR183W'	YPC1	10.1074/jbc.275.10.6876 (Mao) 10.1194/jlr.R800003-JLR200 (dickson)

'cerotic acid [endoplasmic reticulum] + phytosphingosine [endoplasmic reticulum] -> ceramide-2 (C26) [endoplasmic reticulum] '	'YBR183W'	YPC1	10.1074/jbc.275.10.6876 (Mao) 10.1194/jlr.R800003-JLR200 (dickson)
'ceramide-2'' (C24) [endoplasmic reticulum] <=> ceramide-2'' (C24) [Golgi] '	''		modelling reaction
'ceramide-2'' (C26) [endoplasmic reticulum] <=> ceramide-2'' (C26) [Golgi] '	''		modelling reaction
'ceramide-3 (C24) [endoplasmic reticulum] <=> ceramide-3 (C24) [Golgi] '	''		modelling reaction
'ceramide-3 (C26) [endoplasmic reticulum] <=> ceramide-3 (C26) [Golgi] '	''		modelling reaction
'inositol-P-ceramide [Golgi] -> complex sphingolipid [Golgi] '	''		modelling reaction
'mannosylinositol phosphorylceramide [Golgi] -> complex sphingolipid [Golgi] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide [Golgi] -> complex sphingolipid [Golgi] '	''		modelling reaction
'inositol-P-ceramide [mitochondrion] -> complex sphingolipid [mitochondrion] '	''		modelling reaction
'mannosylinositol phosphorylceramide [mitochondrion] -> complex sphingolipid [mitochondrion] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide [mitochondrion] -> complex sphingolipid [mitochondrion] '	''		modelling reaction
'inositol-P-ceramide B'' (C24) [Golgi] -> inositol-P-ceramide [Golgi] '	''		modelling reaction
'inositol-P-ceramide B'' (C26) [Golgi] -> inositol-P-ceramide [Golgi] '	''		modelling reaction
'inositol-P-ceramide C (C24) [Golgi] -> inositol-P-ceramide [Golgi] '	''		modelling reaction

'inositol-P-ceramide C (C26) [Golgi] -> inositol-P-ceramide [Golgi] '	''		modelling reaction
'inositol-P-ceramide A (C24) [Golgi] -> inositol-P-ceramide [Golgi] '	''		modelling reaction
'inositol-P-ceramide A (C26) [Golgi] -> inositol-P-ceramide [Golgi] '	''		modelling reaction
'inositol-P-ceramide B (C24) [Golgi] -> inositol-P-ceramide [Golgi] '	''		modelling reaction
'inositol-P-ceramide B (C26) [Golgi] -> inositol-P-ceramide [Golgi] '	''		modelling reaction
'inositol-P-ceramide D (C24) [Golgi] -> inositol-P-ceramide [Golgi] '	''		modelling reaction
'inositol-P-ceramide D (C26) [Golgi] -> inositol-P-ceramide [Golgi] '	''		modelling reaction
'inositol-P-ceramide B'' (C24) [endoplasmic reticulum] -> inositol-P-ceramide [endoplasmic reticulum] '	''		modelling reaction
'inositol-P-ceramide B'' (C26) [endoplasmic reticulum] -> inositol-P-ceramide [endoplasmic reticulum] '	''		modelling reaction
'inositol-P-ceramide C (C24) [endoplasmic reticulum] -> inositol-P-ceramide [endoplasmic reticulum] '	''		modelling reaction
'inositol-P-ceramide C (C26) [endoplasmic reticulum] -> inositol-P-ceramide [endoplasmic reticulum] '	''		modelling reaction
'inositol-P-ceramide A (C24) [endoplasmic reticulum] -> inositol-P-ceramide [endoplasmic reticulum] '	''		modelling reaction
'inositol-P-ceramide A (C26) [endoplasmic reticulum] -> inositol-P-ceramide [endoplasmic reticulum] '	''		modelling reaction
'inositol-P-ceramide B (C24) [endoplasmic reticulum] ->	''		modelling reaction

inositol-P-ceramide [endoplasmic reticulum] '			
'inositol-P-ceramide B (C26) [endoplasmic reticulum] -> inositol-P-ceramide [endoplasmic reticulum] '	''		modelling reaction
'inositol-P-ceramide D (C24) [endoplasmic reticulum] -> inositol-P-ceramide [endoplasmic reticulum] '	''		modelling reaction
'inositol-P-ceramide D (C26) [endoplasmic reticulum] -> inositol-P-ceramide [endoplasmic reticulum] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide A (C24) [Golgi] -> inositol phosphomannosylinositol phosphoceramide [Golgi] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide A (C26) [Golgi] -> inositol phosphomannosylinositol phosphoceramide [Golgi] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide B'' (C24) [Golgi] -> inositol phosphomannosylinositol phosphoceramide [Golgi] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide B'' (C26) [Golgi] -> inositol phosphomannosylinositol phosphoceramide [Golgi] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide B (C24) [Golgi] -> inositol phosphomannosylinositol phosphoceramide [Golgi] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide B (C26)	''		modelling reaction

[Golgi] -> inositol phosphomannosylinositol phosphoceramide [Golgi] '			
'inositol phosphomannosylinositol phosphoceramide C (C24) [Golgi] -> inositol phosphomannosylinositol phosphoceramide [Golgi] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide C (C26) [Golgi] -> inositol phosphomannosylinositol phosphoceramide [Golgi] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide D (C24) [Golgi] -> inositol phosphomannosylinositol phosphoceramide [Golgi] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide D (C26) [Golgi] -> inositol phosphomannosylinositol phosphoceramide [Golgi] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide A (C24) [endoplasmic reticulum] -> inositol phosphomannosylinositol phosphoceramide [endoplasmic reticulum] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide A (C26) [endoplasmic reticulum] -> inositol phosphomannosylinositol phosphoceramide [endoplasmic reticulum] '	''		modelling reaction

'inositol phosphomannosylinositol phosphoceramide B'' (C24) [endoplasmic reticulum] -> inositol phosphomannosylinositol phosphoceramide [endoplasmic reticulum] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide B'' (C26) [endoplasmic reticulum] -> inositol phosphomannosylinositol phosphoceramide [endoplasmic reticulum] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide B (C24) [endoplasmic reticulum] -> inositol phosphomannosylinositol phosphoceramide [endoplasmic reticulum] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide B (C26) [endoplasmic reticulum] -> inositol phosphomannosylinositol phosphoceramide [endoplasmic reticulum] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide C (C24) [endoplasmic reticulum] -> inositol phosphomannosylinositol phosphoceramide [endoplasmic reticulum] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide C (C26) [endoplasmic reticulum] -> inositol phosphomannosylinositol phosphoceramide	''		modelling reaction

[endoplasmic reticulum] '			
'inositol phosphomannosylinositol phosphoceramide D (C24) [endoplasmic reticulum] -> inositol phosphomannosylinositol phosphoceramide [endoplasmic reticulum] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide D (C26) [endoplasmic reticulum] -> inositol phosphomannosylinositol phosphoceramide [endoplasmic reticulum] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide A (C24) [mitochondrion] -> inositol phosphomannosylinositol phosphoceramide [mitochondrion] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide A (C26) [mitochondrion] -> inositol phosphomannosylinositol phosphoceramide [mitochondrion] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide B'' (C24) [mitochondrion] -> inositol phosphomannosylinositol phosphoceramide [mitochondrion] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide B'' (C26) [mitochondrion] -> inositol phosphomannosylinositol phosphoceramide	''		modelling reaction

[mitochondrion] '			
'inositol phosphomannosylinositol phosphoceramide B (C24) [mitochondrion] -> inositol phosphomannosylinositol phosphoceramide [mitochondrion] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide B (C26) [mitochondrion] -> inositol phosphomannosylinositol phosphoceramide [mitochondrion] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide C (C24) [mitochondrion] -> inositol phosphomannosylinositol phosphoceramide [mitochondrion] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide C (C26) [mitochondrion] -> inositol phosphomannosylinositol phosphoceramide [mitochondrion] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide D (C24) [mitochondrion] -> inositol phosphomannosylinositol phosphoceramide [mitochondrion] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide D (C26) [mitochondrion] -> inositol phosphomannosylinositol phosphoceramide [mitochondrion] '	''		modelling reaction

'mannosylinositol phosphorylceramide B'' (C24) [Golgi] -> mannosylinositol phosphorylceramide [Golgi] '	''		modelling reaction
'mannosylinositol phosphorylceramide B'' (C26) [Golgi] -> mannosylinositol phosphorylceramide [Golgi] '	''		modelling reaction
'mannosylinositol phosphorylceramide C (C24) [Golgi] -> mannosylinositol phosphorylceramide [Golgi] '	''		modelling reaction
'mannosylinositol phosphorylceramide C (C26) [Golgi] -> mannosylinositol phosphorylceramide [Golgi] '	''		modelling reaction
'mannosylinositol phosphorylceramide A (C24) [Golgi] -> mannosylinositol phosphorylceramide [Golgi] '	''		modelling reaction
'mannosylinositol phosphorylceramide A (C26) [Golgi] -> mannosylinositol phosphorylceramide [Golgi] '	''		modelling reaction
'mannosylinositol phosphorylceramide B (C24) [Golgi] -> mannosylinositol phosphorylceramide [Golgi] '	''		modelling reaction
'mannosylinositol phosphorylceramide B (C26) [Golgi] -> mannosylinositol phosphorylceramide [Golgi] '	''		modelling reaction
'mannosylinositol phosphorylceramide D (C24) [Golgi] -> mannosylinositol phosphorylceramide [Golgi] '	''		modelling reaction
'mannosylinositol phosphorylceramide D (C26) [Golgi] -> mannosylinositol phosphorylceramide [Golgi] '	''		modelling reaction
'mannosylinositol phosphorylceramide B'' (C24) [endoplasmic reticulum] -> mannosylinositol phosphorylceramide [endoplasmic reticulum] '	''		modelling reaction

'mannosylinositol phosphorylceramide B'' (C26) [endoplasmic reticulum] -> mannosylinositol phosphorylceramide [endoplasmic reticulum] '	''		modelling reaction
'mannosylinositol phosphorylceramide C (C24) [endoplasmic reticulum] -> mannosylinositol phosphorylceramide [endoplasmic reticulum] '	''		modelling reaction
'mannosylinositol phosphorylceramide C (C26) [endoplasmic reticulum] -> mannosylinositol phosphorylceramide [endoplasmic reticulum] '	''		modelling reaction
'mannosylinositol phosphorylceramide A (C24) [endoplasmic reticulum] -> mannosylinositol phosphorylceramide [endoplasmic reticulum] '	''		modelling reaction
'mannosylinositol phosphorylceramide A (C26) [endoplasmic reticulum] -> mannosylinositol phosphorylceramide [endoplasmic reticulum] '	''		modelling reaction
'mannosylinositol phosphorylceramide B (C24) [endoplasmic reticulum] -> mannosylinositol phosphorylceramide [endoplasmic reticulum] '	''		modelling reaction
'mannosylinositol phosphorylceramide B (C26) [endoplasmic reticulum] -> mannosylinositol phosphorylceramide [endoplasmic reticulum] '	''		modelling reaction
'mannosylinositol phosphorylceramide D (C24) [endoplasmic reticulum] -> mannosylinositol phosphorylceramide	''		modelling reaction

[endoplasmic reticulum] '			
'mannosylinositol phosphorylceramide D (C26) [endoplasmic reticulum] -> mannosylinositol phosphorylceramide [endoplasmic reticulum] '	''		modelling reaction
'mannosylinositol phosphorylceramide B'' (C24) [mitochondrion] -> mannosylinositol phosphorylceramide [mitochondrion] '	''		modelling reaction
'mannosylinositol phosphorylceramide B'' (C26) [mitochondrion] -> mannosylinositol phosphorylceramide [mitochondrion] '	''		modelling reaction
'mannosylinositol phosphorylceramide C (C24) [mitochondrion] -> mannosylinositol phosphorylceramide [mitochondrion] '	''		modelling reaction
'mannosylinositol phosphorylceramide C (C26) [mitochondrion] -> mannosylinositol phosphorylceramide [mitochondrion] '	''		modelling reaction
'mannosylinositol phosphorylceramide A (C24) [mitochondrion] -> mannosylinositol phosphorylceramide [mitochondrion] '	''		modelling reaction
'mannosylinositol phosphorylceramide A (C26) [mitochondrion] -> mannosylinositol phosphorylceramide [mitochondrion] '	''		modelling reaction

'mannosylinositol phosphorylceramide B (C24) [mitochondrion] -> mannosylinositol phosphorylceramide [mitochondrion] '	''		modelling reaction
'mannosylinositol phosphorylceramide B (C26) [mitochondrion] -> mannosylinositol phosphorylceramide [mitochondrion] '	''		modelling reaction
'mannosylinositol phosphorylceramide D (C24) [mitochondrion] -> mannosylinositol phosphorylceramide [mitochondrion] '	''		modelling reaction
'mannosylinositol phosphorylceramide D (C26) [mitochondrion] -> mannosylinositol phosphorylceramide [mitochondrion] '	''		modelling reaction
'ceramide-1 (C24) [Golgi] <=> ceramide-1 (C24) [endoplasmic reticulum] '	''		modelling reaction
'ceramide-2 (C24) [Golgi] <=> ceramide-2 (C24) [endoplasmic reticulum] '	''		modelling reaction
'ceramide-4 (C24) [Golgi] <=> ceramide-4 (C24) [endoplasmic reticulum] '	''		modelling reaction
'ceramide-1 (C26) [Golgi] <=> ceramide-1 (C26) [endoplasmic reticulum] '	''		modelling reaction
'ceramide-2 (C26) [Golgi] <=> ceramide-2 (C26) [endoplasmic reticulum] '	''		modelling reaction
'ceramide-4 (C26) [Golgi] <=> ceramide-4 (C26) [endoplasmic reticulum] '	''		modelling reaction
'inositol-P-ceramide A (C24) [Golgi] <=> inositol-P-ceramide A (C24) [endoplasmic reticulum] '	''		modelling reaction

'inositol-P-ceramide B (C24) [Golgi] <=> inositol-P-ceramide B (C24) [endoplasmic reticulum] '	''		modelling reaction
'inositol-P-ceramide B'' (C24) [Golgi] <=> inositol-P-ceramide B'' (C24) [endoplasmic reticulum] '	''		modelling reaction
'inositol-P-ceramide C (C24) [Golgi] <=> inositol-P-ceramide C (C24) [endoplasmic reticulum] '	''		modelling reaction
'inositol-P-ceramide D (C24) [Golgi] <=> inositol-P-ceramide D (C24) [endoplasmic reticulum] '	''		modelling reaction
'inositol-P-ceramide A (C26) [Golgi] <=> inositol-P-ceramide A (C26) [endoplasmic reticulum] '	''		modelling reaction
'inositol-P-ceramide B (C26) [Golgi] <=> inositol-P-ceramide B (C26) [endoplasmic reticulum] '	''		modelling reaction
'inositol-P-ceramide B'' (C26) [Golgi] <=> inositol-P-ceramide B'' (C26) [endoplasmic reticulum] '	''		modelling reaction
'inositol-P-ceramide C (C26) [Golgi] <=> inositol-P-ceramide C (C26) [endoplasmic reticulum] '	''		modelling reaction
'inositol-P-ceramide D (C26) [Golgi] <=> inositol-P-ceramide D (C26) [endoplasmic reticulum] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide A (C24) [Golgi] <=> inositol phosphomannosylinositol phosphoceramide A (C24) [endoplasmic reticulum] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide B (C24) [Golgi] <=> inositol	''		modelling reaction

phosphomannosylinositol phosphoceramide B (C24) [endoplasmic reticulum] '			
'inositol phosphomannosylinositol phosphoceramide B'' (C24) [Golgi] <=> inositol phosphomannosylinositol phosphoceramide B'' (C24) [endoplasmic reticulum] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide C (C24) [Golgi] <=> inositol phosphomannosylinositol phosphoceramide C (C24) [endoplasmic reticulum] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide D (C24) [Golgi] <=> inositol phosphomannosylinositol phosphoceramide D (C24) [endoplasmic reticulum] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide A (C26) [Golgi] <=> inositol phosphomannosylinositol phosphoceramide A (C26) [endoplasmic reticulum] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide B (C26) [Golgi] <=> inositol phosphomannosylinositol phosphoceramide B (C26) [endoplasmic reticulum] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide B'' (C26) [Golgi] <=> inositol phosphomannosylinositol phosphoceramide B'' (C26) [endoplasmic reticulum] '	''		modelling reaction

'inositol phosphomannosylinositol phosphoceramide C (C26) [Golgi] <=> inositol phosphomannosylinositol phosphoceramide C (C26) [endoplasmic reticulum] '	''		modelling reaction
'inositol phosphomannosylinositol phosphoceramide D (C26) [Golgi] <=> inositol phosphomannosylinositol phosphoceramide D (C26) [endoplasmic reticulum] '	''		modelling reaction
'mannosylinositol phosphorylceramide A (C24) [Golgi] <=> mannosylinositol phosphorylceramide A (C24) [endoplasmic reticulum] '	''		modelling reaction
'mannosylinositol phosphorylceramide B (C24) [Golgi] <=> mannosylinositol phosphorylceramide B (C24) [endoplasmic reticulum] '	''		modelling reaction
'mannosylinositol phosphorylceramide B'' (C24) [Golgi] <=> mannosylinositol phosphorylceramide B'' (C24) [endoplasmic reticulum] '	''		modelling reaction
'mannosylinositol phosphorylceramide C (C24) [Golgi] <=> mannosylinositol phosphorylceramide C (C24) [endoplasmic reticulum] '	''		modelling reaction
'mannosylinositol phosphorylceramide D (C24) [Golgi] <=> mannosylinositol phosphorylceramide D (C24) [endoplasmic reticulum] '	''		modelling reaction
'mannosylinositol phosphorylceramide A (C26) [Golgi] <=> mannosylinositol phosphorylceramide A (C26) [endoplasmic reticulum] '	''		modelling reaction
'mannosylinositol phosphorylceramide B (C26) [Golgi] <=> mannosylinositol	''		modelling reaction

phosphorylceramide B (C26) [endoplasmic reticulum] '			
'mannosylinositol phosphorylceramide B'' (C26) [Golgi] <=> mannosylinositol phosphorylceramide B'' (C26) [endoplasmic reticulum] '	''		modelling reaction
'mannosylinositol phosphorylceramide C (C26) [Golgi] <=> mannosylinositol phosphorylceramide C (C26) [endoplasmic reticulum] '	''		modelling reaction
'mannosylinositol phosphorylceramide D (C26) [Golgi] <=> mannosylinositol phosphorylceramide D (C26) [endoplasmic reticulum] '	''		modelling reaction

Supplementary Table 3.2

The literature-based list of genes involved in sphingolipid metabolism, with the number of reactions annotated with each gene in the model and literature reaction list.

Gene	Locus	# of rxns in YN 4.05	# of rxns in lit review list	comments
LCB1	YMR296C	1	1	component of serine palmitoyltransferase with Lcb2p
LCB2	YDR062W	1	1	component of serine palmitoyltransferase with Lcb1p
TSC3	YBR058C-A	0	1	activator of LCB1:LCB2 complex, not essential
TSC10	YBR265W	1	1	3-ketosphinganine reductase
LAG1	YHL003C	4	4	lag1 & lac1 are functionally equivalent, and associated with lip1
LAC1	YKL008C	4	4	lag1 & lac1 are functionally equivalent, and associated with lip1
LIP1	YMR298W	0	4	lag1 & lac1 are functionally equivalent, and associated with lip1
SUR2	YDR297W	5	3	Sphinganine C4-hydroxylase
LCB4	YOR171C	2	2	Sphingolipid long chain base kinase (sphingosine and phytosphingosine)
LCB5	YLR260W	2	2	Catalyzes 2 reactions, same as LCB4
LCB3 (aka LBP1, YSR2)	YJL134W	2	2	may catalyze 2 reactions - DHSP and PHSP
YSR3	YKR053C	2	2	Like LCB3; catalyze 2 reactions
AUR1	YKL004W	4	10	Catalyzes the addition of a phosphorylinositol group onto ceramide to form inositol phosphorylceramide
CSG1 (aka SUR1)	YPL057C	2	10	complex with CSG2; CSG1/SUR1 is the catalytic subunit
CSG2	YBR036C	8	10	CSG2 is the likely regulatory subunit
CSH1	YBR161W	0	10	Overlapping function with YPL057C:YBR036C complex
IPT1	YDR072C	1	10	Inositolphosphotransferase
SKN1	YGR143W	1	1	This seems fairly hypothetical by Dickson, but would be a duplicate of IPT1, (r_864) above.
ISC1	YER019W	2	60	hydrolyzes complex sphingolipids
DPL1	YDR294C	2	2	Phytosphingosine phosphate lyase
SCS7	YMR272C	4	6	IPC-B -> IPC-C

YDC1	YPL087W	4	4	preferentially hydrolyzes dihydroceramide to a free fatty acid and dihydrosphingosine
YPC1	YBR183W	2	2	ceramides can be hydrolyzed by two ceramidases, YDC1 and YPC1, to yield a fatty acid and a long-chain base
total # rxns		54	152	the sum of reactions annotated by a gene is not equal to the sum of reactions in the pathway because some reactions are annotated with multiple genes

Supplementary Table 3.3

The list of reactions related to sphingolipid metabolism in Yeast v4.05.

rxn index	rxn id	rxn	rxn with names	gene locus	Gene	Note
7	'r_0007'	's_1415 <=> s_0004 + s_1411 '	'UDP-D-glucose [cytoplasm] <=> (1->6)-beta-D-glucan [cell envelope] + UDP [cytoplasm] '	'(YGR143W or YPR159W)'	SKN1	mis-annotated
44	'r_0044'	's_0218 + s_1097 <=> s_1092 + s_1325 '	'3-dehydrosphinganine [endoplasmic reticulum] + NADPH [endoplasmic reticulum] <=> NADP(+) [endoplasmic reticulum] + sphinganine [endoplasmic reticulum] '	'YBR265W'	TSC10	
186	'r_0186'	's_0174 + s_1096 <=> s_0171 + s_1091 '	'2-methylbutanal [cytoplasm] + NADPH [cytoplasm] <=> 2-methylbutan-1-ol [cytoplasm] + NADP(+) [cytoplasm] '	'(YBR036C or YCR105W or YDR368W or YMR318C)'	CSG2	mis-annotated
189	'r_0189'	's_1096 + s_1203 <=> s_0190 + s_1091 '	'NADPH [cytoplasm] + phenylacetaldehyde [cytoplasm] <=> 2-phenylethanol [cytoplasm] + NADP(+) [cytoplasm] '	'(YBR036C or YCR105W or YMR318C)'	CSG2	mis-annotated
199	'r_0199'	's_0240 + s_1096 <=> s_0836 + s_1091 '	'3-methylbutanal [cytoplasm] + NADPH [cytoplasm] <=> isoamylol [cytoplasm] + NADP(+) [cytoplasm] '	'(YBR036C or YCR105W or YMR318C)'	CSG2	mis-annotated
202	'r_0202'	's_0844 + s_1096 <=> s_0839 + s_1091 '	'isobutyraldehyde [cytoplasm] + NADPH [cytoplasm] <=> isobutanol [cytoplasm] + NADP(+) [cytoplasm] '	'(YBR036C or YCR105W or YMR318C)'	CSG2	mis-annotated
286	'r_0286'	's_1080 + s_1097 + s_1161 <=> s_1079 + s_1092 '	'N-tetracosanylsphinganine [endoplasmic reticulum] + NADPH [endoplasmic reticulum] + oxygen [endoplasmic reticulum] <=> N-tetracosanylphytosphingosine [endoplasmic reticulum] + NADP(+) [endoplasmic reticulum] '	YDR297W'	SUR2	

287	'r_0287'	's_1080 + s_1097 + s_1161 <=> s_1060 + s_1092 '	'N-tetracosanylsphinganine [endoplasmic reticulum] + NADPH [endoplasmic reticulum] + oxygen [endoplasmic reticulum] <=> N-(24-hydroxytetracosanyl)sphinganine [endoplasmic reticulum] + NADP(+) [endoplasmic reticulum] '	'YMR272C'	SCS7	
288	'r_0288'	's_1077 + s_1097 + s_1161 <=> s_1076 + s_1092 '	'N-hexacosanylsphinganine [endoplasmic reticulum] + NADPH [endoplasmic reticulum] + oxygen [endoplasmic reticulum] <=> N-hexacosanylphytosphingosine [endoplasmic reticulum] + NADP(+) [endoplasmic reticulum] '	'YDR297W'	SUR2	
289	'r_0289'	's_1077 + s_1097 + s_1161 <=> s_1064 + s_1092 '	'N-hexacosanylsphinganine [endoplasmic reticulum] + NADPH [endoplasmic reticulum] + oxygen [endoplasmic reticulum] <=> N-(26-hydroxyhexacosanyl)sphinganine [endoplasmic reticulum] + NADP(+) [endoplasmic reticulum] '	'YMR272C'	SCS7	
290	'r_0290'	's_1325 + s_1356 <=> s_0515 + s_1080 '	'sphinganine [endoplasmic reticulum] + tetracosanoyl-CoA [endoplasmic reticulum] <=> coenzyme A [endoplasmic reticulum] + N-tetracosanylsphinganine [endoplasmic reticulum] '	'(YHL003C or YKL008C or YPL087W)'	LAG1 or LAC1 or YDC1	
291	'r_0291'	's_0779 + s_1325 <=> s_0515 + s_1077 '	'hexacosanoyl-CoA [endoplasmic reticulum] + sphinganine [endoplasmic reticulum] <=> coenzyme A [endoplasmic reticulum] + N-hexacosanylsphinganine [endoplasmic reticulum] '	'(YHL003C or YKL008C or YPL087W)'	LAG1 or LAC1 or YDC1	
292	'r_0292'	's_1247 + s_1356 <=> s_0515 + s_1079 '	'phytosphingosine [endoplasmic reticulum] + tetracosanoyl-CoA [endoplasmic reticulum] <=> coenzyme A [endoplasmic reticulum] + N-tetracosanylphytosphingosine [endoplasmic reticulum] '	'(YBR183W or YHL003C or YKL008C or YPL087W)'	YPC1 or LAG1 or LAC1 or YDC1	

293	'r_0293'	's_0779 + s_1247 <=> s_0515 + s_1076 '	'hexacosanoyl-CoA [endoplasmic reticulum] + phytosphingosine [endoplasmic reticulum] <=> coenzyme A [endoplasmic reticulum] + N-hexacosanylphytosphingosine [endoplasmic reticulum] '	'(YBR183W or YHL003C or YKL008C or YPL087W)'	YPC1 or LAG1 or LAC1 or YDC1	
294	'r_0294'	's_1079 + s_1097 + s_1161 <=> s_1058 + s_1092 '	'N-tetracosanylphytosphingosine [endoplasmic reticulum] + NADPH [endoplasmic reticulum] + oxygen [endoplasmic reticulum] <=> N-(24-hydroxytetracosanyl)phytosphingosine [endoplasmic reticulum] + NADP(+) [endoplasmic reticulum] '	'YMR272C'	SCS7	
295	'r_0295'	's_1076 + s_1097 + s_1161 <=> s_1062 + s_1092 '	'N-hexacosanylphytosphingosine [endoplasmic reticulum] + NADPH [endoplasmic reticulum] + oxygen [endoplasmic reticulum] <=> N-(26-hydroxyhexacosanyl)phytosphingosine [endoplasmic reticulum] + NADP(+) [endoplasmic reticulum] '	'YMR272C'	SCS7	
618	'r_0618'	's_0129 + s_1013 <=> s_0825 '	'1D-myo-inositol 1-phosphate [Golgi] + mannosylinositol phosphorylceramide [Golgi] <=> inositol phosphomannosylinositol phosphoceramide [Golgi] '	'YDR072C'	IPT1	
621	'r_0621'	's_0129 + s_1061 <=> s_0828 '	'1D-myo-inositol 1-phosphate [Golgi] + N-(24-hydroxytetracosanyl)sphinganine [Golgi] <=> inositol-P-ceramide B [Golgi] '	'YKL004W'	AUR1	
622	'r_0622'	's_0129 + s_1065 <=> s_0828 '	'1D-myo-inositol 1-phosphate [Golgi] + N-(26-hydroxyhexacosanyl)sphinganine [Golgi] <=> inositol-P-ceramide B [Golgi] '	'YKL004W'	AUR1	
623	'r_0623'	's_0129 + s_1059 <=> s_0829 '	'1D-myo-inositol 1-phosphate [Golgi] + N-(24-hydroxytetracosanyl)phytosphingosine [Golgi] <=> inositol-P-ceramide C [Golgi] '	'YKL004W'	AUR1	
624	'r_0624'	's_0129 + s_1063 <=> s_0829 '	'1D-myo-inositol 1-phosphate [Golgi] + N-(26-hydroxyhexacosanyl)phytosphingosine [Golgi] <=> inositol-P-	'YKL004W'	AUR1	

			ceramide C [Golgi] '			
625	'r_0625'	's_0826 <=> s_0130 + s_0489 '	'inositol phosphosphingolipid [endoplasmic reticulum] <=> 1D-myo-inositol 3-phosphate [endoplasmic reticulum] + ceramide [endoplasmic reticulum] '	'YER019W'	ISC1	
626	'r_0626'	's_0827 <=> s_0131 + s_0490 '	'inositol phosphosphingolipid [mitochondrion] <=> 1D-myo-inositol 3-phosphate [mitochondrion] + ceramide [mitochondrion] '	'YER019W'	ISC1	
723	'r_0723'	's_0711 + s_0828 <=> s_1013 '	'GDP-alpha-D-mannose [Golgi] + inositol-P-ceramide B [Golgi] <=> mannosylinositol phosphorylceramide [Golgi] '	'(YBR036C and YPL057C)'	CSG2 and CSG1	
724	'r_0724'	's_0711 + s_0829 <=> s_1013 '	'GDP-alpha-D-mannose [Golgi] + inositol-P-ceramide C [Golgi] <=> mannosylinositol phosphorylceramide [Golgi] '	'(YBR036C and YDR297W)'	CSG2 and SUR2	
894	'r_0894'	's_1248 <=> s_0166 + s_1125 '	'phytosphingosine 1-phosphate [endoplasmic reticulum] <=> 2-hydroxyhexadecanal [endoplasmic reticulum] + O-phosphoethanolamine [endoplasmic reticulum] '	'YDR294C'	DPL1	
895	'r_0895'	's_1097 + s_1161 + s_1325 <=> s_1092 + s_1247 '	'NADPH [endoplasmic reticulum] + oxygen [endoplasmic reticulum] + sphinganine [endoplasmic reticulum] <=> NADP(+) [endoplasmic reticulum] + phytosphingosine [endoplasmic reticulum] '	'YDR297W'	SUR2	
972	'r_0972'	's_0944 + s_1188 <=> s_0218 + s_0471 + s_0515 '	'L-serine [endoplasmic reticulum] + palmitoyl-CoA [endoplasmic reticulum] <=> 3-dehydrosphinganine [endoplasmic reticulum] + carbon dioxide [endoplasmic reticulum] + coenzyme A [endoplasmic reticulum] '	'(YDR062W or YMR296C)'	LCB2 or LCB1	
984	'r_0984'	's_1326 <=> s_0789 + s_1125 '	'sphinganine 1-phosphate [endoplasmic reticulum] <=> hexadecanal [endoplasmic reticulum] + O-phosphoethanolamine [endoplasmic reticulum] '	'YDR294C'	DPL1	

985	'r_0985'	's_1248 <=> s_1208 + s_1247 '	'phytosphingosine 1-phosphate [endoplasmic reticulum] <=> phosphate [endoplasmic reticulum] + phytosphingosine [endoplasmic reticulum] '	'(YJL134W or YKR053C)'	LCB3 or YSR3	
986	'r_0986'	's_1326 <=> s_1208 + s_1325 '	'sphinganine 1-phosphate [endoplasmic reticulum] <=> phosphate [endoplasmic reticulum] + sphinganine [endoplasmic reticulum] '	'(YJL134W or YKR053C)'	LCB3 or YSR3	
987	'r_0987'	's_0447 + s_1247 <=> s_0401 + s_1248 '	'ATP [endoplasmic reticulum] + phytosphingosine [endoplasmic reticulum] <=> ADP [endoplasmic reticulum] + phytosphingosine 1-phosphate [endoplasmic reticulum] '	'(YLR260W or YOR171C)'	LCB5 or LCB4	
988	'r_0988'	's_0447 + s_1325 <=> s_0401 + s_1326 '	'ATP [endoplasmic reticulum] + sphinganine [endoplasmic reticulum] <=> ADP [endoplasmic reticulum] + sphinganine 1-phosphate [endoplasmic reticulum] '	'(YLR260W or YOR171C)'	LCB5 or LCB4	
1175	'r_1175'	's_1060 <=> s_1061 '	'N-(24-hydroxytetracosanyl)sphinganine [endoplasmic reticulum] <=> N-(24-hydroxytetracosanyl)sphinganine [Golgi] '	''		
1176	'r_1176'	's_1064 <=> s_1065 '	'N-(26-hydroxyhexacosanyl)sphinganine [endoplasmic reticulum] <=> N-(26-hydroxyhexacosanyl)sphinganine [Golgi] '	''		
1177	'r_1177'	's_1058 <=> s_1059 '	'N-(24-hydroxytetracosanyl)phytosphingosine [endoplasmic reticulum] <=> N-(24-hydroxytetracosanyl)phytosphingosine [Golgi] '	''		
1178	'r_1178'	's_1062 <=> s_1063 '	'N-(26-hydroxyhexacosanyl)phytosphingosine [endoplasmic reticulum] <=> N-(26-hydroxyhexacosanyl)phytosphingosine [Golgi] '	''		

1334	'r_1334'	's_0824 <=> s_0825 '	'inositol phosphomannosylinositol phosphoceramide [cytoplasm] <=> inositol phosphomannosylinositol phosphoceramide [Golgi] '	''		
1553	'r_1553'	' <=> '	' <=> '	'(YBR036C and YDR297W)'	CSG2	
1554	'r_1554'	' <=> '	' <=> '	'(YBR036C and YPL057C)'	CSG2	
1796	'r_1796'	's_0789 -> s_0665 '	'hexadecanal [endoplasmic reticulum] -> fatty acid [endoplasmic reticulum] '	''		

Supplementary File 4.1: testYeastModel.m MATLAB script

```
function testYeastModel(filename,verbose)

% a formalised test-suite for validating yeast models. reliant on the cobra
% toolbox.

% kieran smallbone and ben heavner: 15 aug 11

%%

if nargin < 2
    verbose = false; % should the script display model errors?
end

ko_tol = 1e-6;

model = readCbModel(filename);

%% create model with complete medium (all exchange reactions switched on)

selExc = findExcRxns(model);

model_maximal = model;
model_maximal.lb(selExc) = -1000;
model_maximal.ub(selExc) = 1000;

%% simple model statistics

fprintf('%s\n\n%g\t%s\n%g\t%s\n%g\t%s\n\n',model.description,...
    length(model.mets),'metabolites',length(model.rxns),'reactions',...
    length(model.genes),'genes');

%% number of blocked reactions (cannot carry any flux)

% blockedReactions = findBlockedReaction(model);
blockedReactions = findBlockedReaction(model_maximal);

fprintf('%g\t(%g%%)\t%s\n\n',length(blockedReactions),...
    roundSF(100*length(blockedReactions)/length(model.rxns)),'blocked
reactions');

if verbose
    disp('list of blocked reactions:');
    for k = 1:length(blockedReactions)
        disp(blockedReactions{k});
    end
    disp('');
end
```

```
%% dubious genes

dubiousORFs = setdiff(model.genes,verifiedORFs);

fprintf('%g\t(%g%%)\t%s\n\n',length(dubiousORFs),...
    roundSF(100*length(dubiousORFs)/length(model.genes)),'dubious ORFs');

if verbose
    disp('list of dubious ORFs:');
    for k = 1:length(dubiousORFs)
        disp(dubiousORFs{k});
    end
    disp('');
end

%% knockout analysis

% combine inviable with auxotrophy list
%inviableORFsAll = union(inviableORFs,auxotrophicORFs); %auxotrophs should
%be viable in complex media
inviableORFsAll = inviableORFs; %only essential genes

exp_retarded = intersect(model.genes,inviableORFsAll);
exp_retarded = intersect(exp_retarded,verifiedORFs);

exp_viable = setdiff(model.genes,inviableORFsAll);
exp_viable = intersect(exp_viable,verifiedORFs);

%grRatio = singleGeneDeletion(model);
grRatio = singleGeneDeletion(model_maximal); %bh changed to use same
essential def.

mod_viable  = model.genes(grRatio >= ko_tol);
mod_viable = intersect(mod_viable,verifiedORFs);
mod_retarded = model.genes(grRatio < ko_tol);
mod_retarded = intersect(mod_retarded,verifiedORFs);

tp = intersect(exp_viable,mod_viable); n_tp = length(tp)
tn = intersect(exp_retarded,mod_retarded); n_tn = length(tn)
fp = intersect(exp_retarded,mod_viable); n_fp = length(fp)
fn = intersect(exp_viable,mod_retarded); n_fn = length(fn)

n_genes = length(intersect(model.genes,verifiedORFs)); %#ok<NASGU>

disp('knockout analysis (positive = viable):');

% fprintf('%g\t[%.3g %%]\t%s\n%g\t[%.3g %%]\t%s\n%g\t[%.3g %%]\t%s\n%g\t[%.3g
%%]\t%s\n\n',...
%     n_tp,roundSF(100*n_tp/n_genes),'true positive',...
%     n_tn,roundSF(100*n_tn/n_genes),'true negative',...
%     n_fp,roundSF(100*n_fp/n_genes),'false positive',...
%     n_fn,roundSF(100*n_fn/n_genes),'false negative');

fprintf('%g%%\t%s\n%g%%\t%s\n%g%%\t%s\n%g%%\t%s\n%',...
    roundSF(100*n_tp/(n_tp+n_fn)),'sensitivity = recall = tp/(tp+fn)',...
    roundSF(100*n_tn/(n_tn+n_fp)),'specificity = tn/(tn+fp)',...
    roundSF(100*n_tp/(n_tp+n_fp)),'positive predictive value = precision =
```

```
tp/(tp+fp)',...
    roundSF(100*n_tn/(n_fn+n_tn)),'negative predictive value = tn/(fn+tn)'...
    );

% fprintf('g%%\t%s\n%g%%\t%s\n',...
%     roundSF(100*(n_tn+n_tp)/n_genes),'accuracy = (tp+tn)/(tp+tn+fp+fn)',...
%     roundSF(100*2*(n_tp/(n_tp+n_fn) * n_tp/(n_tp+n_fp))/(n_tp/(n_tp+n_fn) +
n_tp/(n_tp+n_fp))),'F-measure = 2.(precision.recall)/(precision+recall)'...
%     );

disp(' ');

if verbose
    disp('list of ko false positives:');
    for k = 1:n_fp
        disp(fp{k});
    end
    disp(' ');

    disp('list of ko false negatives:');
    for k = 1:n_fn
        disp(fn{k});
    end
    disp(' ');
end

%% testing auxotrophy

% or at least an approximation thereof

auxotrophs = intersect(model.genes,auxotrophicORFs);
auxotrophs = intersect(auxotrophs,verifiedORFs);
n_aux = length(auxotrophs);

grRatio = singleGeneDeletion(model_maximal);

aux_fail = intersect(auxotrophs,mod_viable);

aux_viable  = model.genes(grRatio >= ko_tol);
aux_viable = intersect(aux_viable,auxotrophs);
aux_viable = setdiff(aux_viable,aux_fail);
aux_retarded = model.genes(grRatio < ko_tol);
aux_retarded = intersect(aux_retarded,auxotrophs);
aux_retarded = setdiff(aux_retarded,aux_fail);

fprintf('auxotrophy analysis:\n%g\t%s\n%g\t[%.3g %%]\t%s\n%g\t[%.3g
%%]\t%s\n%g\t[%.3g %%]\t%s\n\n',...
    n_aux,'genes',...
    length(aux_viable),roundSF(100*length(aux_viable)/n_aux),...
    'correctly predicted',...
    length(aux_fail),roundSF(100*length(aux_fail)/n_aux),...
    'viable in minimal medium',...
    length(aux_retarded),roundSF(100*length(aux_retarded)/n_aux),...
    'inviable in maximal medium');

if verbose
    disp('list of auxotrophs viable in minimal medium:');
```

```
    for k = 1:length(aux_fail)
        disp(aux_fail{k});
    end
    disp(' ');

    disp('list of auxotrophs inviable in maximal medium:');
    for k = 1:length(aux_retarded)
        disp(aux_retarded{k});
    end
    disp(' ');
end

disp('==')
disp(' ');

%% required functions

function blockedReactions = findBlockedReaction(model)
%findBlockedReaction determines those reactions which cannot carry any
%flux in the given simulation conditions.
%
% BlockedReaction = findBlockedReaction(model)
%
%INPUT
% model              COBRA model structure
%
%OUTPUT
% blockedReactions   List of blocked reactions
%
%
% Ines Thiele 02/09

% kieran: 15 aug 11

tol = 1e-10;
% << changed by ks
% blockedReactions =[];
blockedReactions = {};
% >>
[minMax(:,1),minMax(:,2)] = fluxVariability(model,0);
cnt = 1;
for i=1:length(minMax)
    if (minMax(i,2) < tol && minMax(i,2) > -tol && minMax(i,1) < tol &&
minMax(i,1) > -tol)
        blockedReactions(cnt) = model.rxns(i); %#ok<AGROW>
        cnt = cnt + 1;
    end
end

function y = roundSF(x,nSF)

%ROUNDSF
% y = roundSF(x,nSF) returns x rounded to nSF significant figures
% input x must be a real array
%
% kieran: 7 nov 08
% kieran: 15 aug 11
```

```
if nargin < 2, nSF = 3; end

y = zeros(size(x));

for k = 1:numel(x)

    x0 = x(k);

    if ~isfinite(x0) || x0 == 0
        y0 = x0;
    else
        N = nSF-floor(log10(abs(x0)))-1;
        y0 = round(x0.*10.^N)./10.^N;
    end

    y(k) = y0;
end

function genes = inviableORFs

% list of 1191 unique inviable ORFs taken from the Yeast Deletion Project (14
aug
% 11)
% http://www-
sequence.stanford.edu/group/yeast_deletion_project/downloads.html

genes = {'YAL001C';'YAL003W';'YAL025C';'YAL032C';'YAL033W';'YAL034W-
a';'YAL035C-
A';'YAL038W';'YAL041W';'YAL043C';'YAR007C';'YAR008W';'YAR019C';'YBL004W';'YBL
014C';'YBL018C';'YBL020W';'YBL023C';'YBL026W';'YBL030C';'YBL034C';'YBL035C';'
YBL040C';'YBL041W';'YBL050W';'YBL073W';'YBL074C';'YBL076C';'YBL077W';'YBL084C
';'YBL092W';'YBL097W';'YBL105C';'YBR002C';'YBR004C';'YBR011C';'YBR029C';'YBR0
38W';'YBR049C';'YBR055C';'YBR060C';'YBR070C';'YBR079C';'YBR080C';'YBR087W';'Y
BR087W';'YBR088C';'YBR089W';'YBR091C';'YBR102C';'YBR109C';'YBR110W';'YBR123C'
;'YBR124W';'YBR135W';'YBR136W';'YBR140C';'YBR142W';'YBR143C';'YBR152W';'YBR15
3W';'YBR154C';'YBR155W';'YBR160W';'YBR167C';'YBR190W';'YBR192W';'YBR193C';'YB
R196C';'YBR198C';'YBR202W';'YBR211C';'YBR233W-A';'YBR233W-
A';'YBR234C';'YBR236C';'YBR237W';'YBR243C';'YBR247C';'YBR252W';'YBR253W';'YBR
254C';'YBR256C';'YBR257W';'YBR265W';'YCL003W';'YCL004W';'YCL017C';'YCL031C';'
YCL031C';'YCL031C';'YCL043C';'YCL052C';'YCL053C';'YCL054W';'YCL059C';'YCR012W
';'YCR012W';'YCR013C';'YCR013C';'YCR035C';'YCR052W';'YCR054C';'YCR057C';'YCR0
72C';'YCR093W';'YDL003W';'YDL004W';'YDL007W';'YDL008W';'YDL014W';'YDL015C';'Y
DL016C';'YDL017W';'YDL028C';'YDL029W';'YDL030W';'YDL031W';'YDL043C';'YDL045C'
;'YDL055C';'YDL058W';'YDL060W';'YDL064W';'YDL084W';'YDL087C';'YDL092W';'YDL09
7C';'YDL098C';'YDL102W';'YDL103C';'YDL105W';'YDL108W';'YDL111C';'YDL120W';'YD
L126C';'YDL132W';'YDL139C';'YDL140C';'YDL141W';'YDL143W';'YDL145C';'YDL147W';
'YDL148C';'YDL150W';'YDL152W';'YDL153C';'YDL163W';'YDL164C';'YDL165W';'YDL166
C';'YDL193W';'YDL195W';'YDL196W';'YDL205C';'YDL207W';'YDL208W';'YDL209C';'YDL
212W';'YDL217C';'YDL220C';'YDL221W';'YDL235C';'YDR002W';'YDR013W';'YDR016C';'
YDR021W';'YDR023W';'YDR037W';'YDR041W';'YDR044W';'YDR045C';'YDR047W';'YDR050C
';'YDR050C';'YDR052C';'YDR053W';'YDR054C';'YDR060W';'YDR062W';'YDR064W';'YDR0
81C';'YDR082W';'YDR086C';'YDR087C';'YDR088C';'YDR091C';'YDR113C';'YDR118W';'Y
DR141C';'YDR145W';'YDR160W';'YDR164C';'YDR166C';'YDR167W';'YDR168W';'YDR170C'
;'YDR172W';'YDR177W';'YDR180W';'YDR182W';'YDR187C';'YDR188W';'YDR189W';'YDR19
0C';'YDR196C';'YDR201W';'YDR208W';'YDR211W';'YDR212W';'YDR224C';'YDR224C';'YD
R224C';'YDR228C';'YDR232W';'YDR235W';'YDR236C';'YDR238C';'YDR240C';'YDR243C';
```

'YDR246W';'YDR267C';'YDR280W';'YDR288W';'YDR292C';'YDR299W';'YDR301W';'YDR302
W';'YDR303C';'YDR308C';'YDR311W';'YDR320C-A';'YDR320C-
A';'YDR324C';'YDR325W';'YDR327W';'YDR328C';'YDR331W';'YDR339C';'YDR341C';'YDR
353W';'YDR355C';'YDR356W';'YDR361C';'YDR362C';'YDR365C';'YDR367W';'YDR373W';'
YDR376W';'YDR381W';'YDR390C';'YDR394W';'YDR396W';'YDR397C';'YDR398W';'YDR404C
';'YDR407C';'YDR412W';'YDR413C';'YDR416W';'YDR427W';'YDR427W';'YDR429C';'YDR4
34W';'YDR437W';'YDR449C';'YDR454C';'YDR460W';'YDR464W';'YDR468C';'YDR472W';'Y
DR473C';'YDR478W';'YDR487C';'YDR489W';'YDR498C';'YDR499W';'YDR510W';'YDR526C'
;'YDR527W';'YDR531W';'YEL002C';'YEL019C';'YEL026W';'YEL032W';'YEL034W';'YEL03
5C';'YEL055C';'YEL058W';'YER003C';'YER006W';'YER008C';'YER009W';'YER012W';'YE
R013W';'YER018C';'YER021W';'YER022W';'YER023W';'YER025W';'YER029C';'YER029C';
'YER036C';'YER038C';'YER043C';'YER048W-A';'YER074W-A';'YER074W-
A';'YER082C';'YER093C';'YER094C';'YER104W';'YER112W';'YER125W';'YER126C';'YER
127W';'YER133W';'YER136W';'YER146W';'YER147C';'YER148W';'YER157W';'YER159C';'
YER165W';'YER168C';'YER171W';'YER172C';'YFL002C';'YFL005W';'YFL008W';'YFL009W
';'YFL017C';'YFL018W-A';'YFL022C';'YFL024C';'YFL029C';'YFL035C';'YFL035C-
A';'YFL037W';'YFL038C';'YFL039C';'YFL045C';'YFR002W';'YFR003C';'YFR004W';'YFR
005C';'YFR027W';'YFR028C';'YFR029W';'YFR031C';'YFR037C';'YFR042W';'YFR050C';'
YFR051C';'YFR052W';'YGL001C';'YGL008C';'YGL011C';'YGL018C';'YGL022W';'YGL030W
';'YGL040C';'YGL044C';'YGL044C';'YGL044C';'YGL047W';'YGL048C';'YGL055W';'YGL0
61C';'YGL065C';'YGL068W';'YGL069C';'YGL073W';'YGL074C';'YGL075C';'YGL091C';'Y
GL092W';'YGL093W';'YGL097W';'YGL097W';'YGL098W';'YGL099W';'YGL102C';'YGL103W'
;'YGL111W';'YGL112C';'YGL113W';'YGL116W';'YGL120C';'YGL122C';'YGL123W';'YGL12
8C';'YGL130W';'YGL137W';'YGL142C';'YGL145W';'YGL145W';'YGL145W';'YGL150C';'YG
L155W';'YGL169W';'YGL171W';'YGL172W';'YGL201C';'YGL207W';'YGL225W';'YGL233W';
'YGL238W';'YGL239C';'YGL239C';'YGL245W';'YGL247W';'YGR002C';'YGR005C';'YGR009
C';'YGR013W';'YGR024C';'YGR029W';'YGR029W';'YGR029W';'YGR030C';'YGR046W';'YGR
047C';'YGR048W';'YGR060W';'YGR065C';'YGR073C';'YGR074W';'YGR075C';'YGR082W';'
YGR083C';'YGR090W';'YGR091W';'YGR094W';'YGR095C';'YGR098C';'YGR099W';'YGR103W
';'YGR113W';'YGR114C';'YGR115C';'YGR116W';'YGR119C';'YGR120C';'YGR128C';'YGR1
40W';'YGR145W';'YGR147C';'YGR156W';'YGR158C';'YGR172C';'YGR175C';'YGR179C';'Y
GR185C';'YGR186W';'YGR190C';'YGR191W';'YGR195W';'YGR198W';'YGR211W';'YGR216C'
;'YGR218W';'YGR245C';'YGR246C';'YGR251W';'YGR253C';'YGR264C';'YGR265W';'YGR26
7C';'YGR274C';'YGR277C';'YGR278W';'YGR280C';'YHL015W';'YHR005C-
A';'YHR007C';'YHR019C';'YHR020W';'YHR023W';'YHR024C';'YHR036W';'YHR040W';'YHR
042W';'YHR058C';'YHR062C';'YHR065C';'YHR068W';'YHR069C';'YHR070W';'YHR072W';'
YHR072W-A';'YHR072W-
A';'YHR074W';'YHR083W';'YHR085W';'YHR088W';'YHR088W';'YHR088W';'YHR089C';'YHR
089C';'YHR101C';'YHR102W';'YHR102W';'YHR107C';'YHR118C';'YHR122W';'YHR128W';'
YHR128W';'YHR143W-
A';'YHR148W';'YHR164C';'YHR165C';'YHR165C';'YHR166C';'YHR169W';'YHR169W';'YHR
170W';'YHR172W';'YHR186C';'YHR188C';'YHR188C';'YHR188C';'YHR190W';'YHR196W';'
YHR197W';'YHR197W';'YHR197W';'YHR199C-A';'YHR199C-
A';'YIL003W';'YIL004C';'YIL019W';'YIL021W';'YIL022W';'YIL026C';'YIL031W';'YIL
046W';'YIL048W';'YIL051C';'YIL061C';'YIL062C';'YIL063C';'YIL068C';'YIL075C';'
YIL078W';'YIL083C';'YIL091C';'YIL104C';'YIL106W';'YIL106W';'YIL106W';'YIL109C
';'YIL115C';'YIL118W';'YIL126W';'YIL129C';'YIL142W';'YIL143C';'YIL144W';'YIL1
47C';'YIL150C';'YIL171W';'YIR006C';'YIR008C';'YIR010W';'YIR011C';'YIR012W';'Y
IR015W';'YIR022W';'YJL001W';'YJL002C';'YJL005W';'YJL008C';'YJL008C';'YJL008C'
;'YJL009W';'YJL010C';'YJL011C';'YJL014W';'YJL015C';'YJL018W';'YJL019W';'YJL02
5W';'YJL026W';'YJL031C';'YJL032W';'YJL033W';'YJL034W';'YJL035C';'YJL039C';'YJ
L041W';'YJL050W';'YJL054W';'YJL061W';'YJL069C';'YJL072C';'YJL074C';'YJL076W';
'YJL081C';'YJL085W';'YJL086C';'YJL087C';'YJL090C';'YJL091C';'YJL097W';'YJL104
W';'YJL109C';'YJL111W';'YJL125C';'YJL143W';'YJL156C';'YJL167W';'YJL173C';'YJL
174W';'YJL194W';'YJL195C';'YJL195C';'YJL202C';'YJL202C';'YJL203W';'YJR002W';'
YJR006W';'YJR007W';'YJR012C';'YJR013W';'YJR016C';'YJR017C';'YJR022W';'YJR023C
';'YJR041C';'YJR042W';'YJR045C';'YJR046W';'YJR046W';'YJR046W';'YJR057W';'YJR0

64W';'YJR065C';'YJR067C';'YJR068W';'YJR072C';'YJR076C';'YJR089W';'YJR089W';'Y
JR093C';'YJR112W';'YJR123W';'YJR141W';'YKL004W';'YKL006C-
A';'YKL012W';'YKL013C';'YKL014C';'YKL018W';'YKL019W';'YKL021C';'YKL022C';'YKL
024C';'YKL028W';'YKL033W';'YKL035W';'YKL036C';'YKL042W';'YKL045W';'YKL049C';'
YKL049C';'YKL049C';'YKL052C';'YKL058W';'YKL059C';'YKL060C';'YKL078W';'YKL082C
';'YKL083W';'YKL088W';'YKL089W';'YKL095W';'YKL099C';'YKL104C';'YKL108W';'YKL1
11C';'YKL112W';'YKL122C';'YKL125W';'YKL138C-A';'YKL138C-
A';'YKL141W';'YKL144C';'YKL145W';'YKL152C';'YKL153W';'YKL154W';'YKL165C';'YKL
172W';'YKL172W';'YKL172W';'YKL173W';'YKL180W';'YKL182W';'YKL186C';'YKL189W';'
YKL192C';'YKL193C';'YKL195W';'YKL196C';'YKL203C';'YKL210W';'YKR002W';'YKR004C
';'YKR008W';'YKR022C';'YKR025W';'YKR037C';'YKR038C';'YKR062W';'YKR063C';'YKR0
68C';'YKR071C';'YKR079C';'YKR081C';'YKR083C';'YKR086W';'YLL003W';'YLL004W';'Y
LL008W';'YLL011W';'YLL018C';'YLL031C';'YLL034C';'YLL035W';'YLL036C';'YLL037W'
;'YLL050C';'YLR002C';'YLR005W';'YLR007W';'YLR008C';'YLR009W';'YLR010C';'YLR02
2C';'YLR026C';'YLR029C';'YLR033W';'YLR045C';'YLR051C';'YLR060W';'YLR066W';'YL
R071C';'YLR075W';'YLR076C';'YLR078C';'YLR086W';'YLR088W';'YLR099W-
A';'YLR099W-
A';'YLR100W';'YLR101C';'YLR103C';'YLR105C';'YLR106C';'YLR115W';'YLR116W';'YLR
117C';'YLR127C';'YLR129W';'YLR132C';'YLR140W';'YLR141W';'YLR145W';'YLR147C';'
YLR153C';'YLR163C';'YLR166C';'YLR167W';'YLR175W';'YLR186W';'YLR186W';'YLR195C
';'YLR196W';'YLR197W';'YLR198C';'YLR208W';'YLR212C';'YLR215C';'YLR222C';'YLR2
23C';'YLR229C';'YLR230W';'YLR243W';'YLR249W';'YLR259C';'YLR272C';'YLR274W';'Y
LR275W';'YLR276C';'YLR277C';'YLR291C';'YLR293C';'YLR298C';'YLR305C';'YLR310C'
;'YLR314C';'YLR316C';'YLR316C';'YLR316C';'YLR317W';'YLR321C';'YLR323C';'YLR33
6C';'YLR339C';'YLR340W';'YLR347C';'YLR355C';'YLR359W';'YLR378C';'YLR379W';'YL
R383W';'YLR397C';'YLR409C';'YLR424W';'YLR430W';'YLR438C-
A';'YLR440C';'YLR457C';'YLR458W';'YLR459W';'YML010W';'YML015C';'YML015C';'YML
023C';'YML023C';'YML025C';'YML031W';'YML043C';'YML046W';'YML049C';'YML064C';'
YML065W';'YML069W';'YML077W';'YML085C';'YML091C';'YML092C';'YML092C';'YML092C
';'YML093W';'YML098W';'YML105C';'YML114C';'YML125C';'YML126C';'YML127W';'YML1
30C';'YMR001C';'YMR005W';'YMR005W';'YMR013C';'YMR028W';'YMR033W';'YMR033W';'Y
MR043W';'YMR047C';'YMR047C';'YMR049C';'YMR059W';'YMR059W';'YMR059W';'YMR061W'
;'YMR076C';'YMR079W';'YMR093W';'YMR094W';'YMR108W';'YMR108W';'YMR112C';'YMR11
3W';'YMR117C';'YMR128W';'YMR131C';'YMR134W';'YMR146C';'YMR149W';'YMR168C';'YM
R197C';'YMR200W';'YMR203W';'YMR208W';'YMR211W';'YMR213W';'YMR218C';'YMR220W';
'YMR227C';'YMR229C';'YMR235C';'YMR236W';'YMR239C';'YMR240C';'YMR260C';'YMR268
C';'YMR270C';'YMR277W';'YMR281W';'YMR288W';'YMR290C';'YMR290W-
A';'YMR296C';'YMR298W';'YMR301C';'YMR308C';'YMR309C';'YMR314W';'YNL002C';'YNL
006W';'YNL007C';'YNL024C-A';'YNL024C-
A';'YNL026W';'YNL036W';'YNL036W';'YNL038W';'YNL039W';'YNL061W';'YNL062C';'YNL
075W';'YNL088W';'YNL102W';'YNL103W';'YNL110C';'YNL112W';'YNL112W';'YNL112W';'
YNL113W';'YNL114C';'YNL118C';'YNL124W';'YNL126W';'YNL126W';'YNL131W';'YNL132W
';'YNL137C';'YNL138W-A';'YNL138W-
A';'YNL149C';'YNL150W';'YNL151C';'YNL152W';'YNL158W';'YNL161W';'YNL163C';'YNL
172W';'YNL178W';'YNL181W';'YNL182C';'YNL188W';'YNL189W';'YNL207W';'YNL216W';'
YNL221C';'YNL222W';'YNL232W';'YNL240C';'YNL244C';'YNL245C';'YNL247W';'YNL251C
';'YNL256W';'YNL258C';'YNL260C';'YNL261W';'YNL262W';'YNL263C';'YNL267W';'YNL2
72C';'YNL282W';'YNL287W';'YNL290W';'YNL306W';'YNL308C';'YNL310C';'YNL312W';'Y
NL313C';'YNL317W';'YNR003C';'YNR011C';'YNR016C';'YNR017W';'YNR026C';'YNR035C'
;'YNR038W';'YNR043W';'YNR046W';'YNR053C';'YNR054C';'YOL005C';'YOL010W';'YOL02
1C';'YOL022C';'YOL026C';'YOL034W';'YOL038W';'YOL040C';'YOL066C';'YOL069W';'YO
L077C';'YOL078W';'YOL094C';'YOL097C';'YOL102C';'YOL120C';'YOL123W';'YOL127W';
'YOL130W';'YOL133W';'YOL134C';'YOL135C';'YOL139C';'YOL142W';'YOL142W';'YOL144
W';'YOL146W';'YOL149W';'YOR004W';'YOR020C';'YOR046C';'YOR048C';'YOR056C';'YOR
057W';'YOR060C';'YOR063W';'YOR074C';'YOR075W';'YOR077W';'YOR095C';'YOR098C';'
YOR102W';'YOR103C';'YOR110W';'YOR116C';'YOR117W';'YOR119C';'YOR122C';'YOR143C
';'YOR145C';'YOR146W';'YOR148C';'YOR149C';'YOR151C';'YOR157C';'YOR159C';'YOR1

```
60W';'YOR168W';'YOR169C';'YOR174W';'YOR176W';'YOR181W';'YOR194C';'YOR203W';'Y
OR204W';'YOR206W';'YOR207C';'YOR210W';'YOR217W';'YOR218C';'YOR224C';'YOR232W'
;'YOR236W';'YOR244W';'YOR249C';'YOR250C';'YOR254C';'YOR256C';'YOR257W';'YOR25
9C';'YOR260W';'YOR261C';'YOR262W';'YOR272W';'YOR278W';'YOR281C';'YOR282W';'YO
R287C';'YOR294W';'YOR310C';'YOR319W';'YOR326W';'YOR329C';'YOR335C';'YOR336W';
'YOR340C';'YOR341W';'YOR353C';'YOR361C';'YOR362C';'YOR370C';'YOR372C';'YOR373
W';'YPL007C';'YPL010W';'YPL011C';'YPL012W';'YPL016W';'YPL020C';'YPL028W';'YPL
043W';'YPL044C';'YPL063W';'YPL076W';'YPL082C';'YPL083C';'YPL085W';'YPL093W';'
YPL094C';'YPL117C';'YPL122C';'YPL124W';'YPL126W';'YPL128C';'YPL131W';'YPL142C
';'YPL143W';'YPL146C';'YPL151C';'YPL153C';'YPL160W';'YPL169C';'YPL175W';'YPL1
90C';'YPL204W';'YPL209C';'YPL210C';'YPL211W';'YPL217C';'YPL218W';'YPL228W';'Y
PL231W';'YPL233W';'YPL235W';'YPL237W';'YPL238C';'YPL242C';'YPL243W';'YPL251W'
;'YPL252C';'YPL255W';'YPL266W';'YPR010C';'YPR016C';'YPR019W';'YPR025C';'YPR03
3C';'YPR034W';'YPR035W';'YPR041W';'YPR048W';'YPR055W';'YPR056W';'YPR082C';'YP
R085C';'YPR086W';'YPR088C';'YPR094W';'YPR103W';'YPR104C';'YPR105C';'YPR107C';
'YPR108W';'YPR110C';'YPR112C';'YPR113W';'YPR133C';'YPR136C';'YPR137W';'YPR142
C';'YPR143W';'YPR144C';'YPR161C';'YPR162C';'YPR165W';'YPR168W';'YPR169W';'YPR
175W';'YPR176C';'YPR177C';'YPR178W';'YPR180W';'YPR181C';'YPR182W';'YPR183W';'
YPR186C';'YPR187W';'YPR190C';};

function genes = auxotrophicORFs

% list of 432 auxotrophic ORFs (669 annotated auxotroph minus 52 on YKO
% essential list minus 87 which are re either temp-sensitive ino
% auxotrophs or not auxotrophs in the nature study):

genes={'YAL021C';'YAL024C';'YAL026C';'YAL040C';'YAL051W';'YAL056W';'YAL058W';
'YAR003W';'YAR015W';'YAR069W-A';'YAR070W-
A';'YBL027W';'YBL033C';'YBL047C';'YBL058W';'YBL061C';'YBL091C-
A';'YBL098W';'YBL102W';'YBL103C';'YBR015C';'YBR058C';'YBR077C';'YBR106W';'YBR
107C';'YBR115C';'YBR126C';'YBR127C';'YBR133C';'YBR175W';'YBR176W';'YBR189W';'
YBR191W';'YBR248C';'YBR272C';'YBR279W';'YCL018W';'YCL030C';'YCL032W';'YCL033C
';'YCL045C';'YCR021C';'YCR045C';'YCR047C';'YCR053W';'YCR076C';'YCR089W';'YCR0
94W';'YDL001W';'YDL002C';'YDL006W';'YDL010W';'YDL020C';'YDL021W';'YDL033C';'Y
DL040C';'YDL048C';'YDL069C';'YDL073W';'YDL074C';'YDL077C';'YDL081C';'YDL083C'
;'YDL106C';'YDL130W';'YDL131W';'YDL173W';'YDL190C';'YDL191W';'YDL192W';'YDL19
4W';'YDL201W';'YDL203C';'YDR007W';'YDR043C';'YDR049W';'YDR057W';'YDR071C';'YD
R074W';'YDR080W';'YDR138W';'YDR162C';'YDR173C';'YDR174W';'YDR176W';'YDR200C';
'YDR226W';'YDR260C';'YDR266C';'YDR276C';'YDR277C';'YDR283C';'YDR289C';'YDR335
W';'YDR346C';'YDR348C';'YDR351W';'YDR354W';'YDR358W';'YDR363W';'YDR379W';'YDR
385W';'YDR389W';'YDR392W';'YDR395W';'YDR411C';'YDR422C';'YDR432W';'YDR439W';'
YDR448W';'YDR469W';'YDR477W';'YDR482C';'YDR486C';'YDR540C';'YEL004W';'YEL013W
';'YEL021W';'YEL027W';'YEL029C';'YEL031W';'YEL037C';'YEL040W';'YEL044W';'YEL0
48C';'YEL051W';'YER001W';'YER007C-
A';'YER026C';'YER027C';'YER052C';'YER055C';'YER059W';'YER069W';'YER083C';'YER
090W';'YER091C';'YER092W';'YER095W';'YER101C';'YER116C';'YER118C';'YER120W';'
YER122C';'YER129W';'YER130C';'YER149C';'YER150W';'YER152C';'YER167W';'YER169W
';'YER177W';'YFL013C';'YFL031W';'YFR010W';'YFR019W';'YFR025C';'YFR040W';'YFR0
47C';'YFR048W';'YGL009C';'YGL012W';'YGL020C';'YGL025C';'YGL026C';'YGL031C';'Y
GL049C';'YGL054C';'YGL058W';'YGL060W';'YGL062W';'YGL066W';'YGL070C';'YGL115W'
;'YGL126W';'YGL127C';'YGL154C';'YGL167C';'YGL168W';'YGL175C';'YGL179C';'YGL18
0W';'YGL181W';'YGL203C';'YGL211W';'YGL219C';'YGL234W';'YGL244W';'YGR014W';'YG
R020C';'YGR056W';'YGR057C';'YGR061C';'YGR063C';'YGR092W';'YGR104C';'YGR105W';
'YGR108W';'YGR135W';'YGR144W';'YGR162W';'YGR166W';'YGR204W';'YGR223C';'YGR227
W';'YGR229C';'YGR241C';'YGR250C';'YGR252W';'YHR010W';'YHR013C';'YHR018C';'YHR
025W';'YHR026W';'YHR030C';'YHR038W';'YHR060W';'YHR079C';'YHR111W';'YHR142W';'
YHR162W';'YHR178W';'YHR179W';'YHR199C';'YHR206W';'YHR208W';'YIL017C';'YIL020C
```

```
';'YIL027C';'YIL029C';'YIL036W';'YIL044C';'YIL072W';'YIL077C';'YIL105C';'YIL1
16W';'YIL119C';'YIL128W';'YIL153W';'YIR017C';'YIR034C';'YJL088W';'YJL095W';'Y
JL115W';'YJL117W';'YJL128C';'YJL130C';'YJL140W';'YJL151C';'YJL153C';'YJL158C'
;'YJL172W';'YJL192C';'YJL193W';'YJL204C';'YJL208C';'YJR066W';'YJR075W';'YJR08
3C';'YJR104C';'YJR122W';'YKL001C';'YKL006W';'YKL027W';'YKL032C';'YKL048C';'YK
L053C-
A';'YKL056C';'YKL064W';'YKL077W';'YKL079W';'YKL119C';'YKL121W';'YKL160W';'YKL
176C';'YKL190W';'YKL211C';'YKL212W';'YKL213C';'YKL216W';'YKR007W';'YKR026C';'
YKR036C';'YKR070W';'YKR099W';'YLL019C';'YLL021W';'YLL027W';'YLL039C';'YLR015W
';'YLR016C';'YLR021W';'YLR048W';'YLR055C';'YLR061W';'YLR074C';'YLR079W';'YLR0
87C';'YLR113W';'YLR150W';'YLR192C';'YLR199C';'YLR226W';'YLR242C';'YLR262C';'Y
LR268W';'YLR292C';'YLR315W';'YLR320W';'YLR324W';'YLR332W';'YLR357W';'YLR371W'
;'YLR373C';'YLR396C';'YLR417W';'YLR418C';'YLR420W';'YLR426W';'YLR436C';'YML00
8C';'YML013W';'YML014W';'YML028W';'YML034W';'YML055W';'YML071C';'YML103C';'YM
L115C';'YML117W';'YMR010W';'YMR014W';'YMR016C';'YMR029C';'YMR038C';'YMR052W';
'YMR062C';'YMR067C';'YMR068W';'YMR092C';'YMR099C';'YMR104C';'YMR123W';'YMR165
C';'YMR190C';'YMR202W';'YMR214W';'YMR217W';'YMR242C';'YMR247C';'YMR276W';'YMR
300C';'YMR304W';'YMR307W';'YMR312W';'YNL003C';'YNL041C';'YNL051W';'YNL079C';'
YNL080C';'YNL119W';'YNL127W';'YNL133C';'YNL148C';'YNL215W';'YNL219C';'YNL220W
';'YNL229C';'YNL236W';'YNL241C';'YNL277W';'YNL307C';'YNL316C';'YNL322C';'YNR0
55C';'YOL018C';'YOL067C';'YOL087C';'YOL090W';'YOL093W';'YOL098C';'YOL107W';'Y
OL109W';'YOL111C';'YOL116W';'YOL121C';'YOL122C';'YOL124C';'YOL143C';'YOL145C'
;'YOR002W';'YOR008C';'YOR012W';'YOR067C';'YOR070C';'YOR078W';'YOR096W';'YOR10
6W';'YOR123C';'YOR128C';'YOR189W';'YOR202W';'YOR216C';'YOR246C';'YOR290C';'YO
R320C';'YOR322C';'YOR359W';'YOR371C';'YPL055C';'YPL065W';'YPL089C';'YPL138C';
'YPL140C';'YPL144W';'YPL157W';'YPL159C';'YPL174C';'YPL177C';'YPL226W';'YPL241
C';'YPL254W';'YPL264C';'YPR036W';'YPR043W';'YPR060C';'YPR067W';'YPR139C';'YPR
167C';'YPR173C';'YPR179C';'YPR201W';};

function genes = verifiedORFs

% list of 4941 verified ORFs taken from SGD (14 aug 11)
% http://www.yeastgenome.org/cgi-
bin/search/featureSearch?featuretype=ORF&qualifier=Verified

genes =
{'Q0045';'Q0050';'Q0055';'Q0060';'Q0065';'Q0070';'Q0080';'Q0085';'Q0105';'Q01
10';'Q0115';'Q0120';'Q0130';'Q0140';'Q0160';'Q0250';'Q0275';'R0010W';'R0020C'
;'R0030W';'R0040C';'YAL001C';'YAL002W';'YAL003W';'YAL005C';'YAL007C';'YAL008W
';'YAL009W';'YAL010C';'YAL011W';'YAL012W';'YAL013W';'YAL014C';'YAL015C';'YAL0
16W';'YAL017W';'YAL019W';'YAL020C';'YAL021C';'YAL022C';'YAL023C';'YAL024C';'Y
AL025C';'YAL026C';'YAL027W';'YAL028W';'YAL029C';'YAL030W';'YAL031C';'YAL032C'
;'YAL033W';'YAL034C';'YAL034W-
A';'YAL035W';'YAL036C';'YAL038W';'YAL039C';'YAL040C';'YAL041W';'YAL042W';'YAL
043C';'YAL044C';'YAL046C';'YAL047C';'YAL048C';'YAL049C';'YAL051W';'YAL053W';'
YAL054C';'YAL055W';'YAL056W';'YAL058W';'YAL059W';'YAL060W';'YAL062W';'YAL063C
';'YAL064W';'YAL067C';'YAL068C';'YAR002C-
A';'YAR002W';'YAR003W';'YAR007C';'YAR008W';'YAR014C';'YAR015W';'YAR018C';'YAR
019C';'YAR020C';'YAR027W';'YAR031W';'YAR033W';'YAR035W';'YAR042W';'YAR050W';'
YAR071W';'YBL001C';'YBL002W';'YBL003C';'YBL004W';'YBL005W';'YBL006C';'YBL007C
';'YBL008W';'YBL009W';'YBL011W';'YBL013W';'YBL014C';'YBL015W';'YBL016W';'YBL0
17C';'YBL018C';'YBL019W';'YBL020W';'YBL021C';'YBL022C';'YBL023C';'YBL024W';'Y
BL025W';'YBL026W';'YBL027W';'YBL028C';'YBL030C';'YBL031W';'YBL032W';'YBL033C'
;'YBL034C';'YBL035C';'YBL036C';'YBL037W';'YBL038W';'YBL039C';'YBL040C';'YBL04
1W';'YBL042C';'YBL043W';'YBL045C';'YBL046W';'YBL047C';'YBL049W';'YBL050W';'YB
L051C';'YBL052C';'YBL054W';'YBL055C';'YBL056W';'YBL057C';'YBL058W';'YBL059C-
A';'YBL060W';'YBL061C';'YBL063W';'YBL064C';'YBL066C';'YBL067C';'YBL068W';'YBL
```

```
069W';'YBL071W-
A';'YBL072C';'YBL074C';'YBL075C';'YBL076C';'YBL078C';'YBL079W';'YBL080C';'YBL
082C';'YBL084C';'YBL085W';'YBL087C';'YBL088C';'YBL089W';'YBL090W';'YBL091C';'
YBL091C-
A';'YBL092W';'YBL093C';'YBL097W';'YBL098W';'YBL099W';'YBL101C';'YBL102W';'YBL
103C';'YBL104C';'YBL105C';'YBL106C';'YBL108C-
A';'YBR001C';'YBR002C';'YBR003W';'YBR004C';'YBR005W';'YBR006W';'YBR008C';'YBR
009C';'YBR010W';'YBR011C';'YBR014C';'YBR015C';'YBR016W';'YBR017C';'YBR018C';'
YBR019C';'YBR020W';'YBR021W';'YBR022W';'YBR023C';'YBR024W';'YBR025C';'YBR026C
';'YBR028C';'YBR029C';'YBR030W';'YBR031W';'YBR034C';'YBR035C';'YBR036C';'YBR0
37C';'YBR038W';'YBR039W';'YBR040W';'YBR041W';'YBR042C';'YBR043C';'YBR044C';'Y
BR045C';'YBR046C';'YBR048W';'YBR049C';'YBR050C';'YBR052C';'YBR054W';'YBR055C'
;'YBR057C';'YBR058C';'YBR058C-
A';'YBR059C';'YBR060C';'YBR061C';'YBR065C';'YBR066C';'YBR067C';'YBR068C';'YBR
069C';'YBR070C';'YBR071W';'YBR072W';'YBR073W';'YBR076W';'YBR077C';'YBR078W';'
YBR079C';'YBR080C';'YBR081C';'YBR082C';'YBR083W';'YBR084C-
A';'YBR084W';'YBR085W';'YBR086C';'YBR087W';'YBR088C';'YBR089C-
A';'YBR091C';'YBR092C';'YBR093C';'YBR094W';'YBR095C';'YBR097W';'YBR098W';'YBR
101C';'YBR102C';'YBR103W';'YBR104W';'YBR105C';'YBR106W';'YBR107C';'YBR108W';'
YBR109C';'YBR110W';'YBR111C';'YBR111W-
A';'YBR112C';'YBR114W';'YBR115C';'YBR117C';'YBR118W';'YBR119W';'YBR120C';'YBR
121C';'YBR122C';'YBR123C';'YBR125C';'YBR126C';'YBR127C';'YBR128C';'YBR129C';'
YBR130C';'YBR131W';'YBR132C';'YBR133C';'YBR135W';'YBR136W';'YBR137W';'YBR139W
';'YBR140C';'YBR142W';'YBR143C';'YBR145W';'YBR146W';'YBR147W';'YBR148W';'YBR1
49W';'YBR150C';'YBR151W';'YBR152W';'YBR153W';'YBR154C';'YBR155W';'YBR156C';'Y
BR157C';'YBR158W';'YBR159W';'YBR160W';'YBR161W';'YBR162C';'YBR162W-
A';'YBR163W';'YBR164C';'YBR165W';'YBR166C';'YBR167C';'YBR168W';'YBR169C';'YBR
170C';'YBR171W';'YBR172C';'YBR173C';'YBR175W';'YBR176W';'YBR177C';'YBR179C';'
YBR180W';'YBR181C';'YBR182C';'YBR183W';'YBR185C';'YBR186W';'YBR188C';'YBR189W
';'YBR191W';'YBR192W';'YBR193C';'YBR194W';'YBR195C';'YBR196C';'YBR198C';'YBR1
99W';'YBR200W';'YBR201W';'YBR202W';'YBR203W';'YBR204C';'YBR205W';'YBR207W';'Y
BR208C';'YBR210W';'YBR211C';'YBR212W';'YBR213W';'YBR214W';'YBR215W';'YBR216C'
;'YBR217W';'YBR218C';'YBR221C';'YBR222C';'YBR223C';'YBR227C';'YBR228W';'YBR22
9C';'YBR230C';'YBR231C';'YBR233W';'YBR233W-
A';'YBR234C';'YBR236C';'YBR237W';'YBR238C';'YBR240C';'YBR243C';'YBR244W';'YBR
245C';'YBR247C';'YBR248C';'YBR249C';'YBR250W';'YBR251W';'YBR252W';'YBR253W';'
YBR254C';'YBR255W';'YBR256C';'YBR257W';'YBR258C';'YBR260C';'YBR261C';'YBR262C
';'YBR263W';'YBR264C';'YBR265W';'YBR267W';'YBR268W';'YBR272C';'YBR273C';'YBR2
74W';'YBR275C';'YBR276C';'YBR278W';'YBR279W';'YBR280C';'YBR281C';'YBR282W';'Y
BR283C';'YBR286W';'YBR288C';'YBR289W';'YBR290W';'YBR291C';'YBR293W';'YBR294W'
;'YBR295W';'YBR296C';'YBR297W';'YBR298C';'YBR299W';'YBR301W';'YBR302C';'YCL00
1W';'YCL004W';'YCL005W';'YCL005W-
A';'YCL008C';'YCL009C';'YCL010C';'YCL011C';'YCL012C';'YCL014W';'YCL016C';'YCL
017C';'YCL018W';'YCL024W';'YCL025C';'YCL026C-
A';'YCL027W';'YCL028W';'YCL029C';'YCL030C';'YCL031C';'YCL032W';'YCL033C';'YCL
034W';'YCL035C';'YCL036W';'YCL037C';'YCL038C';'YCL039W';'YCL040W';'YCL043C';'
YCL044C';'YCL045C';'YCL047C';'YCL048W';'YCL050C';'YCL051W';'YCL052C';'YCL054W
';'YCL055W';'YCL056C';'YCL057W';'YCL058C';'YCL058W-
A';'YCL059C';'YCL061C';'YCL063W';'YCL064C';'YCL066W';'YCL067C';'YCL069W';'YCL
073C';'YCR002C';'YCR003W';'YCR004C';'YCR005C';'YCR008W';'YCR009C';'YCR010C';'
YCR011C';'YCR012W';'YCR014C';'YCR017C';'YCR018C';'YCR019W';'YCR020C';'YCR020C
-A';'YCR020W-B';'YCR021C';'YCR023C';'YCR024C';'YCR024C-
A';'YCR026C';'YCR027C';'YCR028C';'YCR028C-
A';'YCR030C';'YCR031C';'YCR032W';'YCR033W';'YCR034W';'YCR035C';'YCR036W';'YCR
037C';'YCR038C';'YCR039C';'YCR040W';'YCR042C';'YCR044C';'YCR045C';'YCR046C';'
YCR047C';'YCR048W';'YCR052W';'YCR053W';'YCR054C';'YCR057C';'YCR059C';'YCR060W
';'YCR063W';'YCR065W';'YCR066W';'YCR067C';'YCR068W';'YCR069W';'YCR071C';'YCR0
```

72C';'YCR073C';'YCR073W-
A';'YCR075C';'YCR077C';'YCR079W';'YCR081W';'YCR082W';'YCR083W';'YCR084C';'YCR
086W';'YCR088W';'YCR089W';'YCR091W';'YCR092C';'YCR093W';'YCR094W';'YCR096C';'
YCR097W';'YCR098C';'YCR104W';'YCR105W';'YCR106W';'YCR107W';'YDL001W';'YDL002C
';'YDL003W';'YDL004W';'YDL005C';'YDL006W';'YDL007W';'YDL008W';'YDL010W';'YDL0
12C';'YDL013W';'YDL014W';'YDL015C';'YDL017W';'YDL018C';'YDL019C';'YDL020C';'Y
DL021W';'YDL022W';'YDL024C';'YDL025C';'YDL028C';'YDL029W';'YDL030W';'YDL031W'
;'YDL033C';'YDL035C';'YDL036C';'YDL037C';'YDL039C';'YDL040C';'YDL042C';'YDL04
3C';'YDL044C';'YDL045C';'YDL045W-
A';'YDL046W';'YDL047W';'YDL048C';'YDL049C';'YDL051W';'YDL052C';'YDL053C';'YDL
054C';'YDL055C';'YDL056W';'YDL058W';'YDL059C';'YDL060W';'YDL061C';'YDL063C';'
YDL064W';'YDL065C';'YDL066W';'YDL067C';'YDL069C';'YDL070W';'YDL072C';'YDL074C
';'YDL075W';'YDL076C';'YDL077C';'YDL078C';'YDL079C';'YDL080C';'YDL081C';'YDL0
82W';'YDL083C';'YDL084W';'YDL085W';'YDL087C';'YDL088C';'YDL089W';'YDL090C';'Y
DL091C';'YDL092W';'YDL093W';'YDL095W';'YDL097C';'YDL098C';'YDL099W';'YDL100C'
;'YDL101C';'YDL102W';'YDL103C';'YDL104C';'YDL105W';'YDL106C';'YDL107W';'YDL10
8W';'YDL110C';'YDL111C';'YDL112W';'YDL113C';'YDL115C';'YDL116W';'YDL117W';'YD
L120W';'YDL122W';'YDL123W';'YDL124W';'YDL125C';'YDL126C';'YDL127W';'YDL128W';
'YDL130W';'YDL130W-A';'YDL131W';'YDL132W';'YDL133C-
A';'YDL133W';'YDL134C';'YDL135C';'YDL136W';'YDL137W';'YDL138W';'YDL139C';'YDL
140C';'YDL141W';'YDL142C';'YDL143W';'YDL145C';'YDL146W';'YDL147W';'YDL148C';'
YDL149W';'YDL150W';'YDL153C';'YDL154W';'YDL155W';'YDL159W';'YDL160C';'YDL161W
';'YDL164C';'YDL165W';'YDL166C';'YDL167C';'YDL168W';'YDL169C';'YDL170W';'YDL1
71C';'YDL173W';'YDL174C';'YDL175C';'YDL176W';'YDL178W';'YDL179W';'YDL181W';'Y
DL182W';'YDL183C';'YDL184C';'YDL185W';'YDL188C';'YDL189W';'YDL190C';'YDL191W'
;'YDL192W';'YDL193W';'YDL194W';'YDL195W';'YDL197C';'YDL198C';'YDL200C';'YDL20
1W';'YDL202W';'YDL203C';'YDL204W';'YDL205C';'YDL207W';'YDL208W';'YDL209C';'YD
L210W';'YDL212W';'YDL213C';'YDL214C';'YDL215C';'YDL216C';'YDL217C';'YDL219W';
'YDL220C';'YDL222C';'YDL223C';'YDL224C';'YDL225W';'YDL226C';'YDL227C';'YDL229
W';'YDL230W';'YDL231C';'YDL232W';'YDL234C';'YDL235C';'YDL236W';'YDL237W';'YDL
238C';'YDL239C';'YDL240W';'YDL243C';'YDL244W';'YDL245C';'YDL247W';'YDL248W';'
YDR001C';'YDR002W';'YDR003W';'YDR004W';'YDR005C';'YDR006C';'YDR007W';'YDR009W
';'YDR011W';'YDR012W';'YDR013W';'YDR014W';'YDR014W-
A';'YDR016C';'YDR017C';'YDR019C';'YDR021W';'YDR022C';'YDR023W';'YDR025W';'YDR
026C';'YDR027C';'YDR028C';'YDR030C';'YDR031W';'YDR032C';'YDR033W';'YDR034C';'
YDR035W';'YDR036C';'YDR037W';'YDR038C';'YDR039C';'YDR040C';'YDR041W';'YDR043C
';'YDR044W';'YDR045C';'YDR046C';'YDR047W';'YDR049W';'YDR050C';'YDR051C';'YDR0
52C';'YDR054C';'YDR055W';'YDR057W';'YDR058C';'YDR059C';'YDR060W';'YDR062W';'Y
DR063W';'YDR064W';'YDR065W';'YDR068W';'YDR069C';'YDR071C';'YDR072C';'YDR073W'
;'YDR074W';'YDR075W';'YDR076W';'YDR077W';'YDR078C';'YDR079C-
A';'YDR079W';'YDR080W';'YDR081C';'YDR082W';'YDR083W';'YDR084C';'YDR085C';'YDR
086C';'YDR087C';'YDR088C';'YDR091C';'YDR092W';'YDR093W';'YDR096W';'YDR097C';'
YDR098C';'YDR099W';'YDR100W';'YDR101C';'YDR103W';'YDR104C';'YDR105C';'YDR106W
';'YDR107C';'YDR108W';'YDR110W';'YDR113C';'YDR116C';'YDR117C';'YDR118W';'YDR1
20C';'YDR121W';'YDR122W';'YDR123C';'YDR125C';'YDR126W';'YDR127W';'YDR128W';'Y
DR129C';'YDR130C';'YDR135C';'YDR137W';'YDR138W';'YDR139C';'YDR140W';'YDR141C'
;'YDR142C';'YDR143C';'YDR144C';'YDR145W';'YDR146C';'YDR147W';'YDR148C';'YDR15
0W';'YDR151C';'YDR152W';'YDR153C';'YDR155C';'YDR156W';'YDR158W';'YDR159W';'YD
R160W';'YDR162C';'YDR163W';'YDR164C';'YDR165W';'YDR166C';'YDR167W';'YDR168W';
'YDR169C';'YDR170C';'YDR171W';'YDR172W';'YDR173C';'YDR174W';'YDR175C';'YDR176
W';'YDR177W';'YDR178W';'YDR179C';'YDR180W';'YDR181C';'YDR182W';'YDR183W';'YDR
184C';'YDR185C';'YDR186C';'YDR188W';'YDR189W';'YDR190C';'YDR191W';'YDR192C';'
YDR194C';'YDR195W';'YDR196C';'YDR197W';'YDR198C';'YDR200C';'YDR201W';'YDR202C
';'YDR204W';'YDR205W';'YDR206W';'YDR207C';'YDR208W';'YDR211W';'YDR212W';'YDR2
13W';'YDR214W';'YDR216W';'YDR217C';'YDR218C';'YDR219C';'YDR221W';'YDR223W';'Y
DR224C';'YDR225W';'YDR226W';'YDR227W';'YDR228C';'YDR229W';'YDR231C';'YDR232W'
;'YDR233C';'YDR234W';'YDR235W';'YDR236C';'YDR237W';'YDR238C';'YDR239C';'YDR24

0C';'YDR242W';'YDR243C';'YDR244W';'YDR245W';'YDR246W';'YDR247W';'YDR251W';'YD
R252W';'YDR253C';'YDR254W';'YDR255C';'YDR256C';'YDR257C';'YDR258C';'YDR259C';
'YDR260C';'YDR261C';'YDR263C';'YDR264C';'YDR265W';'YDR266C';'YDR267C';'YDR268
W';'YDR270W';'YDR272W';'YDR273W';'YDR275W';'YDR276C';'YDR277C';'YDR279W';'YDR
280W';'YDR281C';'YDR283C';'YDR284C';'YDR285W';'YDR287W';'YDR288W';'YDR289C';'
YDR292C';'YDR293C';'YDR294C';'YDR295C';'YDR296W';'YDR297W';'YDR298C';'YDR299W
';'YDR300C';'YDR301W';'YDR302W';'YDR303C';'YDR304C';'YDR305C';'YDR308C';'YDR3
09C';'YDR310C';'YDR311W';'YDR312W';'YDR313C';'YDR314C';'YDR315C';'YDR316W';'Y
DR317W';'YDR318W';'YDR320C';'YDR320C-A';'YDR321W';'YDR322C-
A';'YDR322W';'YDR323C';'YDR324C';'YDR325W';'YDR326C';'YDR328C';'YDR329C';'YDR
330W';'YDR331W';'YDR332W';'YDR334W';'YDR335W';'YDR337W';'YDR339C';'YDR341C';'
YDR342C';'YDR343C';'YDR345C';'YDR346C';'YDR347W';'YDR349C';'YDR350C';'YDR351W
';'YDR353W';'YDR354W';'YDR356W';'YDR358W';'YDR359C';'YDR361C';'YDR362C';'YDR3
63W';'YDR363W-
A';'YDR364C';'YDR365C';'YDR367W';'YDR368W';'YDR369C';'YDR372C';'YDR373W';'YDR
375C';'YDR376W';'YDR377W';'YDR378C';'YDR379C-A';'YDR379W';'YDR380W';'YDR381C-
A';'YDR381W';'YDR382W';'YDR383C';'YDR384C';'YDR385W';'YDR386W';'YDR388W';'YDR
389W';'YDR390C';'YDR392W';'YDR393W';'YDR394W';'YDR395W';'YDR397C';'YDR398W';'
YDR399W';'YDR400W';'YDR402C';'YDR403W';'YDR404C';'YDR405W';'YDR406W';'YDR407C
';'YDR408C';'YDR409W';'YDR410C';'YDR411C';'YDR412W';'YDR414C';'YDR416W';'YDR4
18W';'YDR419W';'YDR420W';'YDR421W';'YDR422C';'YDR423C';'YDR424C';'YDR425W';'Y
DR427W';'YDR428C';'YDR429C';'YDR430C';'YDR432W';'YDR434W';'YDR435C';'YDR436W'
;'YDR437W';'YDR438W';'YDR439W';'YDR440W';'YDR441C';'YDR443C';'YDR446W';'YDR44
7C';'YDR448W';'YDR449C';'YDR450W';'YDR451C';'YDR452W';'YDR453C';'YDR454C';'YD
R456W';'YDR457W';'YDR458C';'YDR459C';'YDR460W';'YDR461W';'YDR462W';'YDR463W';
'YDR464W';'YDR465C';'YDR466W';'YDR468C';'YDR469W';'YDR470C';'YDR471W';'YDR472
W';'YDR473C';'YDR475C';'YDR477W';'YDR478W';'YDR479C';'YDR480W';'YDR481C';'YDR
482C';'YDR483W';'YDR484W';'YDR485C';'YDR486C';'YDR487C';'YDR488C';'YDR489W';'
YDR490C';'YDR492W';'YDR493W';'YDR494W';'YDR495C';'YDR496C';'YDR497C';'YDR498C
';'YDR499W';'YDR500C';'YDR501W';'YDR502C';'YDR503C';'YDR504C';'YDR505C';'YDR5
07C';'YDR508C';'YDR510W';'YDR511W';'YDR512C';'YDR513W';'YDR514C';'YDR515W';'Y
DR516C';'YDR517W';'YDR518W';'YDR519W';'YDR522C';'YDR523C';'YDR524C';'YDR525W-
A';'YDR527W';'YDR528W';'YDR529C';'YDR530C';'YDR531W';'YDR532C';'YDR533C';'YDR
534C';'YDR536W';'YDR538W';'YDR539W';'YDR540C';'YDR542W';'YDR545W';'YEL001C';'
YEL002C';'YEL003W';'YEL004W';'YEL005C';'YEL006W';'YEL009C';'YEL011W';'YEL012W
';'YEL013W';'YEL015W';'YEL016C';'YEL017C-
A';'YEL017W';'YEL018W';'YEL019C';'YEL020W-
A';'YEL021W';'YEL022W';'YEL024W';'YEL026W';'YEL027W';'YEL029C';'YEL030W';'YEL
031W';'YEL032W';'YEL034W';'YEL036C';'YEL037C';'YEL038W';'YEL039C';'YEL040W';'
YEL041W';'YEL042W';'YEL043W';'YEL044W';'YEL046C';'YEL047C';'YEL048C';'YEL049W
';'YEL050C';'YEL051W';'YEL052W';'YEL053C';'YEL054C';'YEL055C';'YEL056W';'YEL0
58W';'YEL059C-
A';'YEL060C';'YEL061C';'YEL062W';'YEL063C';'YEL064C';'YEL065W';'YEL066W';'YEL
069C';'YEL071W';'YEL072W';'YER001W';'YER002W';'YER003C';'YER004W';'YER005W';'
YER006W';'YER007C-
A';'YER007W';'YER008C';'YER009W';'YER010C';'YER011W';'YER012W';'YER013W';'YER
014C-A';'YER014W';'YER015W';'YER016W';'YER017C';'YER018C';'YER019C-
A';'YER019W';'YER020W';'YER021W';'YER022W';'YER023W';'YER024W';'YER025W';'YER
026C';'YER027C';'YER028C';'YER029C';'YER030W';'YER031C';'YER032W';'YER033C';'
YER035W';'YER036C';'YER037W';'YER038C';'YER039C';'YER040W';'YER041W';'YER042W
';'YER043C';'YER044C';'YER044C-
A';'YER045C';'YER046W';'YER047C';'YER048C';'YER048W-
A';'YER049W';'YER050C';'YER051W';'YER052C';'YER053C';'YER054C';'YER055C';'YER
056C';'YER056C-A';'YER057C';'YER058W';'YER059W';'YER060W';'YER060W-
A';'YER061C';'YER062C';'YER063W';'YER065C';'YER067W';'YER068W';'YER069W';'YER
070W';'YER072W';'YER073W';'YER074W';'YER074W-
A';'YER075C';'YER078C';'YER080W';'YER081W';'YER082C';'YER083C';'YER086W';'YER

```
087C-
B';'YER087W';'YER088C';'YER089C';'YER090W';'YER091C';'YER092W';'YER093C';'YER
093C-
A';'YER094C';'YER095W';'YER096W';'YER098W';'YER099C';'YER100W';'YER101C';'YER
102W';'YER103W';'YER104W';'YER105C';'YER106W';'YER107C';'YER109C';'YER110C';'
YER111C';'YER112W';'YER113C';'YER114C';'YER115C';'YER116C';'YER117W';'YER118C
';'YER119C';'YER120W';'YER122C';'YER123W';'YER124C';'YER125W';'YER126C';'YER1
27W';'YER128W';'YER129W';'YER131W';'YER132C';'YER133W';'YER134C';'YER136W';'Y
ER139C';'YER141W';'YER142C';'YER143W';'YER144C';'YER145C';'YER146W';'YER147C'
;'YER148W';'YER149C';'YER150W';'YER151C';'YER152C';'YER153C';'YER154W';'YER15
5C';'YER157W';'YER159C';'YER161C';'YER162C';'YER164W';'YER165W';'YER166W';'YE
R167W';'YER168C';'YER169W';'YER170W';'YER171W';'YER172C';'YER173W';'YER174C';
'YER175C';'YER176W';'YER177W';'YER178W';'YER179W';'YER180C';'YER180C-
A';'YER183C';'YER185W';'YER190W';'YFL001W';'YFL002C';'YFL003C';'YFL004W';'YFL
005W';'YFL007W';'YFL008W';'YFL009W';'YFL010C';'YFL010W-
A';'YFL011W';'YFL013C';'YFL014W';'YFL016C';'YFL017C';'YFL017W-
A';'YFL018C';'YFL020C';'YFL021W';'YFL022C';'YFL023W';'YFL024C';'YFL025C';'YFL
026W';'YFL027C';'YFL028C';'YFL029C';'YFL030W';'YFL031W';'YFL033C';'YFL034C-
A';'YFL034C-
B';'YFL036W';'YFL037W';'YFL038C';'YFL039C';'YFL041W';'YFL044C';'YFL045C';'YFL
047W';'YFL048C';'YFL049W';'YFL050C';'YFL053W';'YFL055W';'YFL056C';'YFL057C';'
YFL058W';'YFL059W';'YFL060C';'YFL062W';'YFR001W';'YFR002W';'YFR003C';'YFR004W
';'YFR005C';'YFR007W';'YFR008W';'YFR009W';'YFR010W';'YFR011C';'YFR013W';'YFR0
14C';'YFR015C';'YFR016C';'YFR017C';'YFR019W';'YFR021W';'YFR022W';'YFR023W';'Y
FR024C-
A';'YFR025C';'YFR026C';'YFR027W';'YFR028C';'YFR029W';'YFR030W';'YFR031C';'YFR
031C-A';'YFR032C-
A';'YFR033C';'YFR034C';'YFR036W';'YFR037C';'YFR038W';'YFR040W';'YFR041C';'YFR
042W';'YFR043C';'YFR044C';'YFR046C';'YFR047C';'YFR048W';'YFR049W';'YFR050C';'
YFR051C';'YFR052W';'YFR053C';'YGL001C';'YGL002W';'YGL003C';'YGL004C';'YGL005C
';'YGL006W';'YGL008C';'YGL009C';'YGL011C';'YGL012W';'YGL013C';'YGL014W';'YGL0
16W';'YGL017W';'YGL018C';'YGL019W';'YGL020C';'YGL021W';'YGL022W';'YGL023C';'Y
GL025C';'YGL026C';'YGL027C';'YGL028C';'YGL029W';'YGL030W';'YGL031C';'YGL032C'
;'YGL033W';'YGL035C';'YGL037C';'YGL038C';'YGL039W';'YGL040C';'YGL043W';'YGL04
4C';'YGL045W';'YGL047W';'YGL048C';'YGL049C';'YGL050W';'YGL051W';'YGL053W';'YG
L054C';'YGL055W';'YGL056C';'YGL057C';'YGL058W';'YGL059W';'YGL060W';'YGL061C';
'YGL062W';'YGL063W';'YGL064C';'YGL065C';'YGL066W';'YGL067W';'YGL068W';'YGL070
C';'YGL071W';'YGL073W';'YGL075C';'YGL076C';'YGL077C';'YGL078C';'YGL083W';'YGL
084C';'YGL086W';'YGL087C';'YGL089C';'YGL090W';'YGL091C';'YGL092W';'YGL093W';'
YGL094C';'YGL095C';'YGL096W';'YGL097W';'YGL098W';'YGL099W';'YGL100W';'YGL103W
';'YGL104C';'YGL105W';'YGL106W';'YGL107C';'YGL110C';'YGL111W';'YGL112C';'YGL1
13W';'YGL115W';'YGL116W';'YGL119W';'YGL120C';'YGL121C';'YGL122C';'YGL123W';'Y
GL124C';'YGL125W';'YGL126W';'YGL127C';'YGL128C';'YGL129C';'YGL130W';'YGL131C'
;'YGL133W';'YGL134W';'YGL135W';'YGL136C';'YGL137W';'YGL139W';'YGL141W';'YGL14
2C';'YGL143C';'YGL144C';'YGL145W';'YGL147C';'YGL148W';'YGL150C';'YGL151W';'YG
L153W';'YGL154C';'YGL155W';'YGL156W';'YGL157W';'YGL158W';'YGL160W';'YGL161C';
'YGL162W';'YGL163C';'YGL164C';'YGL166W';'YGL167C';'YGL168W';'YGL169W';'YGL170
C';'YGL171W';'YGL172W';'YGL173C';'YGL174W';'YGL175C';'YGL178W';'YGL179C';'YGL
180W';'YGL181W';'YGL183C';'YGL184C';'YGL186C';'YGL187C';'YGL189C';'YGL190C';'
YGL191W';'YGL192W';'YGL194C';'YGL195W';'YGL196W';'YGL197W';'YGL198W';'YGL200C
';'YGL201C';'YGL202W';'YGL203C';'YGL205W';'YGL206C';'YGL207W';'YGL208W';'YGL2
09W';'YGL210W';'YGL211W';'YGL212W';'YGL213C';'YGL215W';'YGL216W';'YGL219C';'Y
GL220W';'YGL221C';'YGL222C';'YGL223C';'YGL224C';'YGL225W';'YGL226C-
A';'YGL226W';'YGL227W';'YGL228W';'YGL229C';'YGL231C';'YGL232W';'YGL233W';'YGL
234W';'YGL236C';'YGL237C';'YGL238W';'YGL240W';'YGL241W';'YGL243W';'YGL244W';'
YGL245W';'YGL246C';'YGL247W';'YGL248W';'YGL249W';'YGL250W';'YGL251C';'YGL252C
';'YGL253W';'YGL254W';'YGL255W';'YGL256W';'YGL257C';'YGL258W';'YGL263W';'YGR0
```

02C';'YGR003W';'YGR004W';'YGR005C';'YGR006W';'YGR007W';'YGR008C';'YGR009C';'Y
GR010W';'YGR012W';'YGR013W';'YGR014W';'YGR019W';'YGR020C';'YGR023W';'YGR024C'
;'YGR027C';'YGR028W';'YGR029W';'YGR030C';'YGR031C-
A';'YGR031W';'YGR032W';'YGR033C';'YGR034W';'YGR036C';'YGR037C';'YGR038W';'YGR
040W';'YGR041W';'YGR043C';'YGR044C';'YGR046W';'YGR047C';'YGR048W';'YGR049W';'
YGR054W';'YGR055W';'YGR056W';'YGR057C';'YGR058W';'YGR059W';'YGR060W';'YGR061C
';'YGR062C';'YGR063C';'YGR065C';'YGR068C';'YGR070W';'YGR072W';'YGR074W';'YGR0
75C';'YGR076C';'YGR077C';'YGR078C';'YGR080W';'YGR081C';'YGR082W';'YGR083C';'Y
GR084C';'YGR085C';'YGR086C';'YGR087C';'YGR088W';'YGR089W';'YGR090W';'YGR091W'
;'YGR092W';'YGR094W';'YGR095C';'YGR096W';'YGR097W';'YGR098C';'YGR099W';'YGR10
0W';'YGR101W';'YGR102C';'YGR103W';'YGR104C';'YGR105W';'YGR106C';'YGR108W';'YG
R109C';'YGR110W';'YGR112W';'YGR113W';'YGR116W';'YGR118W';'YGR119C';'YGR120C';
'YGR121C';'YGR122W';'YGR123C';'YGR124W';'YGR128C';'YGR129W';'YGR130C';'YGR131
W';'YGR132C';'YGR133W';'YGR134W';'YGR135W';'YGR136W';'YGR138C';'YGR140W';'YGR
141W';'YGR142W';'YGR143W';'YGR144W';'YGR145W';'YGR146C';'YGR147C';'YGR148C';'
YGR150C';'YGR152C';'YGR154C';'YGR155W';'YGR156W';'YGR157W';'YGR158C';'YGR159C
';'YGR162W';'YGR163W';'YGR165W';'YGR166W';'YGR167W';'YGR169C';'YGR170W';'YGR1
71C';'YGR172C';'YGR173W';'YGR174C';'YGR175C';'YGR177C';'YGR178C';'YGR179C';'Y
GR180C';'YGR181W';'YGR183C';'YGR184C';'YGR185C';'YGR186W';'YGR187C';'YGR188C'
;'YGR189C';'YGR191W';'YGR192C';'YGR193C';'YGR194C';'YGR195W';'YGR196C';'YGR19
7C';'YGR198W';'YGR199W';'YGR200C';'YGR202C';'YGR203W';'YGR204W';'YGR205W';'YG
R206W';'YGR207C';'YGR208W';'YGR209C';'YGR211W';'YGR212W';'YGR213C';'YGR214W';
'YGR215W';'YGR216C';'YGR217W';'YGR218W';'YGR220C';'YGR221C';'YGR222W';'YGR223
C';'YGR224W';'YGR225W';'YGR227W';'YGR229C';'YGR230W';'YGR231C';'YGR232W';'YGR
233C';'YGR234W';'YGR236C';'YGR238C';'YGR239C';'YGR240C';'YGR241C';'YGR244C';'
YGR245C';'YGR246C';'YGR247W';'YGR248W';'YGR249W';'YGR250C';'YGR251W';'YGR252W
';'YGR253C';'YGR254W';'YGR255C';'YGR256W';'YGR257C';'YGR258C';'YGR260W';'YGR2
61C';'YGR262C';'YGR263C';'YGR264C';'YGR266W';'YGR267C';'YGR268C';'YGR270W';'Y
GR271C-
A';'YGR271W';'YGR274C';'YGR275W';'YGR276C';'YGR277C';'YGR278W';'YGR279C';'YGR
280C';'YGR281W';'YGR282C';'YGR283C';'YGR284C';'YGR285C';'YGR286C';'YGR287C';'
YGR288W';'YGR289C';'YGR292W';'YGR294W';'YGR295C';'YGR296W';'YHL001W';'YHL002W
';'YHL003C';'YHL004W';'YHL006C';'YHL007C';'YHL009C';'YHL010C';'YHL011C';'YHL0
13C';'YHL014C';'YHL015W';'YHL016C';'YHL019C';'YHL020C';'YHL021C';'YHL022C';'Y
HL023C';'YHL024W';'YHL025W';'YHL027W';'YHL028W';'YHL030W';'YHL031C';'YHL032C'
;'YHL033C';'YHL034C';'YHL035C';'YHL036W';'YHL038C';'YHL039W';'YHL040C';'YHL04
3W';'YHL046C';'YHL047C';'YHL048W';'YHR001W';'YHR001W-
A';'YHR002W';'YHR003C';'YHR004C';'YHR005C';'YHR005C-
A';'YHR006W';'YHR007C';'YHR008C';'YHR010W';'YHR011W';'YHR012W';'YHR013C';'YHR
014W';'YHR015W';'YHR016C';'YHR017W';'YHR018C';'YHR019C';'YHR020W';'YHR021C';'
YHR023W';'YHR024C';'YHR025W';'YHR026W';'YHR027C';'YHR028C';'YHR029C';'YHR030C
';'YHR031C';'YHR032W';'YHR034C';'YHR036W';'YHR037W';'YHR038W';'YHR039C';'YHR0
39C-
A';'YHR040W';'YHR041C';'YHR042W';'YHR043C';'YHR044C';'YHR046C';'YHR047C';'YHR
049W';'YHR050W';'YHR051W';'YHR052W';'YHR053C';'YHR055C';'YHR056C';'YHR057C';'
YHR058C';'YHR059W';'YHR060W';'YHR061C';'YHR062C';'YHR063C';'YHR064C';'YHR065C
';'YHR066W';'YHR067W';'YHR068W';'YHR069C';'YHR070W';'YHR071W';'YHR072W';'YHR0
72W-A';'YHR073W';'YHR074W';'YHR075C';'YHR076W';'YHR077C';'YHR079C';'YHR079C-
A';'YHR080C';'YHR081W';'YHR082C';'YHR083W';'YHR084W';'YHR085W';'YHR086W';'YHR
087W';'YHR088W';'YHR089C';'YHR090C';'YHR091C';'YHR092C';'YHR094C';'YHR096C';'
YHR098C';'YHR099W';'YHR100C';'YHR101C';'YHR102W';'YHR103W';'YHR104W';'YHR105W
';'YHR106W';'YHR107C';'YHR108W';'YHR109W';'YHR110W';'YHR111W';'YHR112C';'YHR1
13W';'YHR114W';'YHR115C';'YHR116W';'YHR117W';'YHR118C';'YHR119W';'YHR120W';'Y
HR121W';'YHR122W';'YHR123W';'YHR124W';'YHR127W';'YHR128W';'YHR129C';'YHR132C'
;'YHR132W-
A';'YHR133C';'YHR134W';'YHR135C';'YHR136C';'YHR137W';'YHR139C';'YHR141C';'YHR
142W';'YHR143W';'YHR143W-

A';'YHR144C';'YHR146W';'YHR147C';'YHR148W';'YHR149C';'YHR150W';'YHR151C';'YHR
152W';'YHR153C';'YHR154W';'YHR155W';'YHR156C';'YHR157W';'YHR158C';'YHR160C';'
YHR161C';'YHR163W';'YHR164C';'YHR165C';'YHR166C';'YHR167W';'YHR168W';'YHR169W
';'YHR170W';'YHR171W';'YHR172W';'YHR174W';'YHR175W';'YHR176W';'YHR178W';'YHR1
79W';'YHR181W';'YHR183W';'YHR184W';'YHR185C';'YHR186C';'YHR187W';'YHR188C';'Y
HR189W';'YHR190W';'YHR191C';'YHR193C';'YHR194W';'YHR195W';'YHR196W';'YHR197W'
;'YHR198C';'YHR199C';'YHR199C-
A';'YHR200W';'YHR201C';'YHR203C';'YHR204W';'YHR205W';'YHR206W';'YHR207C';'YHR
208W';'YHR209W';'YHR211W';'YHR215W';'YHR216W';'YIL002C';'YIL003W';'YIL004C';'
YIL005W';'YIL006W';'YIL007C';'YIL008W';'YIL009C-
A';'YIL009W';'YIL010W';'YIL011W';'YIL013C';'YIL014W';'YIL015W';'YIL016W';'YIL
017C';'YIL018W';'YIL019W';'YIL020C';'YIL021W';'YIL022W';'YIL023C';'YIL026C';'
YIL027C';'YIL030C';'YIL031W';'YIL033C';'YIL034C';'YIL035C';'YIL036W';'YIL037C
';'YIL038C';'YIL039W';'YIL040W';'YIL041W';'YIL042C';'YIL043C';'YIL044C';'YIL0
45W';'YIL046W';'YIL047C';'YIL048W';'YIL049W';'YIL050W';'YIL051C';'YIL052C';'Y
IL053W';'YIL056W';'YIL057C';'YIL061C';'YIL062C';'YIL063C';'YIL064W';'YIL065C'
;'YIL066C';'YIL068C';'YIL069C';'YIL070C';'YIL071C';'YIL072W';'YIL073C';'YIL07
4C';'YIL075C';'YIL076W';'YIL078W';'YIL079C';'YIL083C';'YIL084C';'YIL085C';'YI
L087C';'YIL088C';'YIL089W';'YIL090W';'YIL091C';'YIL093C';'YIL094C';'YIL095W';
'YIL097W';'YIL098C';'YIL099W';'YIL101C';'YIL103W';'YIL104C';'YIL105C';'YIL106
W';'YIL107C';'YIL109C';'YIL110W';'YIL111W';'YIL112W';'YIL113W';'YIL114C';'YIL
115C';'YIL116W';'YIL117C';'YIL118W';'YIL119C';'YIL120W';'YIL121W';'YIL122W';'
YIL123W';'YIL124W';'YIL125W';'YIL126W';'YIL128W';'YIL129C';'YIL130W';'YIL131C
';'YIL132C';'YIL133C';'YIL134W';'YIL135C';'YIL136W';'YIL137C';'YIL138C';'YIL1
39C';'YIL140W';'YIL142W';'YIL143C';'YIL144W';'YIL145C';'YIL146C';'YIL147C';'Y
IL148W';'YIL149C';'YIL150C';'YIL153W';'YIL154C';'YIL155C';'YIL156W';'YIL157C'
;'YIL158W';'YIL159W';'YIL160C';'YIL162W';'YIL164C';'YIL172C';'YIL173W';'YIR00
1C';'YIR002C';'YIR003W';'YIR004W';'YIR005W';'YIR006C';'YIR008C';'YIR009W';'YI
R010W';'YIR011C';'YIR012W';'YIR013C';'YIR015W';'YIR017C';'YIR018W';'YIR019C';
'YIR021W';'YIR022W';'YIR023W';'YIR024C';'YIR025W';'YIR026C';'YIR027C';'YIR028
W';'YIR029W';'YIR030C';'YIR031C';'YIR032C';'YIR033W';'YIR034C';'YIR037W';'YIR
038C';'YIR039C';'YIR041W';'YJL001W';'YJL002C';'YJL003W';'YJL004C';'YJL005W';'
YJL006C';'YJL008C';'YJL010C';'YJL011C';'YJL012C';'YJL013C';'YJL014W';'YJL019W
';'YJL020C';'YJL023C';'YJL024C';'YJL025W';'YJL026W';'YJL028W';'YJL029C';'YJL0
30W';'YJL031C';'YJL033W';'YJL034W';'YJL035C';'YJL036W';'YJL037W';'YJL038C';'Y
JL039C';'YJL041W';'YJL042W';'YJL044C';'YJL045W';'YJL046W';'YJL047C';'YJL048C'
;'YJL050W';'YJL051W';'YJL052W';'YJL053W';'YJL054W';'YJL056C';'YJL057C';'YJL05
8C';'YJL059W';'YJL060W';'YJL061W';'YJL062W';'YJL062W-
A';'YJL063C';'YJL065C';'YJL066C';'YJL068C';'YJL069C';'YJL071W';'YJL072C';'YJL
073W';'YJL074C';'YJL076W';'YJL077C';'YJL078C';'YJL079C';'YJL080C';'YJL081C';'
YJL082W';'YJL083W';'YJL084C';'YJL085W';'YJL087C';'YJL088W';'YJL089W';'YJL090C
';'YJL091C';'YJL092W';'YJL093C';'YJL094C';'YJL095W';'YJL096W';'YJL097W';'YJL0
98W';'YJL099W';'YJL100W';'YJL101C';'YJL102W';'YJL103C';'YJL104W';'YJL105W';'Y
JL106W';'YJL108C';'YJL109C';'YJL110C';'YJL111W';'YJL112W';'YJL115W';'YJL116C'
;'YJL117W';'YJL118W';'YJL121C';'YJL122W';'YJL123C';'YJL124C';'YJL125C';'YJL12
6W';'YJL127C';'YJL128C';'YJL129C';'YJL130C';'YJL131C';'YJL133W';'YJL134W';'YJ
L136C';'YJL137C';'YJL138C';'YJL139C';'YJL140W';'YJL141C';'YJL143W';'YJL144W';
'YJL145W';'YJL146W';'YJL148W';'YJL149W';'YJL151C';'YJL153C';'YJL154C';'YJL155
C';'YJL156C';'YJL157C';'YJL158C';'YJL159W';'YJL162C';'YJL164C';'YJL165C';'YJL
166W';'YJL167W';'YJL168C';'YJL170C';'YJL171C';'YJL172W';'YJL173C';'YJL174W';'
YJL176C';'YJL177W';'YJL178C';'YJL179W';'YJL180C';'YJL183W';'YJL184W';'YJL186W
';'YJL187C';'YJL189W';'YJL190C';'YJL191W';'YJL192C';'YJL194W';'YJL196C';'YJL1
97W';'YJL198W';'YJL200C';'YJL201W';'YJL203W';'YJL204C';'YJL205C';'YJL207C';'Y
JL208C';'YJL209W';'YJL210W';'YJL212C';'YJL213W';'YJL214W';'YJL216C';'YJL217W'
;'YJL219W';'YJL221C';'YJL222W';'YJL223C';'YJR001W';'YJR002W';'YJR004C';'YJR00
5W';'YJR006W';'YJR007W';'YJR009C';'YJR010C-
A';'YJR010W';'YJR013W';'YJR014W';'YJR016C';'YJR017C';'YJR019C';'YJR021C';'YJR

022W';'YJR024C';'YJR025C';'YJR031C';'YJR032W';'YJR033C';'YJR034W';'YJR035W';'
YJR036C';'YJR040W';'YJR041C';'YJR042W';'YJR043C';'YJR044C';'YJR045C';'YJR046W
';'YJR047C';'YJR048W';'YJR049C';'YJR050W';'YJR051W';'YJR052W';'YJR053W';'YJR0
55W';'YJR057W';'YJR058C';'YJR059W';'YJR060W';'YJR062C';'YJR063W';'YJR064W';'Y
JR065C';'YJR066W';'YJR067C';'YJR068W';'YJR069C';'YJR070C';'YJR072C';'YJR073C'
;'YJR074W';'YJR075W';'YJR076C';'YJR077C';'YJR078W';'YJR080C';'YJR082C';'YJR08
3C';'YJR084W';'YJR086W';'YJR088C';'YJR089W';'YJR090C';'YJR091C';'YJR092W';'YJ
R093C';'YJR094C';'YJR094W-
A';'YJR095W';'YJR096W';'YJR097W';'YJR099W';'YJR100C';'YJR101W';'YJR102C';'YJR
103W';'YJR104C';'YJR105W';'YJR106W';'YJR108W';'YJR109C';'YJR110W';'YJR112W';'
YJR113C';'YJR117W';'YJR118C';'YJR119C';'YJR120W';'YJR121W';'YJR122W';'YJR123W
';'YJR125C';'YJR126C';'YJR127C';'YJR130C';'YJR131W';'YJR132W';'YJR133W';'YJR1
34C';'YJR135C';'YJR135W-
A';'YJR136C';'YJR137C';'YJR138W';'YJR139C';'YJR140C';'YJR143C';'YJR144W';'YJR
145C';'YJR147W';'YJR148W';'YJR150C';'YJR151C';'YJR152W';'YJR153W';'YJR155W';'
YJR156C';'YJR158W';'YJR159W';'YJR160C';'YJR161C';'YKL001C';'YKL002W';'YKL003C
';'YKL004W';'YKL005C';'YKL006C-
A';'YKL006W';'YKL007W';'YKL008C';'YKL009W';'YKL010C';'YKL011C';'YKL012W';'YKL
013C';'YKL014C';'YKL015W';'YKL016C';'YKL017C';'YKL018W';'YKL019W';'YKL020C';'
YKL021C';'YKL022C';'YKL024C';'YKL025C';'YKL026C';'YKL027W';'YKL028W';'YKL029C
';'YKL032C';'YKL033W';'YKL034W';'YKL035W';'YKL037W';'YKL038W';'YKL039W';'YKL0
40C';'YKL041W';'YKL042W';'YKL043W';'YKL045W';'YKL046C';'YKL048C';'YKL049C';'Y
KL050C';'YKL051W';'YKL052C';'YKL053C-
A';'YKL054C';'YKL055C';'YKL056C';'YKL057C';'YKL058W';'YKL059C';'YKL060C';'YKL
062W';'YKL064W';'YKL065C';'YKL067W';'YKL068W';'YKL069W';'YKL072W';'YKL073W';'
YKL074C';'YKL078W';'YKL079W';'YKL080W';'YKL081W';'YKL082C';'YKL084W';'YKL085W
';'YKL086W';'YKL087C';'YKL088W';'YKL089W';'YKL090W';'YKL091C';'YKL092C';'YKL0
93W';'YKL094W';'YKL095W';'YKL096W';'YKL096W-
A';'YKL098W';'YKL099C';'YKL101W';'YKL103C';'YKL104C';'YKL106W';'YKL108W';'YKL
109W';'YKL110C';'YKL112W';'YKL113C';'YKL114C';'YKL116C';'YKL117W';'YKL119C';'
YKL120W';'YKL122C';'YKL124W';'YKL125W';'YKL126W';'YKL127W';'YKL128C';'YKL129C
';'YKL130C';'YKL132C';'YKL134C';'YKL135C';'YKL137W';'YKL138C';'YKL138C-
A';'YKL139W';'YKL140W';'YKL141W';'YKL142W';'YKL143W';'YKL144C';'YKL145W';'YKL
146W';'YKL148C';'YKL149C';'YKL150W';'YKL152C';'YKL154W';'YKL155C';'YKL156W';'
YKL157W';'YKL159C';'YKL160W';'YKL161C';'YKL163W';'YKL164C';'YKL165C';'YKL166C
';'YKL167C';'YKL168C';'YKL170W';'YKL171W';'YKL172W';'YKL173W';'YKL174C';'YKL1
75W';'YKL176C';'YKL178C';'YKL179C';'YKL180W';'YKL181W';'YKL182W';'YKL183W';'Y
KL184W';'YKL185W';'YKL186C';'YKL188C';'YKL189W';'YKL190W';'YKL191W';'YKL192C'
;'YKL193C';'YKL194C';'YKL195W';'YKL196C';'YKL197C';'YKL198C';'YKL201C';'YKL20
3C';'YKL204W';'YKL205W';'YKL206C';'YKL207W';'YKL208W';'YKL209C';'YKL210W';'YK
L211C';'YKL212W';'YKL213C';'YKL214C';'YKL215C';'YKL216W';'YKL217W';'YKL218C';
'YKL219W';'YKL220C';'YKL221W';'YKL222C';'YKL224C';'YKR001C';'YKR002W';'YKR003
W';'YKR004C';'YKR006C';'YKR007W';'YKR008W';'YKR009C';'YKR010C';'YKR013W';'YKR
014C';'YKR016W';'YKR019C';'YKR020W';'YKR021W';'YKR022C';'YKR024C';'YKR025W';'
YKR026C';'YKR027W';'YKR028W';'YKR029C';'YKR030W';'YKR031C';'YKR034W';'YKR035W
-
A';'YKR036C';'YKR037C';'YKR038C';'YKR039W';'YKR041W';'YKR042W';'YKR043C';'YKR
044W';'YKR046C';'YKR048C';'YKR049C';'YKR050W';'YKR052C';'YKR053C';'YKR054C';'
YKR055W';'YKR056W';'YKR057W';'YKR058W';'YKR059W';'YKR060W';'YKR061W';'YKR062W
';'YKR063C';'YKR064W';'YKR065C';'YKR066C';'YKR067W';'YKR068C';'YKR069W';'YKR0
71C';'YKR072C';'YKR074W';'YKR076W';'YKR077W';'YKR078W';'YKR079C';'YKR080W';'Y
KR081C';'YKR082W';'YKR083C';'YKR084C';'YKR085C';'YKR086W';'YKR087C';'YKR088C'
;'YKR089C';'YKR090W';'YKR091W';'YKR092C';'YKR093W';'YKR094C';'YKR095W';'YKR09
5W-
A';'YKR096W';'YKR097W';'YKR098C';'YKR099W';'YKR100C';'YKR101W';'YKR102W';'YKR
103W';'YKR104W';'YKR106W';'YLL001W';'YLL002W';'YLL003W';'YLL004W';'YLL005C';'
YLL006W';'YLL008W';'YLL009C';'YLL010C';'YLL011W';'YLL012W';'YLL013C';'YLL014W

';'YLL015W';'YLL018C';'YLL018C-
A';'YLL019C';'YLL021W';'YLL022C';'YLL023C';'YLL024C';'YLL025W';'YLL026W';'YLL
027W';'YLL028W';'YLL029W';'YLL031C';'YLL032C';'YLL033W';'YLL034C';'YLL035W';'
YLL036C';'YLL038C';'YLL039C';'YLL040C';'YLL041C';'YLL042C';'YLL043W';'YLL045C
';'YLL046C';'YLL048C';'YLL049W';'YLL050C';'YLL051C';'YLL052C';'YLL055W';'YLL0
57C';'YLL060C';'YLL061W';'YLL062C';'YLL063C';'YLL064C';'YLR002C';'YLR003C';'Y
LR004C';'YLR005W';'YLR006C';'YLR007W';'YLR008C';'YLR009W';'YLR010C';'YLR011W'
;'YLR013W';'YLR014C';'YLR015W';'YLR016C';'YLR017W';'YLR018C';'YLR019W';'YLR02
0C';'YLR021W';'YLR022C';'YLR023C';'YLR024C';'YLR025W';'YLR026C';'YLR027C';'YL
R028C';'YLR029C';'YLR032W';'YLR033W';'YLR034C';'YLR035C';'YLR037C';'YLR038C';
'YLR039C';'YLR043C';'YLR044C';'YLR045C';'YLR047C';'YLR048W';'YLR051C';'YLR052
W';'YLR054C';'YLR055C';'YLR056W';'YLR058C';'YLR059C';'YLR060W';'YLR061W';'YLR
064W';'YLR066W';'YLR067C';'YLR068W';'YLR069C';'YLR070C';'YLR071C';'YLR073C';'
YLR074C';'YLR075W';'YLR077W';'YLR078C';'YLR079W';'YLR080W';'YLR081W';'YLR082C
';'YLR083C';'YLR084C';'YLR085C';'YLR086W';'YLR087C';'YLR088W';'YLR089C';'YLR0
90W';'YLR091W';'YLR092W';'YLR093C';'YLR094C';'YLR095C';'YLR096W';'YLR097C';'Y
LR098C';'YLR099C';'YLR100W';'YLR102C';'YLR103C';'YLR105C';'YLR106C';'YLR107W'
;'YLR109W';'YLR110C';'YLR113W';'YLR114C';'YLR115W';'YLR116W';'YLR117C';'YLR11
8C';'YLR119W';'YLR120C';'YLR121C';'YLR127C';'YLR128W';'YLR129W';'YLR130C';'YL
R131C';'YLR133W';'YLR134W';'YLR135W';'YLR136C';'YLR137W';'YLR138W';'YLR139C';
'YLR141W';'YLR142W';'YLR144C';'YLR145W';'YLR146C';'YLR147C';'YLR148W';'YLR150
W';'YLR151C';'YLR153C';'YLR154C';'YLR154W-
C';'YLR155C';'YLR157C';'YLR158C';'YLR160C';'YLR162W';'YLR163C';'YLR164W';'YLR
165C';'YLR166C';'YLR167W';'YLR168C';'YLR170C';'YLR172C';'YLR174W';'YLR175W';'
YLR176C';'YLR178C';'YLR179C';'YLR180W';'YLR181C';'YLR182W';'YLR183C';'YLR185W
';'YLR186W';'YLR188W';'YLR189C';'YLR190W';'YLR191W';'YLR192C';'YLR193C';'YLR1
94C';'YLR195C';'YLR196W';'YLR197W';'YLR199C';'YLR200W';'YLR201C';'YLR203C';'Y
LR204W';'YLR205C';'YLR206W';'YLR207W';'YLR208W';'YLR209C';'YLR210W';'YLR212C'
;'YLR213C';'YLR214W';'YLR215C';'YLR216C';'YLR218C';'YLR219W';'YLR220W';'YLR22
1C';'YLR222C';'YLR223C';'YLR226W';'YLR227C';'YLR228C';'YLR229C';'YLR231C';'YL
R233C';'YLR234W';'YLR237W';'YLR238W';'YLR239C';'YLR240W';'YLR242C';'YLR244C';
'YLR245C';'YLR246W';'YLR247C';'YLR248W';'YLR249W';'YLR250W';'YLR251W';'YLR254
C';'YLR256W';'YLR258W';'YLR259C';'YLR260W';'YLR262C';'YLR262C-
A';'YLR263W';'YLR264W';'YLR265C';'YLR266C';'YLR268W';'YLR270W';'YLR272C';'YLR
273C';'YLR274W';'YLR275W';'YLR276C';'YLR277C';'YLR284C';'YLR285W';'YLR286C';'
YLR287C-
A';'YLR288C';'YLR289W';'YLR291C';'YLR292C';'YLR293C';'YLR295C';'YLR298C';'YLR
299W';'YLR300W';'YLR301W';'YLR303W';'YLR304C';'YLR305C';'YLR306W';'YLR307W';'
YLR308W';'YLR309C';'YLR310C';'YLR312W-
A';'YLR313C';'YLR314C';'YLR315W';'YLR316C';'YLR318W';'YLR319C';'YLR320W';'YLR
321C';'YLR323C';'YLR324W';'YLR325C';'YLR327C';'YLR328W';'YLR329W';'YLR330W';'
YLR332W';'YLR333C';'YLR335W';'YLR336C';'YLR337C';'YLR340W';'YLR341W';'YLR342W
';'YLR343W';'YLR344W';'YLR347C';'YLR348C';'YLR350W';'YLR351C';'YLR353W';'YLR3
54C';'YLR355C';'YLR356W';'YLR357W';'YLR359W';'YLR360W';'YLR361C';'YLR362W';'Y
LR363C';'YLR364W';'YLR367W';'YLR368W';'YLR369W';'YLR370C';'YLR371W';'YLR372W'
;'YLR373C';'YLR375W';'YLR376C';'YLR377C';'YLR378C';'YLR380W';'YLR381W';'YLR38
2C';'YLR383W';'YLR384C';'YLR385C';'YLR386W';'YLR387C';'YLR388W';'YLR389C';'YL
R390W';'YLR390W-
A';'YLR392C';'YLR393W';'YLR394W';'YLR395C';'YLR396C';'YLR397C';'YLR398C';'YLR
399C';'YLR401C';'YLR403W';'YLR404W';'YLR405W';'YLR406C';'YLR409C';'YLR410W';'
YLR411W';'YLR412W';'YLR414C';'YLR417W';'YLR418C';'YLR420W';'YLR421C';'YLR423C
';'YLR424W';'YLR425W';'YLR427W';'YLR429W';'YLR430W';'YLR431C';'YLR432W';'YLR4
33C';'YLR435W';'YLR436C';'YLR437C';'YLR438C-
A';'YLR438W';'YLR439W';'YLR440C';'YLR441C';'YLR442C';'YLR443W';'YLR447C';'YLR
448W';'YLR449W';'YLR450W';'YLR451W';'YLR452C';'YLR453C';'YLR457C';'YLR459W';'
YLR461W';'YLR466W';'YLR467W';'YML001W';'YML004C';'YML005W';'YML006C';'YML007W
';'YML008C';'YML009C';'YML010W';'YML011C';'YML012W';'YML013W';'YML014W';'YML0

15C';'YML016C';'YML017W';'YML019W';'YML021C';'YML022W';'YML023C';'YML024W';'Y
ML025C';'YML026C';'YML027W';'YML028W';'YML029W';'YML030W';'YML031W';'YML032C'
;'YML034W';'YML035C';'YML036W';'YML038C';'YML041C';'YML042W';'YML043C';'YML04
6W';'YML047C';'YML048W';'YML049C';'YML050W';'YML051W';'YML052W';'YML054C';'YM
L055W';'YML056C';'YML057W';'YML058W';'YML058W-
A';'YML059C';'YML060W';'YML061C';'YML062C';'YML063W';'YML064C';'YML065W';'YML
066C';'YML067C';'YML068W';'YML069W';'YML070W';'YML071C';'YML072C';'YML073C';'
YML074C';'YML075C';'YML076C';'YML077W';'YML078W';'YML080W';'YML081C-
A';'YML081W';'YML085C';'YML086C';'YML087C';'YML088W';'YML091C';'YML092C';'YML
093W';'YML094W';'YML095C';'YML097C';'YML098W';'YML099C';'YML100W';'YML101C';'
YML102W';'YML103C';'YML104C';'YML105C';'YML106W';'YML107C';'YML109W';'YML110C
';'YML111W';'YML112W';'YML113W';'YML114C';'YML115C';'YML116W';'YML117W';'YML1
18W';'YML120C';'YML121W';'YML123C';'YML124C';'YML125C';'YML126C';'YML127W';'Y
ML128C';'YML129C';'YML130C';'YML132W';'YMR001C';'YMR002W';'YMR003W';'YMR004W'
;'YMR005W';'YMR006C';'YMR008C';'YMR009W';'YMR011W';'YMR012W';'YMR013C';'YMR01
4W';'YMR015C';'YMR016C';'YMR017W';'YMR019W';'YMR020W';'YMR021C';'YMR022W';'YM
R023C';'YMR024W';'YMR025W';'YMR026C';'YMR028W';'YMR029C';'YMR030W';'YMR031C';
'YMR032W';'YMR033W';'YMR035W';'YMR036C';'YMR037C';'YMR038C';'YMR039C';'YMR040
W';'YMR041C';'YMR042W';'YMR043W';'YMR044W';'YMR047C';'YMR048W';'YMR049C';'YMR
052W';'YMR053C';'YMR054W';'YMR055C';'YMR056C';'YMR058W';'YMR059W';'YMR060C';'
YMR061W';'YMR062C';'YMR063W';'YMR064W';'YMR065W';'YMR066W';'YMR067C';'YMR068W
';'YMR069W';'YMR070W';'YMR071C';'YMR072W';'YMR073C';'YMR074C';'YMR075W';'YMR0
76C';'YMR077C';'YMR078C';'YMR079W';'YMR080C';'YMR081C';'YMR083W';'YMR086W';'Y
MR087W';'YMR088C';'YMR089C';'YMR091C';'YMR092C';'YMR093W';'YMR094W';'YMR095C'
;'YMR096W';'YMR097C';'YMR098C';'YMR099C';'YMR100W';'YMR101C';'YMR104C';'YMR10
5C';'YMR106C';'YMR107W';'YMR108W';'YMR109W';'YMR110C';'YMR112C';'YMR113W';'YM
R114C';'YMR115W';'YMR116C';'YMR117C';'YMR119W';'YMR120C';'YMR121C';'YMR123W';
'YMR125W';'YMR127C';'YMR128W';'YMR129W';'YMR131C';'YMR133W';'YMR135C';'YMR136
W';'YMR137C';'YMR138W';'YMR139W';'YMR140W';'YMR142C';'YMR143W';'YMR145C';'YMR
146C';'YMR148W';'YMR149W';'YMR150C';'YMR152W';'YMR153W';'YMR154C';'YMR156C';'
YMR157C';'YMR158W';'YMR159C';'YMR161W';'YMR162C';'YMR163C';'YMR164C';'YMR165C
';'YMR167W';'YMR168C';'YMR169C';'YMR170C';'YMR171C';'YMR172W';'YMR173W';'YMR1
74C';'YMR175W';'YMR176W';'YMR177W';'YMR179W';'YMR180C';'YMR182C';'YMR183C';'Y
MR184W';'YMR186W';'YMR188C';'YMR189W';'YMR190C';'YMR191W';'YMR192W';'YMR193W'
;'YMR194C-
B';'YMR194W';'YMR195W';'YMR197C';'YMR198W';'YMR199W';'YMR200W';'YMR201C';'YMR
202W';'YMR203W';'YMR204C';'YMR205C';'YMR207C';'YMR208W';'YMR210W';'YMR211W';'
YMR212C';'YMR213W';'YMR214W';'YMR215W';'YMR216C';'YMR217W';'YMR218C';'YMR219W
';'YMR220W';'YMR222C';'YMR223W';'YMR224C';'YMR225C';'YMR226C';'YMR227C';'YMR2
28W';'YMR229C';'YMR230W';'YMR231W';'YMR232W';'YMR233W';'YMR234W';'YMR235C';'Y
MR236W';'YMR237W';'YMR238W';'YMR239C';'YMR240C';'YMR241W';'YMR242C';'YMR243C'
;'YMR246W';'YMR247C';'YMR250W';'YMR251W';'YMR251W-
A';'YMR255W';'YMR256C';'YMR257C';'YMR258C';'YMR260C';'YMR261C';'YMR263W';'YMR
264W';'YMR266W';'YMR267W';'YMR268C';'YMR269W';'YMR270C';'YMR271C';'YMR272C';'
YMR273C';'YMR274C';'YMR275C';'YMR276W';'YMR277W';'YMR278W';'YMR279C';'YMR280C
';'YMR281W';'YMR282C';'YMR283C';'YMR284W';'YMR285C';'YMR286W';'YMR287C';'YMR2
88W';'YMR289W';'YMR290C';'YMR291W';'YMR292W';'YMR293C';'YMR294W';'YMR295C';'Y
MR296C';'YMR297W';'YMR298W';'YMR299C';'YMR300C';'YMR301C';'YMR302C';'YMR303C'
;'YMR304W';'YMR305C';'YMR306W';'YMR307W';'YMR308C';'YMR309C';'YMR311C';'YMR31
2W';'YMR313C';'YMR314W';'YMR315W';'YMR316W';'YMR318C';'YMR319C';'YMR323W';'YM
R325W';'YNL001W';'YNL002C';'YNL003C';'YNL004W';'YNL005C';'YNL006W';'YNL007C';
'YNL008C';'YNL009W';'YNL012W';'YNL014W';'YNL015W';'YNL016W';'YNL020C';'YNL021
W';'YNL023C';'YNL024C-
A';'YNL025C';'YNL026W';'YNL027W';'YNL029C';'YNL030W';'YNL031C';'YNL032W';'YNL
036W';'YNL037C';'YNL038W';'YNL039W';'YNL041C';'YNL042W';'YNL044W';'YNL045W';'
YNL047C';'YNL048W';'YNL049C';'YNL051W';'YNL052W';'YNL053W';'YNL054W';'YNL055C
';'YNL056W';'YNL059C';'YNL061W';'YNL062C';'YNL063W';'YNL064C';'YNL065W';'YNL0

66W';'YNL067W';'YNL068C';'YNL069C';'YNL070W';'YNL071W';'YNL072W';'YNL073W';'Y
NL074C';'YNL075W';'YNL076W';'YNL077W';'YNL078W';'YNL079C';'YNL080C';'YNL081C'
;'YNL082W';'YNL083W';'YNL084C';'YNL085W';'YNL087W';'YNL088W';'YNL090W';'YNL09
1W';'YNL093W';'YNL094W';'YNL096C';'YNL097C';'YNL098C';'YNL099C';'YNL100W';'YN
L101W';'YNL102W';'YNL103W';'YNL104C';'YNL106C';'YNL107W';'YNL110C';'YNL111C';
'YNL112W';'YNL113W';'YNL116W';'YNL117W';'YNL118C';'YNL119W';'YNL121C';'YNL123
W';'YNL124W';'YNL125C';'YNL126W';'YNL127W';'YNL128W';'YNL129W';'YNL130C';'YNL
131W';'YNL132W';'YNL133C';'YNL135C';'YNL136W';'YNL137C';'YNL138W';'YNL138W-
A';'YNL139C';'YNL141W';'YNL142W';'YNL145W';'YNL147W';'YNL148C';'YNL149C';'YNL
151C';'YNL152W';'YNL153C';'YNL154C';'YNL156C';'YNL157W';'YNL158W';'YNL159C';'
YNL160W';'YNL161W';'YNL162W';'YNL163C';'YNL164C';'YNL166C';'YNL167C';'YNL169C
';'YNL172W';'YNL173C';'YNL175C';'YNL177C';'YNL178W';'YNL180C';'YNL182C';'YNL1
83C';'YNL185C';'YNL186W';'YNL187W';'YNL188W';'YNL189W';'YNL191W';'YNL192W';'Y
NL194C';'YNL197C';'YNL199C';'YNL201C';'YNL202W';'YNL204C';'YNL206C';'YNL207W'
;'YNL208W';'YNL209W';'YNL210W';'YNL212W';'YNL213C';'YNL214W';'YNL215W';'YNL21
6W';'YNL218W';'YNL219C';'YNL220W';'YNL221C';'YNL222W';'YNL223W';'YNL224C';'YN
L225C';'YNL227C';'YNL229C';'YNL230C';'YNL231C';'YNL232W';'YNL233W';'YNL234W';
'YNL236W';'YNL237W';'YNL238W';'YNL239W';'YNL240C';'YNL241C';'YNL242W';'YNL243
W';'YNL244C';'YNL245C';'YNL246W';'YNL247W';'YNL248C';'YNL249C';'YNL250W';'YNL
251C';'YNL252C';'YNL253W';'YNL254C';'YNL255C';'YNL256W';'YNL257C';'YNL258C';'
YNL259C';'YNL261W';'YNL262W';'YNL263C';'YNL264C';'YNL265C';'YNL267W';'YNL268W
';'YNL269W';'YNL270C';'YNL271C';'YNL272C';'YNL273W';'YNL274C';'YNL275W';'YNL2
77W';'YNL278W';'YNL279W';'YNL280C';'YNL281W';'YNL282W';'YNL283C';'YNL284C';'Y
NL286W';'YNL287W';'YNL288W';'YNL289W';'YNL290W';'YNL291C';'YNL292W';'YNL293W'
;'YNL294C';'YNL297C';'YNL298W';'YNL299W';'YNL301C';'YNL302C';'YNL304W';'YNL30
5C';'YNL306W';'YNL307C';'YNL308C';'YNL309W';'YNL310C';'YNL311C';'YNL312W';'YN
L313C';'YNL314W';'YNL315C';'YNL316C';'YNL317W';'YNL318C';'YNL321W';'YNL322C';
'YNL323W';'YNL325C';'YNL326C';'YNL327W';'YNL328C';'YNL329C';'YNL330C';'YNL331
C';'YNL332W';'YNL333W';'YNL334C';'YNL336W';'YNL339C';'YNR001C';'YNR002C';'YNR
003C';'YNR006W';'YNR007C';'YNR008W';'YNR009W';'YNR010W';'YNR011C';'YNR012W';'
YNR013C';'YNR015W';'YNR016C';'YNR017W';'YNR018W';'YNR019W';'YNR020C';'YNR022C
';'YNR023W';'YNR024W';'YNR026C';'YNR027W';'YNR028W';'YNR030W';'YNR031C';'YNR0
32C-
A';'YNR032W';'YNR033W';'YNR034W';'YNR035C';'YNR036C';'YNR037C';'YNR038W';'YNR
039C';'YNR041C';'YNR043W';'YNR044W';'YNR045W';'YNR046W';'YNR047W';'YNR048W';'
YNR049C';'YNR050C';'YNR051C';'YNR052C';'YNR053C';'YNR054C';'YNR055C';'YNR056C
';'YNR057C';'YNR058W';'YNR059W';'YNR060W';'YNR064C';'YNR067C';'YNR069C';'YNR0
72W';'YNR074C';'YNR075W';'YNR076W';'YOL001W';'YOL002C';'YOL003C';'YOL004W';'Y
OL005C';'YOL006C';'YOL007C';'YOL008W';'YOL009C';'YOL010W';'YOL011W';'YOL012C'
;'YOL013C';'YOL015W';'YOL016C';'YOL017W';'YOL018C';'YOL020W';'YOL021C';'YOL02
2C';'YOL023W';'YOL025W';'YOL026C';'YOL027C';'YOL028C';'YOL030W';'YOL031C';'YO
L032W';'YOL033W';'YOL034W';'YOL038W';'YOL039W';'YOL040C';'YOL041C';'YOL042W';
'YOL043C';'YOL044W';'YOL045W';'YOL049W';'YOL051W';'YOL052C';'YOL052C-
A';'YOL053W';'YOL054W';'YOL055C';'YOL056W';'YOL057W';'YOL058W';'YOL059W';'YOL
060C';'YOL061W';'YOL062C';'YOL063C';'YOL064C';'YOL065C';'YOL066C';'YOL067C';'
YOL068C';'YOL069W';'YOL070C';'YOL071W';'YOL072W';'YOL076W';'YOL077C';'YOL077W
-
A';'YOL078W';'YOL080C';'YOL081W';'YOL082W';'YOL083W';'YOL084W';'YOL086C';'YOL
088C';'YOL089C';'YOL090W';'YOL091W';'YOL093W';'YOL094C';'YOL095C';'YOL096C';'
YOL097C';'YOL100W';'YOL101C';'YOL102C';'YOL103W';'YOL104C';'YOL105C';'YOL108C
';'YOL109W';'YOL110W';'YOL111C';'YOL112W';'YOL113W';'YOL115W';'YOL116W';'YOL1
17W';'YOL119C';'YOL120C';'YOL121C';'YOL122C';'YOL123W';'YOL124C';'YOL125W';'Y
OL126C';'YOL127W';'YOL128C';'YOL129W';'YOL130W';'YOL132W';'YOL133W';'YOL135C'
;'YOL136C';'YOL137W';'YOL138C';'YOL139C';'YOL140W';'YOL141W';'YOL142W';'YOL14
3C';'YOL144W';'YOL145C';'YOL146W';'YOL147C';'YOL148C';'YOL149W';'YOL151W';'YO
L152W';'YOL154W';'YOL155C';'YOL156W';'YOL157C';'YOL158C';'YOL159C';'YOL159C-
A';'YOL161C';'YOL164W';'YOL165C';'YOR001W';'YOR002W';'YOR003W';'YOR004W';'YOR

005C';'YOR006C';'YOR007C';'YOR008C';'YOR009W';'YOR010C';'YOR011W';'YOR014W';'
YOR016C';'YOR017W';'YOR018W';'YOR019W';'YOR020C';'YOR023C';'YOR025W';'YOR026W
';'YOR027W';'YOR028C';'YOR030W';'YOR031W';'YOR032C';'YOR033C';'YOR034C';'YOR0
35C';'YOR036W';'YOR037W';'YOR038C';'YOR039W';'YOR040W';'YOR042W';'YOR043W';'Y
OR044W';'YOR045W';'YOR046C';'YOR047C';'YOR048C';'YOR049C';'YOR051C';'YOR052C'
;'YOR054C';'YOR056C';'YOR057W';'YOR058C';'YOR060C';'YOR061W';'YOR063W';'YOR06
4C';'YOR065W';'YOR066W';'YOR067C';'YOR068C';'YOR069W';'YOR070C';'YOR071C';'YO
R073W';'YOR074C';'YOR075W';'YOR076C';'YOR077W';'YOR078W';'YOR079C';'YOR080W';
'YOR081C';'YOR083W';'YOR084W';'YOR085W';'YOR086C';'YOR087W';'YOR089C';'YOR090
C';'YOR091W';'YOR092W';'YOR094W';'YOR095C';'YOR096W';'YOR098C';'YOR099W';'YOR
100C';'YOR101W';'YOR103C';'YOR104W';'YOR106W';'YOR107W';'YOR108W';'YOR109W';'
YOR110W';'YOR112W';'YOR113W';'YOR115C';'YOR116C';'YOR117W';'YOR118W';'YOR119C
';'YOR120W';'YOR122C';'YOR123C';'YOR124C';'YOR125C';'YOR126C';'YOR127W';'YOR1
28C';'YOR129C';'YOR130C';'YOR132W';'YOR133W';'YOR134W';'YOR136W';'YOR137C';'Y
OR138C';'YOR140W';'YOR141C';'YOR142W';'YOR143C';'YOR144C';'YOR145C';'YOR147W'
;'YOR148C';'YOR149C';'YOR150W';'YOR151C';'YOR153W';'YOR155C';'YOR156C';'YOR15
7C';'YOR158W';'YOR159C';'YOR160W';'YOR161C';'YOR162C';'YOR163W';'YOR164C';'YO
R165W';'YOR166C';'YOR167C';'YOR168W';'YOR171C';'YOR172W';'YOR173W';'YOR174W';
'YOR175C';'YOR176W';'YOR177C';'YOR178C';'YOR179C';'YOR180C';'YOR181W';'YOR182
C';'YOR184W';'YOR185C';'YOR187W';'YOR188W';'YOR189W';'YOR190W';'YOR191W';'YOR
192C';'YOR193W';'YOR194C';'YOR195W';'YOR196C';'YOR197W';'YOR198C';'YOR201C';'
YOR202W';'YOR204W';'YOR205C';'YOR206W';'YOR207C';'YOR208W';'YOR209C';'YOR210W
';'YOR211C';'YOR212W';'YOR213C';'YOR215C';'YOR216C';'YOR217W';'YOR219C';'YOR2
21C';'YOR222W';'YOR223W';'YOR224C';'YOR226C';'YOR227W';'YOR228C';'YOR229W';'Y
OR230W';'YOR231W';'YOR232W';'YOR233W';'YOR234C';'YOR236W';'YOR237W';'YOR239W'
;'YOR241W';'YOR242C';'YOR243C';'YOR244W';'YOR245C';'YOR247W';'YOR249C';'YOR25
0C';'YOR251C';'YOR252W';'YOR253W';'YOR254C';'YOR255W';'YOR256C';'YOR257W';'YO
R258W';'YOR259C';'YOR260W';'YOR261C';'YOR262W';'YOR264W';'YOR265W';'YOR266W';
'YOR267C';'YOR269W';'YOR270C';'YOR272W';'YOR273C';'YOR274W';'YOR275C';'YOR276
W';'YOR278W';'YOR279C';'YOR280C';'YOR281C';'YOR283W';'YOR284W';'YOR285W';'YOR
286W';'YOR287C';'YOR288C';'YOR290C';'YOR291W';'YOR293W';'YOR294W';'YOR295W';'
YOR297C';'YOR298C-
A';'YOR298W';'YOR299W';'YOR301W';'YOR302W';'YOR303W';'YOR304W';'YOR305W';'YOR
306C';'YOR307C';'YOR308C';'YOR310C';'YOR311C';'YOR312C';'YOR313C';'YOR315W';'
YOR316C';'YOR317W';'YOR319W';'YOR320C';'YOR321W';'YOR322C';'YOR323C';'YOR324C
';'YOR326W';'YOR327C';'YOR328W';'YOR329C';'YOR330C';'YOR332W';'YOR334W';'YOR3
35C';'YOR336W';'YOR337W';'YOR339C';'YOR340C';'YOR341W';'YOR344C';'YOR346W';'Y
OR347C';'YOR348C';'YOR349W';'YOR350C';'YOR351C';'YOR353C';'YOR354C';'YOR355W'
;'YOR356W';'YOR357C';'YOR358W';'YOR359W';'YOR360C';'YOR361C';'YOR362C';'YOR36
3C';'YOR367W';'YOR368W';'YOR369C';'YOR370C';'YOR371C';'YOR372C';'YOR373W';'YO
R374W';'YOR375C';'YOR377W';'YOR380W';'YOR381W';'YOR382W';'YOR383C';'YOR384W';
'YOR386W';'YOR388C';'YOR391C';'YOR393W';'YOR394W';'YPL001W';'YPL002C';'YPL003
W';'YPL004C';'YPL005W';'YPL006W';'YPL007C';'YPL008W';'YPL009C';'YPL010W';'YPL
011C';'YPL012W';'YPL013C';'YPL015C';'YPL016W';'YPL017C';'YPL018W';'YPL019C';'
YPL020C';'YPL021W';'YPL022W';'YPL023C';'YPL024W';'YPL026C';'YPL027W';'YPL028W
';'YPL029W';'YPL030W';'YPL031C';'YPL032C';'YPL033C';'YPL036W';'YPL037C';'YPL0
38W';'YPL040C';'YPL042C';'YPL043W';'YPL045W';'YPL046C';'YPL047W';'YPL048W';'Y
PL049C';'YPL050C';'YPL051W';'YPL052W';'YPL053C';'YPL054W';'YPL055C';'YPL057C'
;'YPL058C';'YPL059W';'YPL060W';'YPL061W';'YPL063W';'YPL064C';'YPL065W';'YPL06
9C';'YPL070W';'YPL072W';'YPL074W';'YPL075W';'YPL076W';'YPL078C';'YPL079W';'YP
L081W';'YPL082C';'YPL083C';'YPL084W';'YPL085W';'YPL086C';'YPL087W';'YPL089C';
'YPL090C';'YPL091W';'YPL092W';'YPL093W';'YPL094C';'YPL095C';'YPL096C-
A';'YPL096W';'YPL097W';'YPL098C';'YPL099C';'YPL100W';'YPL101W';'YPL103C';'YPL
104W';'YPL105C';'YPL106C';'YPL110C';'YPL111W';'YPL112C';'YPL115C';'YPL116W';'
YPL117C';'YPL118W';'YPL119C';'YPL120W';'YPL121C';'YPL122C';'YPL123C';'YPL124W
';'YPL125W';'YPL126W';'YPL127C';'YPL128C';'YPL129W';'YPL130W';'YPL131W';'YPL1
32W';'YPL133C';'YPL134C';'YPL135W';'YPL137C';'YPL138C';'YPL139C';'YPL140C';'Y

```
PL141C';'YPL143W';'YPL144W';'YPL145C';'YPL146C';'YPL147W';'YPL148C';'YPL149W'
;'YPL151C';'YPL152W';'YPL153C';'YPL154C';'YPL155C';'YPL156C';'YPL157W';'YPL15
8C';'YPL159C';'YPL160W';'YPL161C';'YPL163C';'YPL164C';'YPL165C';'YPL166W';'YP
L167C';'YPL169C';'YPL170W';'YPL171C';'YPL172C';'YPL173W';'YPL174C';'YPL175W';
'YPL176C';'YPL177C';'YPL178W';'YPL179W';'YPL180W';'YPL181W';'YPL183C';'YPL183
W-A';'YPL184C';'YPL186C';'YPL187W';'YPL188W';'YPL189C-
A';'YPL189W';'YPL190C';'YPL192C';'YPL193W';'YPL194W';'YPL195W';'YPL196W';'YPL
198W';'YPL200W';'YPL201C';'YPL202C';'YPL203W';'YPL204W';'YPL206C';'YPL207W';'
YPL208W';'YPL209C';'YPL210C';'YPL211W';'YPL212C';'YPL213W';'YPL214C';'YPL215W
';'YPL217C';'YPL218W';'YPL219W';'YPL220W';'YPL221W';'YPL223C';'YPL224C';'YPL2
25W';'YPL226W';'YPL227C';'YPL228W';'YPL230W';'YPL231W';'YPL232W';'YPL233W';'Y
PL234C';'YPL235W';'YPL237W';'YPL239W';'YPL240C';'YPL241C';'YPL242C';'YPL243W'
;'YPL244C';'YPL246C';'YPL248C';'YPL249C';'YPL249C-
A';'YPL250C';'YPL252C';'YPL253C';'YPL254W';'YPL255W';'YPL256C';'YPL258C';'YPL
259C';'YPL262W';'YPL263C';'YPL265W';'YPL266W';'YPL267W';'YPL268W';'YPL269W';'
YPL270W';'YPL271W';'YPL273W';'YPL274W';'YPL281C';'YPL282C';'YPL283C';'YPR001W
';'YPR002W';'YPR004C';'YPR005C';'YPR006C';'YPR007C';'YPR008W';'YPR009W';'YPR0
10C';'YPR016C';'YPR017C';'YPR018W';'YPR019W';'YPR020W';'YPR021C';'YPR023C';'Y
PR024W';'YPR025C';'YPR026W';'YPR028W';'YPR029C';'YPR030W';'YPR031W';'YPR032W'
;'YPR033C';'YPR034W';'YPR035W';'YPR036W';'YPR036W-
A';'YPR037C';'YPR040W';'YPR041W';'YPR042C';'YPR043W';'YPR045C';'YPR046W';'YPR
047W';'YPR048W';'YPR049C';'YPR051W';'YPR052C';'YPR054W';'YPR055W';'YPR056W';'
YPR057W';'YPR058W';'YPR060C';'YPR061C';'YPR062W';'YPR065W';'YPR066W';'YPR067W
';'YPR068C';'YPR069C';'YPR070W';'YPR072W';'YPR073C';'YPR074C';'YPR075C';'YPR0
79W';'YPR080W';'YPR081C';'YPR082C';'YPR083W';'YPR085C';'YPR086W';'YPR088C';'Y
PR091C';'YPR093C';'YPR094W';'YPR095C';'YPR096C';'YPR097W';'YPR098C';'YPR100W'
;'YPR101W';'YPR102C';'YPR103W';'YPR104C';'YPR105C';'YPR106W';'YPR107C';'YPR10
8W';'YPR110C';'YPR111W';'YPR112C';'YPR113W';'YPR115W';'YPR116W';'YPR118W';'YP
R119W';'YPR120C';'YPR121W';'YPR122W';'YPR124W';'YPR125W';'YPR127W';'YPR128C';
'YPR129W';'YPR131C';'YPR132W';'YPR133C';'YPR133W-
A';'YPR134W';'YPR135W';'YPR137W';'YPR138C';'YPR139C';'YPR140W';'YPR141C';'YPR
143W';'YPR144C';'YPR145W';'YPR148C';'YPR149W';'YPR151C';'YPR152C';'YPR153W';'
YPR154W';'YPR155C';'YPR156C';'YPR158W';'YPR159W';'YPR160W';'YPR161C';'YPR162C
';'YPR163C';'YPR164W';'YPR165W';'YPR166C';'YPR167C';'YPR168W';'YPR169W';'YPR1
71W';'YPR173C';'YPR175W';'YPR176C';'YPR178W';'YPR179C';'YPR180W';'YPR181C';'Y
PR182W';'YPR183W';'YPR184W';'YPR185W';'YPR186C';'YPR187W';'YPR188C';'YPR189W'
;'YPR190C';'YPR191W';'YPR192W';'YPR193C';'YPR194C';'YPR198W';'YPR199C';'YPR20
0C';'YPR201W';'YPR204W';};

function genes = kuepfer_glucose_aerobic

% list of 62 genes found to be essential in defined media + glucose + aerobic
%Doi: 10.1101/gr.3992505

genes = {'YDR300C';'YJR122W';'YAL012W';'YER068C-
A';'YHR018C';'YAR015W';'YBL033C';'YBR115C';'YDR007W';'YDR127W';'YDR158W';'YDR
234W';'YDR354W';'YEL024W';'YER052C';'YER086W';'YER090W';'YGL026C';'YGL148W';'
YGL154C';'YGL234W';'YGR061C';'YGR204W';'YHR025W';'YIL094C';'YIR034C';'YJL088W
';'YJR139C';'YKL211C';'YLR304C';'YML022W';'YMR300C';'YNL220W';'YNL316C';'YOL0
58W';'YOL143C';'YOR128C';'YPR060C';'YBR209W';'YCR066W';'YDL072C';'YDL075W';'Y
DR008C';'YDR477W';'YFL032W';'YGR168C';'YHL005C';'YKR058W';'YPL006W';'YPR067W'
;'YGR155W';'YCR053W';'YNR050C';'YIL069C';'YKL135C';'YER057C';'YER069W';'YHL00
3C';'YNL218W';'YDR523C';'YOR184W';'YMR135W-A';};

function genes = kuepfer_galactose_aerobic

% list of 311 genes found to be essential in defined media + galactose +
```

```
% aerobic
%Doi: 10.1101/gr.3992505

genes = {'YDR300C';'YJR122W';'YAL012W';'YER068C-
A';'YAR015W';'YBL033C';'YBR115C';'YDR007W';'YDR127W';'YDR158W';'YDR234W';'YDR
354W';'YEL024W';'YER052C';'YER086W';'YER090W';'YGL026C';'YGL148W';'YGL154C';'
YGL234W';'YGR061C';'YGR204W';'YHR025W';'YIL094C';'YIR034C';'YJL088W';'YJR139C
';'YKL211C';'YLR304C';'YML022W';'YMR300C';'YNL220W';'YNL316C';'YOL058W';'YOL1
43C';'YOR128C';'YPR060C';'YBR209W';'YCR066W';'YDL072C';'YDL075W';'YDR008C';'Y
DR477W';'YFL032W';'YGR168C';'YHL005C';'YKR058W';'YPL006W';'YPR067W';'YGR155W'
;'YCR053W';'YNR050C';'YIL069C';'YKL135C';'YER057C';'YDR523C';'YOR184W';'YMR13
5W-
A';'YOR130C';'YHR177W';'YHR177W';'YBL045C';'YBL099W';'YBR003W';'YDL067C';'YDR
298C';'YDR529C';'YEL024W';'YER070W';'YGL012W';'YGR180C';'YGR255C';'YHR051W';'
YLR295C';'YML110C';'YMR256C';'YMR267W';'YNR041C';'YOL096C';'YOR065W';'YOR125C
';'YOR241W';'YPL059W';'YPL078C';'YPL078C';'YPL148C';'YPL271W';'YAL009W';'YAL0
39C';'YAL048C';'YBL002W';'YBL012C';'YBL022C';'YBL038W';'YBL090W';'YBL100C';'Y
BR037C';'YBR112C';'YBR120C';'YBR122C';'YBR132C';'YBR179C';'YBR251W';'YBR268W'
;'YBR282W';'YCR003W';'YCR004C';'YCR024C';'YCR028C-
A';'YCR046C';'YCR071C';'YDL044C';'YDL045W-
A';'YDL049C';'YDL057W';'YDL062W';'YDL063C';'YDL068W';'YDL069C';'YDL104C';'YDL
107W';'YDL113C';'YDL135C';'YDL146W';'YDL167C';'YDL181W';'YDL198C';'YDL202W';'
YDR042C';'YDR065W';'YDR078C';'YDR079W';'YDR114C';'YDR194C';'YDR197W';'YDR204W
';'YDR230W';'YDR231C';'YDR237W';'YDR268W';'YDR283C';'YDR295C';'YDR296W';'YDR3
22W';'YDR337W';'YDR507C';'YDR518W';'YEL029C';'YEL036C';'YEL050C';'YER017C';'Y
ER050C';'YER058W';'YER087W';'YER103W';'YER110C';'YER122C';'YER141W';'YER153C'
;'YER154W';'YER169W';'YFL016C';'YFL036W';'YGL038C';'YGL064C';'YGL095C';'YGL10
7C';'YGL129C';'YGL135W';'YGL218W';'YGL220W';'YGL240W';'YGR062C';'YGR076C';'YG
R102C';'YGR112W';'YGR150C';'YGR171C';'YGR215W';'YGR219W';'YGR220C';'YGR222W';
'YGR257C';'YHL038C';'YHR011W';'YHR038W';'YHR091C';'YHR116W';'YHR120W';'YHR147
C';'YHR168W';'YIR021W';'YJL003W';'YJL063C';'YJL096W';'YJL102W';'YJL180C';'YJL
209W';'YJR004C';'YJR144W';'YKL003C';'YKL134C';'YKL138C';'YKL155C';'YKL169C';'
YKL170W';'YKL194C';'YKR057W';'YKR061W';'YKR085C';'YLL018C-
A';'YLR067C';'YLR069C';'YLR091W';'YLR114C';'YLR139C';'YLR202C';'YLR203C';'YLR
204W';'YLR238W';'YLR255C';'YLR257W';'YLR312W-
A';'YLR369W';'YML061C';'YML088W';'YML129C';'YMR024W';'YMR064W';'YMR066W';'YMR
071C';'YMR072W';'YMR084W';'YMR089C';'YMR097C';'YMR098C';'YMR150C';'YMR151W';'
YMR158W';'YMR184W';'YMR193W';'YMR228W';'YMR257C';'YMR282C';'YMR286W';'YMR287C
';'YNL003C';'YNL005C';'YNL073W';'YNL081C';'YNL084C';'YNL160W';'YNL170W';'YNL1
77C';'YNL184C';'YNL213C';'YNL225C';'YNL252C';'YNL284C';'YNR036C';'YNR037C';'Y
OL023W';'YOL095C';'YOL100W';'YOR033C';'YOR150W';'YOR158W';'YOR199W';'YOR200W'
;'YOR201C';'YOR205C';'YOR211C';'YOR304C-
A';'YOR305W';'YOR318C';'YPL005W';'YPL013C';'YPL029W';'YPL031C';'YPL040C';'YPL
097W';'YPL104W';'YPL118W';'YPL132W';'YPL172C';'YPL173W';'YPL215W';'YPR047W';'
YPR072W';'YPR099C';'YPR100W';'YPR124W';'YPR166C';'YOR187W';'YKL016C';'YCR084C
';'YDR115W';'YDR175C';'YMR231W';'YPR116W';'YGL206C';'YGR262C';'YGL223C';'YLR3
82C';'YMR021C';'YBR018C';'YBR019C';'YBR020W';'YBR020W';'YDR009W';'YPL248C';};
```

Supplementary File 4.2: modelToReconstruction.m MATLAB script

```
function modelToReconstruction(filename)

% a function for converting an SBML model to a reconstruction, using SBO
% terms. reliant on the SBML toolbox.

% kieran smallbone and ben heavner: 18 aug 11

if ~strncmp(fliplr(filename),fliplr('.xml'),length('.xml'))
    filename = [filename,'.xml'];
end

m = TranslateSBML(filename);

% only real reactions
J = ismember([m.reaction.sboTerm],[395,397]);
m.reaction(J) = [];

% no modelling parameters
for k = 1:length(m.reaction)
    m.reaction(k).kineticLaw(1:end) = [];
end

% % no modelling notes
% for k = 1:length(m.speciesType)
%     m.speciesType(k).notes = '';
% end
% for k = 1:length(m.reaction)
%     m.reaction(k).notes = '';
% end

% find unused species
used_species = {};
for k = 1:length(m.reaction)
    used_species = union(used_species,{m.reaction(k).reactant.species});
    used_species = union(used_species,{m.reaction(k).product.species});
    used_species = union(used_species,{m.reaction(k).modifier.species});
end
J = ~ismember({m.species.id},used_species);
m.species(J) = [];

% ... and species types
J = ~ismember({m.speciesType.id},{m.species.speciesType});
m.speciesType(J) = [];

% ... and compartments
J = ~ismember({m.compartment.id},{m.species.compartment});
m.compartment(J) = [];

% save
filename = strrep(filename,'_model','');
filename = strrep(filename,'.xml','_recon.xml');
OutputSBML(m,filename);
```

```
% validate
[~,errors] = TranslateSBML(filename,1,0);
if isempty(errors)
    disp('Results: This document is valid SBML!');
else
    for k = 1:length(errors)
        fprintf('%s:\tline %g:\t(SBML Validation Rules #%g)\t%s',...

strtrim(errors(k).severity),errors(k).line,errors(k).errorId,errors(k).messag
e);
    end
end
```

Supplementary File 4.3: fluxDistribution.m MATLAB script

```
function fluxDistribution(filename)

model = readCbModel(filename,inf);

% define reactions of interest
rxnNames = {
    % glycolysis
    'glucose transport'
    'hexokinase (D-glucose:ATP)'
    'glucose-6-phosphate isomerase'
    'phosphofructokinase'
    'fructose-bisphosphate aldolase'
    'glyceraldehyde-3-phosphate dehydrogenase'
    'phosphoglycerate kinase'
    'phosphoglycerate mutase'
    'enolase'
    'pyruvate kinase'
    'pyruvate decarboxylase'
    %'alcohol dehydrogenase, reverse rxn (acetaldehyde -> ethanol)' % should
be 'alcohol dehydrogenase (ethanol)'
    'alcohol dehydrogenase (ethanol)'
    'ethanol transport'
    ' ' % TCA
    'pyruvate dehydrogenase'
    'citrate synthase'
    'citrate to cis-aconitate(3-)'
    'cis-aconitate(3-) to isocitrate'
    'isocitrate dehydrogenase (NAD+)'
    'oxoglutarate dehydrogenase (lipoamide)'
    'oxoglutarate dehydrogenase (dihydrolipoamide S-succinyltransferase)'
    'succinate-CoA ligase (ADP-forming)'
    'succinate dehydrogenase (ubiquinone-6)'
    'fumarase'
    'malate dehydrogenase'
    ' ' % PPP
    'glucose 6-phosphate dehydrogenase'
    '6-phosphogluconolactonase'
    'phosphogluconate dehydrogenase'
    'ribose-5-phosphate isomerase'
    'ribulose 5-phosphate 3-epimerase'
    'transketolase 1'
    'transaldolase'
    'transketolase 2'
    ' ' % other
    };

FBAsolution = optimizeCbModel(model,[],'one');
fprintf('aerobic growth\nrate:\t%.2f\n\n',FBAsolution.f);

for k = 1:length(rxnNames)
   ind = strcmp(rxnNames{k},model.rxnNames);
   if sum(ind)>1, disp(rxnNames{k}); end
```

```
   fprintf('%.2f\t%s\n',FBAsolution.x(ind),rxnNames{k});
end

ind = strcmp('oxygen exchange',model.rxnNames); model.lb(ind) = 0;

ind = ismember(model.rxnNames,{...
    'lipid pseudoreaction [no 14-demethyllanosterol, no ergosta-
5,7,22,24(28)-tetraen-3beta-ol]'
    'ergosterol exchange'
    'lanosterol exchange'
    'zymosterol exchange'
    'phosphatidate exchange'
    });
model.lb(ind) = -Inf;
model.ub(ind) = Inf;

ind = strcmp('lipid pseudoreaction',model.rxnNames); model.ub(ind) = 0;

FBAsolution = optimizeCbModel(model,[],'one');
fprintf('anaerobic growth\nrate:\t%.2f\n\n',FBAsolution.f);

for k = 1:length(rxnNames)
   ind = find(strcmp(rxnNames{k},model.rxnNames));
   fprintf('%.2f\t%s\n',FBAsolution.x(ind),rxnNames{k}); %#ok<FNDSB>
end
```

Supplementary Table 4.1: Essential Gene List

Essential ORFs in SGD:

YAL001C; YAL003W; YAL025C; YAL032C; YAL033W; YAL038W; YAL041W; YAL043C;
YAR007C; YAR008W; YAR019C; YBL004W; YBL014C; YBL018C; YBL020W; YBL023C;
YBL026W; YBL030C; YBL034C; YBL035C; YBL040C; YBL041W; YBL050W; YBL074C;
YBL076C; YBL084C; YBL092W; YBL097W; YBL105C; YBR002C; YBR004C; YBR011C;
YBR029C; YBR038W; YBR049C; YBR055C; YBR060C; YBR070C; YBR079C; YBR080C;
YBR087W; YBR088C; YBR091C; YBR102C; YBR109C; YBR110W; YBR123C; YBR135W;
YBR136W; YBR140C; YBR142W; YBR143C; YBR152W; YBR153W; YBR154C; YBR155W;
YBR160W; YBR167C; YBR192W; YBR193C; YBR196C; YBR198C; YBR202W; YBR211C;
YBR233W-A; YBR234C; YBR236C; YBR237W; YBR243C; YBR247C; YBR252W; YBR253W;
YBR254C; YBR256C; YBR257W; YBR265W; YCL004W; YCL017C; YCL031C; YCL043C;
YCL052C; YCL054W; YCL059C; YCR012W; YCR035C; YCR052W; YCR054C; YCR057C;
YCR072C; YCR093W; YDL003W; YDL004W; YDL007W; YDL008W; YDL014W; YDL015C;
YDL017W; YDL028C; YDL029W; YDL030W; YDL031W; YDL043C; YDL045C; YDL055C;
YDL058W; YDL060W; YDL064W; YDL084W; YDL087C; YDL092W; YDL097C; YDL098C;
YDL102W; YDL103C; YDL105W; YDL108W; YDL111C; YDL120W; YDL126C; YDL132W;
YDL139C; YDL140C; YDL141W; YDL143W; YDL145C; YDL147W; YDL148C; YDL150W;
YDL153C; YDL164C; YDL165W; YDL166C; YDL193W; YDL195W; YDL205C; YDL207W;
YDL208W; YDL209C; YDL212W; YDL217C; YDL220C; YDL235C; YDR002W; YDR013W;
YDR016C; YDR021W; YDR023W; YDR037W; YDR041W; YDR044W; YDR045C; YDR047W;
YDR050C; YDR052C; YDR054C; YDR060W; YDR062W; YDR064W; YDR081C; YDR082W;
YDR086C; YDR087C; YDR088C; YDR091C; YDR113C; YDR118W; YDR141C; YDR145W;
YDR160W; YDR164C; YDR166C; YDR167W; YDR168W; YDR170C; YDR172W; YDR177W;
YDR180W; YDR182W; YDR188W; YDR189W; YDR190C; YDR196C; YDR201W; YDR208W;
YDR211W; YDR212W; YDR224C; YDR228C; YDR232W; YDR235W; YDR236C; YDR238C;
YDR240C; YDR243C; YDR246W; YDR267C; YDR280W; YDR288W; YDR292C; YDR299W;
YDR301W; YDR302W; YDR303C; YDR308C; YDR311W; YDR320C-A; YDR324C; YDR325W;
YDR328C; YDR331W; YDR339C; YDR341C; YDR353W; YDR356W; YDR361C; YDR362C;
YDR365C; YDR367W; YDR373W; YDR376W; YDR381W; YDR390C; YDR394W; YDR397C;
YDR398W; YDR404C; YDR407C; YDR412W; YDR416W; YDR427W; YDR429C; YDR434W;
YDR437W; YDR449C; YDR454C; YDR460W; YDR464W; YDR468C; YDR472W; YDR473C;
YDR478W; YDR487C; YDR489W; YDR498C; YDR499W; YDR510W; YDR527W; YDR531W;
YEL002C; YEL019C; YEL026W; YEL032W; YEL034W; YEL055C; YEL058W; YER003C;
YER006W; YER008C; YER009W; YER012W; YER013W; YER018C; YER021W; YER022W;
YER023W; YER025W; YER029C; YER036C; YER038C; YER043C; YER048W-A; YER074W-A;
YER082C; YER093C; YER094C; YER104W; YER112W; YER125W; YER126C; YER127W;
YER133W; YER136W; YER146W; YER147C; YER148W; YER157W; YER159C; YER165W;
YER168C; YER171W; YER172C; YFL002C; YFL005W; YFL008W; YFL009W; YFL017C;
YFL022C; YFL024C; YFL029C; YFL037W; YFL038C; YFL039C; YFL045C; YFR002W;
YFR003C; YFR004W; YFR005C; YFR027W; YFR028C; YFR029W; YFR031C; YFR037C;
YFR042W; YFR050C; YFR051C; YFR052W; YGL001C; YGL008C; YGL011C; YGL018C;
YGL022W; YGL030W; YGL040C; YGL044C; YGL047W; YGL048C; YGL055W; YGL061C;
YGL065C; YGL068W; YGL073W; YGL075C; YGL091C; YGL092W; YGL093W; YGL097W;
YGL098W; YGL099W; YGL103W; YGL111W; YGL112C; YGL113W; YGL116W; YGL120C;
YGL122C; YGL123W; YGL128C; YGL130W; YGL137W; YGL142C; YGL145W; YGL150C;
YGL155W; YGL169W; YGL171W; YGL172W; YGL201C; YGL207W; YGL225W; YGL233W;
YGL238W; YGL245W; YGL247W; YGR002C; YGR005C; YGR009C; YGR013W; YGR024C;
YGR029W; YGR030C; YGR046W; YGR047C; YGR048W; YGR060W; YGR065C; YGR074W;
YGR075C; YGR082W; YGR083C; YGR090W; YGR091W; YGR094W; YGR095C; YGR098C;
YGR099W; YGR103W; YGR113W; YGR116W; YGR119C; YGR120C; YGR128C; YGR140W;
YGR145W; YGR147C; YGR156W; YGR158C; YGR172C; YGR175C; YGR179C; YGR185C;
YGR186W; YGR191W; YGR195W; YGR198W; YGR211W; YGR216C; YGR218W; YGR245C;
YGR246C; YGR251W; YGR253C; YGR264C; YGR267C; YGR274C; YGR277C; YGR278W;
YGR280C; YHL015W; YHR005C-A; YHR007C; YHR019C; YHR020W; YHR023W; YHR024C;
YHR036W; YHR040W; YHR042W; YHR058C; YHR062C; YHR065C; YHR068W; YHR069C;

YHR070W; YHR072W; YHR072W-A; YHR074W; YHR083W; YHR085W; YHR088W; YHR089C;
YHR101C; YHR102W; YHR107C; YHR118C; YHR122W; YHR128W; YHR143W-A; YHR148W;
YHR164C; YHR165C; YHR166C; YHR169W; YHR170W; YHR172W; YHR186C; YHR188C;
YHR190W; YHR196W; YHR197W; YHR199C-A; YIL003W; YIL004C; YIL019W; YIL021W;
YIL022W; YIL026C; YIL031W; YIL046W; YIL048W; YIL051C; YIL061C; YIL062C;
YIL063C; YIL068C; YIL075C; YIL078W; YIL083C; YIL091C; YIL104C; YIL106W;
YIL109C; YIL115C; YIL118W; YIL126W; YIL129C; YIL142W; YIL143C; YIL144W;
YIL147C; YIL150C; YIR006C; YIR008C; YIR010W; YIR011C; YIR012W; YIR015W;
YIR022W; YJL001W; YJL002C; YJL005W; YJL008C; YJL010C; YJL011C; YJL014W;
YJL019W; YJL025W; YJL026W; YJL031C; YJL033W; YJL034W; YJL035C; YJL039C;
YJL041W; YJL050W; YJL054W; YJL061W; YJL069C; YJL072C; YJL074C; YJL076W;
YJL081C; YJL085W; YJL087C; YJL090C; YJL091C; YJL097W; YJL104W; YJL109C;
YJL111W; YJL125C; YJL143W; YJL156C; YJL167W; YJL173C; YJL174W; YJL194W;
YJL203W; YJR002W; YJR006W; YJR007W; YJR013W; YJR016C; YJR017C; YJR022W;
YJR041C; YJR042W; YJR045C; YJR046W; YJR057W; YJR064W; YJR065C; YJR067C;
YJR068W; YJR072C; YJR076C; YJR089W; YJR093C; YJR112W; YJR123W; YKL004W;
YKL006C-A; YKL012W; YKL013C; YKL014C; YKL018W; YKL019W; YKL021C; YKL022C;
YKL024C; YKL028W; YKL033W; YKL035W; YKL042W; YKL045W; YKL049C; YKL052C;
YKL058W; YKL059C; YKL060C; YKL078W; YKL082C; YKL088W; YKL089W; YKL095W;
YKL099C; YKL104C; YKL108W; YKL112W; YKL122C; YKL125W; YKL138C-A; YKL141W;
YKL144C; YKL145W; YKL152C; YKL154W; YKL165C; YKL172W; YKL173W; YKL180W;
YKL182W; YKL186C; YKL189W; YKL192C; YKL193C; YKL195W; YKL196C; YKL203C;
YKL210W; YKR002W; YKR004C; YKR008W; YKR022C; YKR025W; YKR037C; YKR038C;
YKR062W; YKR063C; YKR068C; YKR071C; YKR079C; YKR081C; YKR083C; YKR086W;
YLL003W; YLL004W; YLL008W; YLL011W; YLL018C; YLL031C; YLL034C; YLL035W;
YLL036C; YLL050C; YLR002C; YLR005W; YLR007W; YLR008C; YLR009W; YLR010C;
YLR022C; YLR026C; YLR029C; YLR033W; YLR045C; YLR051C; YLR060W; YLR066W;
YLR071C; YLR075W; YLR078C; YLR086W; YLR088W; YLR100W; YLR103C; YLR105C;
YLR106C; YLR115W; YLR116W; YLR117C; YLR127C; YLR129W; YLR141W; YLR145W;
YLR147C; YLR153C; YLR163C; YLR166C; YLR167W; YLR175W; YLR186W; YLR195C;
YLR196W; YLR197W; YLR208W; YLR212C; YLR215C; YLR222C; YLR223C; YLR229C;
YLR249W; YLR259C; YLR272C; YLR274W; YLR275W; YLR276C; YLR277C; YLR291C;
YLR293C; YLR298C; YLR305C; YLR310C; YLR314C; YLR316C; YLR321C; YLR323C;
YLR336C; YLR340W; YLR347C; YLR355C; YLR359W; YLR378C; YLR383W; YLR397C;
YLR409C; YLR424W; YLR430W; YLR438C-A; YLR440C; YLR457C; YLR459W; YML010W;
YML015C; YML023C; YML025C; YML031W; YML043C; YML046W; YML049C; YML064C;
YML065W; YML069W; YML077W; YML085C; YML091C; YML092C; YML093W; YML098W;
YML105C; YML114C; YML125C; YML126C; YML127W; YML130C; YMR001C; YMR005W;
YMR013C; YMR028W; YMR033W; YMR043W; YMR047C; YMR049C; YMR059W; YMR061W;
YMR076C; YMR079W; YMR093W; YMR094W; YMR108W; YMR112C; YMR113W; YMR117C;
YMR128W; YMR131C; YMR146C; YMR149W; YMR168C; YMR197C; YMR200W; YMR203W;
YMR208W; YMR211W; YMR213W; YMR218C; YMR220W; YMR227C; YMR229C; YMR235C;
YMR236W; YMR239C; YMR240C; YMR260C; YMR268C; YMR270C; YMR277W; YMR281W;
YMR288W; YMR290C; YMR296C; YMR298W; YMR301C; YMR308C; YMR309C; YMR314W;
YNL002C; YNL006W; YNL007C; YNL024C-A; YNL026W; YNL036W; YNL038W; YNL039W;
YNL061W; YNL062C; YNL075W; YNL088W; YNL102W; YNL103W; YNL110C; YNL112W;
YNL113W; YNL118C; YNL124W; YNL126W; YNL131W; YNL132W; YNL137C; YNL138W-A;
YNL149C; YNL151C; YNL152W; YNL158W; YNL161W; YNL163C; YNL172W; YNL178W;
YNL182C; YNL188W; YNL189W; YNL207W; YNL216W; YNL221C; YNL222W; YNL232W;
YNL240C; YNL244C; YNL245C; YNL247W; YNL251C; YNL256W; YNL258C; YNL261W;
YNL262W; YNL263C; YNL267W; YNL272C; YNL282W; YNL287W; YNL290W; YNL306W;
YNL308C; YNL310C; YNL312W; YNL313C; YNL317W; YNR003C; YNR011C; YNR016C;
YNR017W; YNR026C; YNR035C; YNR038W; YNR043W; YNR046W; YNR053C; YNR054C;
YOL005C; YOL010W; YOL021C; YOL022C; YOL026C; YOL034W; YOL038W; YOL040C;
YOL066C; YOL069W; YOL077C; YOL078W; YOL094C; YOL097C; YOL102C; YOL120C;
YOL123W; YOL127W; YOL130W; YOL133W; YOL135C; YOL139C; YOL142W; YOL144W;
YOL146W; YOL149W; YOR004W; YOR020C; YOR046C; YOR048C; YOR056C; YOR057W;

YOR060C; YOR063W; YOR074C; YOR075W; YOR077W; YOR095C; YOR098C; YOR103C;
YOR110W; YOR116C; YOR117W; YOR119C; YOR122C; YOR143C; YOR145C; YOR148C;
YOR149C; YOR151C; YOR157C; YOR159C; YOR160W; YOR168W; YOR174W; YOR176W;
YOR181W; YOR194C; YOR204W; YOR206W; YOR207C; YOR210W; YOR217W; YOR224C;
YOR232W; YOR236W; YOR244W; YOR249C; YOR250C; YOR254C; YOR256C; YOR257W;
YOR259C; YOR260W; YOR261C; YOR262W; YOR272W; YOR278W; YOR281C; YOR287C;
YOR294W; YOR310C; YOR319W; YOR326W; YOR329C; YOR335C; YOR336W; YOR340C;
YOR341W; YOR353C; YOR361C; YOR362C; YOR370C; YOR372C; YOR373W; YPL007C;
YPL010W; YPL011C; YPL012W; YPL016W; YPL020C; YPL028W; YPL043W; YPL063W;
YPL076W; YPL082C; YPL083C; YPL085W; YPL093W; YPL094C; YPL117C; YPL122C;
YPL124W; YPL126W; YPL128C; YPL131W; YPL143W; YPL146C; YPL151C; YPL153C;
YPL160W; YPL169C; YPL175W; YPL190C; YPL204W; YPL209C; YPL210C; YPL211W;
YPL217C; YPL218W; YPL228W; YPL231W; YPL233W; YPL235W; YPL237W; YPL242C;
YPL243W; YPL252C; YPL255W; YPL266W; YPR010C; YPR016C; YPR019W; YPR025C;
YPR033C; YPR034W; YPR035W; YPR041W; YPR048W; YPR055W; YPR056W; YPR082C;
YPR085C; YPR086W; YPR088C; YPR094W; YPR103W; YPR104C; YPR105C; YPR107C;
YPR108W; YPR110C; YPR112C; YPR113W; YPR133C; YPR137W; YPR143W; YPR144C;
YPR161C; YPR162C; YPR165W; YPR168W; YPR169W; YPR175W; YPR176C; YPR178W;
YPR180W; YPR181C; YPR182W; YPR183W; YPR186C; YPR187W; YPR190C

Essential ORFs in Yeast 5.0:

YAL038W; YBL030C; YBL076C; YBR002C; YBR004C; YBR011C; YBR029C; YBR038W;
YBR110W; YBR153W; YBR192W; YBR196C; YBR252W; YBR256C; YBR265W; YCL004W;
YCL052C; YCR012W; YDL004W; YDL015C; YDL045C; YDL055C; YDL103C; YDL141W;
YDL205C; YDR023W; YDR037W; YDR044W; YDR047W; YDR050C; YDR062W; YDR196C;
YDR208W; YDR232W; YDR236C; YDR302W; YDR341C; YDR353W; YDR367W; YDR376W;
YDR434W; YDR454C; YDR487C; YDR531W; YEL058W; YER003C; YER023W; YER043C;
YFL017C; YFL022C; YFL045C; YGL001C; YGL008C; YGL040C; YGL055W; YGL142C;
YGL155W; YGL225W; YGL245W; YGR060W; YGR065C; YGR094W; YGR175C; YGR185C;
YGR191W; YGR264C; YGR267C; YGR277C; YHR007C; YHR019C; YHR020W; YHR042W;
YHR068W; YHR072W; YHR074W; YHR128W; YHR188C; YHR190W; YIL078W; YIL083C;
YJL005W; YJL026W; YJL091C; YJL097W; YJL167W; YJR013W; YJR016C; YJR057W;
YKL004W; YKL019W; YKL024C; YKL035W; YKL060C; YKL088W; YKL104C; YKL141W;
YKL152C; YKL165C; YKL182W; YKL192C; YLL018C; YLL031C; YLR060W; YLR088W;
YLR100W; YLR153C; YLR195C; YLR305C; YLR355C; YLR359W; YLR459W; YML126C;
YMR013C; YMR108W; YMR113W; YMR208W; YMR220W; YMR281W; YMR296C; YMR298W;
YNL247W; YNL256W; YNL267W; YNR016C; YNR043W; YOL066C; YOL097C; YOR074C;
YOR095C; YOR143C; YOR168W; YOR176W; YOR236W; YOR278W; YOR335C; YPL028W;
YPL117C; YPL160W; YPL231W; YPL252C; YPR033C; YPR035W; YPR113W; YPR183W

Supplementary Table 4.2: SGD Auxotroph List

SGD Auxotroph list	Auxotrophy as annotated by SGD		Selected Auxotroph List		Auxotroph ORFs in Yeast 5
YAL012W	cysteine		YAL021C		YAR015W
YAL021C	myo-inositol		YAL024C		YBL033C
YAL023C	myo-inositol		YAL026C		YBL098W
YAL024C	myo-inositol		YAL040C		YBR115C
YAL026C	myo-inositol		YAL051W		YBR126C
YAL040C	myo-inositol		YAL056W		YBR127C
YAL051W	myo-inositol		YAL058W		YBR176W
YAL056W	myo-inositol		YAR003W		YBR248C
YAL058W	myo-inositol		YAR015W		YCL018W
YAR003W	myo-inositol		YAR069W-A		YCL030C
YAR015W	adenine		YAR070W-A		YCR053W
YAR015W	adenine		YBL027W		YDL131W
YAR069W-A	biotin		YBL033C		YDR007W
YAR069W-A	biotin		YBL047C		YDR074W
YAR070W-A	biotin		YBL058W		YDR173C
YBL011W	myo-inositol		YBL061C		YDR226W
YBL027W	myo-inositol		YBL091C-A		YDR354W
YBL033C	riboflavin		YBL098W		YEL021W
YBL047C	myo-inositol		YBL102W		YEL027W
YBL058W	myo-inositol		YBL103C		YEL051W
YBL061C	myo-inositol		YBR015C		YER026C
YBL091C-A	myo-inositol		YBR058C		YER052C
YBL098W	nicotinic acid		YBR077C		YER055C
YBL102W	myo-inositol		YBR106W		YER069W
YBL103C	glutamate(1-)		YBR107C		YER090W
YBR015C	myo-inositol		YBR115C		YER091C
YBR015C			YBR126C		YFR019W
YBR026C	myo-inositol		YBR127C		YFR025C
YBR058C	myo-inositol		YBR133C		YFR047C
YBR077C	myo-inositol		YBR175W		YGL009C
YBR106W	myo-inositol		YBR176W		YGL012W
YBR107C	myo-inositol		YBR189W		YGL026C
YBR115C	lysine		YBR191W		YGL062W
YBR115C	lysine		YBR248C		YGL154C
YBR115C	lysine		YBR272C		YGL234W
YBR115C	lysine		YBR279W		YGR020C
YBR115C	lysine		YCL018W		YGR061C
YBR126C	myo-inositol		YCL030C		YGR204W
YBR127C	myo-inositol		YCL032W		YGR227W
YBR133C	myo-inositol		YCL033C		YHR018C
YBR135W	myo-inositol		YCL045C		YHR025W

YBR175W	myo-inositol		YCR021C		YHR026W
YBR176W	pantothenate		YCR045C		YHR208W
YBR176W	pantothenic acid		YCR047C		YIL020C
YBR176W	pantothenic acid (0.1 mM)		YCR053W		YIL116W
YBR183W	myo-inositol		YCR076C		YIR034C
YBR189W	myo-inositol		YCR089W		YJL088W
YBR191W	myo-inositol		YCR094W		YJL130C
YBR213W	methionine		YDL001W		YJL153C
YBR221C	myo-inositol		YDL002C		YKL001C
YBR248C	histidine		YDL006W		YKL211C
YBR252W	dTMP		YDL010W		YKL212W
YBR256C	riboflavin		YDL020C		YKL216W
YBR272C	myo-inositol		YDL021W		YLR420W
YBR279W	myo-inositol		YDL033C		YML008C
YCL018W	leucine		YDL040C		YMR062C
YCL030C	histidine		YDL048C		YMR202W
YCL032W	myo-inositol		YDL069C		YMR217W
YCL033C	myo-inositol		YDL073W		YMR300C
YCL045C	myo-inositol		YDL074C		YNL003C
YCR021C	myo-inositol		YDL077C		YNL220W
YCR045C	myo-inositol		YDL081C		YNL241C
YCR047C	myo-inositol		YDL083C		YNL277W
YCR053W	L-threonine		YDL106C		YNL316C
YCR053W	L-threonine		YDL130W		YOL143C
YCR053W	threonine		YDL131W		YOR128C
YCR076C	myo-inositol		YDL173W		YOR202W
YCR089W	myo-inositol		YDL190C		YPR036W
YCR094W	myo-inositol		YDL191W		YPR060C
YCR098C	myo-inositol		YDL192W		YPR167C
YDL001W	myo-inositol		YDL194W		
YDL002C	myo-inositol		YDL201W		
YDL006W	myo-inositol		YDL203C		
YDL010W	myo-inositol		YDR007W		
YDL020C	myo-inositol		YDR043C		
YDL021W	myo-inositol		YDR049W		
YDL022W	myo-inositol		YDR057W		
YDL033C	myo-inositol		YDR071C		
YDL040C	myo-inositol		YDR074W		
YDL048C	myo-inositol		YDR080W		
YDL066W	myo-inositol		YDR138W		
YDL069C	myo-inositol		YDR162C		
YDL073W	myo-inositol		YDR173C		
YDL074C	myo-inositol		YDR174W		
YDL077C	myo-inositol		YDR176W		
YDL081C	myo-inositol		YDR200C		
YDL083C	myo-inositol		YDR226W		
YDL095W	myo-inositol		YDR260C		

YDL106C	adenine		YDR266C		
YDL130W	myo-inositol		YDR276C		
YDL131W	lysine		YDR277C		
YDL140C	inositol		YDR283C		
YDL140C	inositol		YDR289C		
YDL173W	myo-inositol		YDR335W		
YDL185W	myo-inositol		YDR346C		
YDL190C	myo-inositol		YDR348C		
YDL191W	myo-inositol		YDR351W		
YDL192W	myo-inositol		YDR354W		
YDL194W	myo-inositol		YDR358W		
YDL201W	myo-inositol		YDR363W		
YDL203C	myo-inositol		YDR379W		
YDL205C	heme		YDR385W		
YDL205C	heme		YDR389W		
YDR007W	tryptophan		YDR392W		
YDR007W			YDR395W		
YDR017C	myo-inositol		YDR411C		
YDR043C	myo-inositol		YDR422C		
YDR047W	heme		YDR432W		
YDR049W	myo-inositol		YDR439W		
YDR050C	myo-inositol		YDR448W		
YDR057W	myo-inositol		YDR469W		
YDR071C	myo-inositol		YDR477W		
YDR074W	myo-inositol		YDR482C		
YDR080W	myo-inositol		YDR486C		
YDR138W	myo-inositol		YDR540C		
YDR162C	myo-inositol		YEL004W		
YDR173C	arginine		YEL013W		
YDR173C	arginine		YEL021W		
YDR173C	myo-inositol		YEL027W		
YDR173C	ornithine		YEL029C		
YDR174W	myo-inositol		YEL031W		
YDR176W	myo-inositol		YEL037C		
YDR200C	myo-inositol		YEL040W		
YDR226W	myo-inositol		YEL044W		
YDR232W	cysteine		YEL048C		
YDR232W	heme		YEL051W		
YDR260C	myo-inositol		YER001W		
YDR266C	myo-inositol		YER007C-A		
YDR276C	myo-inositol		YER026C		
YDR277C	myo-inositol		YER027C		
YDR283C	myo-inositol		YER052C		
YDR289C	myo-inositol		YER055C		
YDR294C	myo-inositol		YER059W		
YDR297W	myo-inositol		YER069W		
YDR300C	myo-inositol		YER083C		

YDR335W	myo-inositol		YER090W		
YDR346C	myo-inositol		YER091C		
YDR348C	myo-inositol		YER092W		
YDR351W	myo-inositol		YER095W		
YDR354W	myo-inositol		YER101C		
YDR354W	tryptophan		YER116C		
YDR358W	myo-inositol		YER118C		
YDR363W	myo-inositol		YER120W		
YDR368W	myo-inositol		YER122C		
YDR379W	myo-inositol		YER129W		
YDR384C	myo-inositol		YER130C		
YDR385W	myo-inositol		YER149C		
YDR389W	myo-inositol		YER150W		
YDR392W	myo-inositol		YER152C		
YDR395W	myo-inositol		YER167W		
YDR399W	myo-inositol		YER169W		
YDR400W	myo-inositol		YER177W		
YDR400W			YFL013C		
YDR403W	myo-inositol		YFL031W		
YDR408C	adenine		YFR010W		
YDR411C	myo-inositol		YFR019W		
YDR422C	myo-inositol		YFR025C		
YDR432W	myo-inositol		YFR040W		
YDR439W	myo-inositol		YFR047C		
YDR448W	myo-inositol		YFR048W		
YDR469W	myo-inositol		YGL009C		
YDR477W	inositol		YGL012W		
YDR482C	myo-inositol		YGL020C		
YDR486C	myo-inositol		YGL025C		
YDR487C	riboflavin		YGL026C		
YDR540C	myo-inositol		YGL031C		
YEL004W	myo-inositol		YGL049C		
YEL013W	myo-inositol		YGL054C		
YEL021W	uracil		YGL058W		
YEL027W	myo-inositol		YGL060W		
YEL029C	myo-inositol		YGL062W		
YEL031W	myo-inositol		YGL066W		
YEL037C	myo-inositol		YGL070C		
YEL040W	myo-inositol		YGL115W		
YEL041W	myo-inositol		YGL126W		
YEL044W	inositol		YGL127C		
YEL048C	myo-inositol		YGL154C		
YEL051W	myo-inositol		YGL167C		
YEL063C	L-arginine		YGL168W		
YER001W			YGL175C		
YER003C	D-mannose		YGL179C		
YER007C-A	myo-inositol		YGL180W		

YER026C	choline		YGL181W		
YER026C	choline		YGL203C		
YER027C	myo-inositol		YGL211W		
YER052C	homoserine		YGL219C		
YER055C	histidine		YGL234W		
YER059W	myo-inositol		YGL244W		
YER069W	arginine		YGR014W		
YER081W	myo-inositol		YGR020C		
YER083C	myo-inositol		YGR056W		
YER086W	isoleucine		YGR057C		
YER090W	tryptophan		YGR061C		
YER091C	methionine		YGR063C		
YER092W	myo-inositol		YGR092W		
YER095W	myo-inositol		YGR104C		
YER101C	myo-inositol		YGR105W		
YER116C	myo-inositol		YGR108W		
YER118C	myo-inositol		YGR135W		
YER120W	inositol		YGR144W		
YER120W	myo-inositol		YGR162W		
YER120W	myo-inositol		YGR166W		
YER120W	myo-inositol		YGR204W		
YER120W	myo-inositol		YGR223C		
YER120W	myo-inositol		YGR227W		
YER122C	myo-inositol		YGR229C		
YER129W	myo-inositol		YGR241C		
YER130C	myo-inositol		YGR250C		
YER149C	myo-inositol		YGR252W		
YER150W	myo-inositol		YHR010W		
YER152C	myo-inositol		YHR013C		
YER167W	myo-inositol		YHR018C		
YER169W	myo-inositol		YHR025W		
YER174C	myo-inositol		YHR026W		
YER177W	myo-inositol		YHR030C		
YFL013C	myo-inositol		YHR038W		
YFL031W	inositol		YHR060W		
YFL031W	myo-inositol		YHR079C		
YFL031W	myo-inositol		YHR111W		
YFR010W	myo-inositol		YHR142W		
YFR019W	myo-inositol		YHR162W		
YFR025C	histidine		YHR178W		
YFR030W	methionine		YHR179W		
YFR030W	methionine		YHR199C		
YFR030W	methionine		YHR206W		
YFR030W	methionine		YHR208W		
YFR040W	myo-inositol		YIL017C		
YFR047C	nicotinic acid		YIL020C		
YFR048W	myo-inositol		YIL027C		

YGL009C	leucine		YIL029C		
YGL012W	myo-inositol		YIL036W		
YGL020C	myo-inositol		YIL044C		
YGL025C	myo-inositol		YIL072W		
YGL026C	tryptophan		YIL077C		
YGL026C			YIL105C		
YGL031C	myo-inositol		YIL116W		
YGL040C	cysteine		YIL119C		
YGL040C	heme		YIL128W		
YGL049C	myo-inositol		YIL153W		
YGL054C	myo-inositol		YIR017C		
YGL055W	oleic acid		YIR034C		
YGL055W	oleic acid		YJL088W		
YGL058W	myo-inositol		YJL095W		
YGL060W	myo-inositol		YJL115W		
YGL062W	myo-inositol		YJL117W		
YGL066W	myo-inositol		YJL128C		
YGL070C	inositol		YJL130C		
YGL070C	myo-inositol		YJL140W		
YGL115W	inositol		YJL151C		
YGL125W	methionine		YJL153C		
YGL125W	methionine		YJL158C		
YGL125W	methionine		YJL172W		
YGL126W	inositol		YJL192C		
YGL127C	myo-inositol		YJL193W		
YGL150C	inositol		YJL204C		
YGL150C	inositol		YJL208C		
YGL150C	myo-inositol		YJR066W		
YGL154C	lysine		YJR075W		
YGL154C	lysine		YJR083C		
YGL167C	inositol		YJR104C		
YGL167C	myo-inositol		YJR122W		
YGL168W	myo-inositol		YKL001C		
YGL175C	myo-inositol		YKL006W		
YGL179C	myo-inositol		YKL027W		
YGL180W	myo-inositol		YKL032C		
YGL181W	myo-inositol		YKL048C		
YGL203C	myo-inositol		YKL053C-A		
YGL211W	myo-inositol		YKL056C		
YGL219C	myo-inositol		YKL064W		
YGL234W	adenine		YKL077W		
YGL234W	adenine		YKL079W		
YGL244W	myo-inositol		YKL119C		
YGR014W	myo-inositol		YKL121W		
YGR020C	myo-inositol		YKL160W		
YGR056W	myo-inositol		YKL176C		
YGR057C	myo-inositol		YKL190W		

YGR060W	ergosterol		YKL211C		
YGR061C	adenine		YKL212W		
YGR063C	myo-inositol		YKL213C		
YGR092W	myo-inositol		YKL216W		
YGR096W	thiamine		YKR007W		
YGR104C	myo-inositol		YKR026C		
YGR105W	myo-inositol		YKR036C		
YGR108W	myo-inositol		YKR070W		
YGR135W	myo-inositol		YKR099W		
YGR144W	thiamine(1+)		YLL019C		
YGR144W	thiamine(1+)		YLL021W		
YGR155W	cysteine		YLL027W		
YGR157W	choline		YLL039C		
YGR162W	myo-inositol		YLR015W		
YGR166W	myo-inositol		YLR016C		
YGR175C	ergosterol		YLR021W		
YGR204W	adenine;histidine		YLR048W		
YGR208W	serine		YLR055C		
YGR208W	serine		YLR061W		
YGR223C	myo-inositol		YLR074C		
YGR227W	myo-inositol		YLR079W		
YGR229C	myo-inositol		YLR087C		
YGR241C	myo-inositol		YLR113W		
YGR250C	myo-inositol		YLR150W		
YGR252W	myo-inositol		YLR192C		
YGR264C	methionine		YLR199C		
YGR267C	5-formyltetrahydrofolic acid (aka folinic acid)		YLR226W		
YGR286C	biotin		YLR242C		
YHR007C	ergosterol		YLR262C		
YHR010W	myo-inositol		YLR268W		
YHR013C	myo-inositol		YLR292C		
YHR018C			YLR315W		
YHR025W	L-threonine		YLR320W		
YHR025W	L-threonine		YLR324W		
YHR026W	myo-inositol		YLR332W		
YHR030C	inositol		YLR357W		
YHR038W	myo-inositol		YLR371W		
YHR060W	myo-inositol		YLR373C		
YHR072W	ergosterol		YLR396C		
YHR079C	inositol		YLR417W		
YHR079C	myo-inositol		YLR418C		
YHR104W	myo-inositol		YLR420W		
YHR111W	myo-inositol		YLR426W		
YHR142W	myo-inositol		YLR436C		
YHR143W-A	inositol		YML008C		

YHR162W	myo-inositol		YML013W		
YHR178W	myo-inositol		YML014W		
YHR179W	myo-inositol		YML028W		
YHR190W	ergosterol		YML034W		
YHR199C	myo-inositol		YML055W		
YHR206W	myo-inositol		YML071C		
YHR208W	valine;leucine		YML103C		
YIL017C	myo-inositol		YML115C		
YIL020C	histidine		YML117W		
YIL027C	myo-inositol		YMR010W		
YIL029C	myo-inositol		YMR014W		
YIL036W	myo-inositol		YMR016C		
YIL044C	myo-inositol		YMR029C		
YIL051C	isoleucine		YMR038C		
YIL072W	myo-inositol		YMR052W		
YIL077C	myo-inositol		YMR062C		
YIL105C	myo-inositol		YMR067C		
YIL116W	histidine		YMR068W		
YIL119C	myo-inositol		YMR092C		
YIL128W	methionine		YMR099C		
YIL145C	pantothenate		YMR104C		
YIL153W	myo-inositol		YMR123W		
YIR017C	methionine		YMR165C		
YIR034C	lysine		YMR190C		
YJL071W	arginine		YMR202W		
YJL087C	inositol		YMR214W		
YJL088W	arginine		YMR217W		
YJL095W	inositol		YMR242C		
YJL101C	glutathione		YMR247C		
YJL115W	myo-inositol		YMR276W		
YJL117W	myo-inositol		YMR300C		
YJL121C	myo-inositol		YMR304W		
YJL128C	myo-inositol		YMR307W		
YJL130C	uracil		YMR312W		
YJL134W	myo-inositol		YNL003C		
YJL140W	inositol		YNL041C		
YJL151C	myo-inositol		YNL051W		
YJL153C	inositol		YNL079C		
YJL153C	inositol		YNL080C		
YJL155C	myo-inositol		YNL119W		
YJL158C	myo-inositol		YNL127W		
YJL172W	leucine		YNL133C		
YJL192C	myo-inositol		YNL148C		
YJL193W	myo-inositol		YNL215W		
YJL204C	myo-inositol		YNL219C		
YJL208C	myo-inositol		YNL220W		
YJL212C	myo-inositol		YNL229C		

YJL214W	myo-inositol		YNL236W		
YJR010W	methionine		YNL241C		
YJR016C	valine;isoleucine		YNL277W		
YJR017C	galactose		YNL307C		
YJR017C	inositol		YNL316C		
YJR017C	phosphate		YNL322C		
YJR025C	nicotinic acid		YNR055C		
YJR025C	nicotinic acid		YOL018C		
YJR066W	myo-inositol		YOL067C		
YJR075W	myo-inositol		YOL087C		
YJR078W	nicotinic acid		YOL090W		
YJR083C	myo-inositol		YOL093W		
YJR104C	lysine		YOL098C		
YJR104C	lysine		YOL107W		
YJR104C	methionine		YOL109W		
YJR104C	methionine		YOL111C		
YJR109C	arginine		YOL116W		
YJR122W	glutamic acid		YOL121C		
YJR122W	lysine		YOL122C		
YJR137C	methionine		YOL124C		
YJR137C	methionine		YOL143C		
YJR137C	methionine		YOL145C		
YJR137C	methionine		YOR002W		
YKL001C	methionine		YOR008C		
YKL006W	myo-inositol		YOR012W		
YKL027W	myo-inositol		YOR067C		
YKL032C	myo-inositol		YOR070C		
YKL048C	myo-inositol		YOR078W		
YKL053C-A	myo-inositol		YOR096W		
YKL056C	myo-inositol		YOR106W		
YKL064W	myo-inositol		YOR123C		
YKL077W	myo-inositol		YOR128C		
YKL079W	myo-inositol		YOR189W		
YKL104C	D-glucosamine		YOR202W		
YKL119C	myo-inositol		YOR216C		
YKL120W	leucine		YOR246C		
YKL121W	myo-inositol		YOR290C		
YKL160W	myo-inositol		YOR320C		
YKL176C	myo-inositol		YOR322C		
YKL182W	fatty acid		YOR359W		
YKL182W	fatty acid		YOR371C		
YKL184W	beta-alanine (0.1 mM)		YPL055C		
YKL184W	myo-inositol		YPL065W		
YKL184W	pantothenic acid (0.1 mM)		YPL089C		
YKL184W	putrescine (0.25 mM)		YPL138C		

YKL184W	spermidine (10 uM)		YPL140C		
YKL184W	spermine (0.25 mM)		YPL144W		
YKL190W	myo-inositol		YPL157W		
YKL211C	tryptophan		YPL159C		
YKL212W	myo-inositol		YPL174C		
YKL212W	myo-inositol		YPL177C		
YKL212W	myo-inositol		YPL226W		
YKL213C	myo-inositol		YPL241C		
YKL216W	uracil		YPL254W		
YKR007W	myo-inositol		YPL264C		
YKR026C	myo-inositol		YPR036W		
YKR036C	myo-inositol		YPR043W		
YKR069W	methionine		YPR060C		
YKR070W	myo-inositol		YPR067W		
YKR080W	adenine		YPR139C		
YKR099W	adenine		YPR167C		
YLL015W	myo-inositol		YPR173C		
YLL019C	myo-inositol		YPR179C		
YLL021W	myo-inositol		YPR201W		
YLL027W	glutamate(1-)				
YLL027W	lysine				
YLL039C	myo-inositol				
YLL043W	myo-inositol				
YLR015W	myo-inositol				
YLR016C	myo-inositol				
YLR021W	myo-inositol				
YLR048W	myo-inositol				
YLR055C	myo-inositol				
YLR056W	myo-inositol				
YLR061W	myo-inositol				
YLR074C	myo-inositol				
YLR079W	myo-inositol				
YLR087C	myo-inositol				
YLR089C	myo-inositol				
YLR100W	cholesterol				
YLR100W	ergosterol				
YLR113W	myo-inositol				
YLR134W	myo-inositol				
YLR146C	beta-alanine (0.1 mM)				
YLR146C	pantothenic acid (0.1 mM)				
YLR146C	spermine (0.25 mM)				
YLR150W	myo-inositol				
YLR192C	myo-inositol				
YLR199C	myo-inositol				

YLR208W	myo-inositol				
YLR226W	myo-inositol				
YLR226W	myo-inositol				
YLR231C	nicotinic acid				
YLR237W	myo-inositol				
YLR242C	myo-inositol				
YLR262C	myo-inositol				
YLR268W	myo-inositol				
YLR292C	myo-inositol				
YLR303W	methionine				
YLR303W	methionine				
YLR304C	glutamic acid				
YLR315W	myo-inositol				
YLR320W	myo-inositol				
YLR324W	myo-inositol				
YLR332W	myo-inositol				
YLR342W	myo-inositol				
YLR354C	myo-inositol				
YLR355C	valine;isoleucine				
YLR357W	myo-inositol				
YLR359W					
YLR359W					
YLR371W	myo-inositol				
YLR373C	myo-inositol				
YLR396C	methionine				
YLR417W	ethanolamine				
YLR418C	myo-inositol				
YLR420W	uracil				
YLR426W	myo-inositol				
YLR436C	myo-inositol				
YML008C	myo-inositol				
YML013W	myo-inositol				
YML014W	myo-inositol				
YML022W	myo-inositol				
YML028W	myo-inositol				
YML034W	myo-inositol				
YML055W	myo-inositol				
YML059C	inositol				
YML071C	myo-inositol				
YML103C	myo-inositol				
YML106W					
YML115C	myo-inositol				
YML117W	myo-inositol				
YML126C	ergosterol				
YMR010W	myo-inositol				
YMR014W	myo-inositol				
YMR015C	myo-inositol				

YMR016C	myo-inositol				
YMR020W	beta-alanine (0.1 mM)				
YMR020W	pantothenic acid (0.1 mM)				
YMR029C	myo-inositol				
YMR038C	methionine;lysine				
YMR052W	myo-inositol				
YMR062C	arginine				
YMR067C	myo-inositol				
YMR068W	myo-inositol				
YMR079W	myo-inositol				
YMR092C	myo-inositol				
YMR099C	myo-inositol				
YMR104C	myo-inositol				
YMR108W	isoleucine;valine				
YMR113W	5-formyltetrahydrofolic acid				
YMR113W	5-formyltetrahydrofolic acid				
YMR123W	myo-inositol				
YMR165C	myo-inositol				
YMR170C	beta-alanine				
YMR170C	pantothenic acid				
YMR190C	myo-inositol				
YMR202W	myo-inositol				
YMR208W	ergosterol				
YMR214W	myo-inositol				
YMR217W	guanine				
YMR220W	ergosterol				
YMR242C	myo-inositol				
YMR247C	inositol				
YMR272C	myo-inositol				
YMR276W	myo-inositol				
YMR289W	4-aminobenzoic acid				
YMR289W	folic acid				
YMR300C	adenine				
YMR304W	myo-inositol				
YMR307W	myo-inositol				
YMR312W	myo-inositol				
YNL003C	biotin				
YNL041C	myo-inositol				
YNL051W	myo-inositol				
YNL079C	myo-inositol				
YNL080C	myo-inositol				

YNL103W	methionine				
YNL104C	leucine				
YNL119W	myo-inositol				
YNL127W	myo-inositol				
YNL133C	myo-inositol				
YNL148C	myo-inositol				
YNL215W	myo-inositol				
YNL219C	myo-inositol				
YNL220W					
YNL220W					
YNL229C	myo-inositol				
YNL236W	myo-inositol				
YNL241C	methionine				
YNL241C	methionine				
YNL241C	methionine				
YNL256W	folic acid				
YNL267W	myo-inositol				
YNL277W	methionine				
YNL277W	methionine				
YNL307C	myo-inositol				
YNL316C					
YNL322C	myo-inositol				
YNL323W	myo-inositol				
YNR055C	histidinol				
YOL018C	myo-inositol				
YOL052C	beta-alanine (0.1 mM)				
YOL052C	myo-inositol				
YOL052C	pantothenic acid (0.1 mM)				
YOL052C	spermine (0.25 mM)				
YOL059W	myo-inositol				
YOL064C	methionine				
YOL066C	riboflavin				
YOL067C	glutamate(1-)				
YOL067C	myo-inositol				
YOL068C	myo-inositol				
YOL087C	myo-inositol				
YOL090W	myo-inositol				
YOL093W	myo-inositol				
YOL098C	myo-inositol				
YOL107W	myo-inositol				
YOL109W	myo-inositol				
YOL110W	myo-inositol				
YOL111C	myo-inositol				
YOL116W	myo-inositol				
YOL121C	myo-inositol				

YOL122C	myo-inositol				
YOL124C	myo-inositol				
YOL143C	riboflavin				
YOL145C	myo-inositol				
YOR002W	myo-inositol				
YOR008C	myo-inositol				
YOR012W	myo-inositol				
YOR067C	myo-inositol				
YOR070C	myo-inositol				
YOR078W	myo-inositol				
YOR096W	myo-inositol				
YOR106W	myo-inositol				
YOR123C	myo-inositol				
YOR128C	adenine				
YOR151C	inositol				
YOR184W	serine				
YOR184W	serine				
YOR189W	myo-inositol				
YOR202W	histidine				
YOR210W	inositol				
YOR216C	myo-inositol				
YOR224C	inositol				
YOR236W	adenine				
YOR236W	dTMP				
YOR236W	histidine				
YOR236W	methionine				
YOR241W	methionine				
YOR241W	methionine				
YOR246C	myo-inositol				
YOR270C	myo-inositol				
YOR278W	heme				
YOR290C	inositol				
YOR290C	myo-inositol				
YOR303W	arginine				
YOR320C	myo-inositol				
YOR322C	myo-inositol				
YOR359W	myo-inositol				
YOR371C	myo-inositol				
YPL028W	ergosterol				
YPL055C	myo-inositol				
YPL065W	myo-inositol				
YPL089C	myo-inositol				
YPL138C	myo-inositol				
YPL140C	myo-inositol				
YPL144W	myo-inositol				
YPL157W	myo-inositol				
YPL159C	myo-inositol				

YPL174C	myo-inositol				
YPL177C	myo-inositol				
YPL214C	thiamine				
YPL226W	myo-inositol				
YPL231W	fatty acid				
YPL241C	myo-inositol				
YPL254W	myo-inositol				
YPL264C	myo-inositol				
YPL268W	ornithine				
YPR035W	glutamine				
YPR036W	myo-inositol				
YPR043W	myo-inositol				
YPR056W	inositol				
YPR060C	phenylalanine				
YPR067W	glutamate(1-)				
YPR067W	lysine				
YPR069C	beta-alanine (0.1 mM)				
YPR069C	pantothenic acid (0.1 mM)				
YPR069C	spermidine (10 uM)				
YPR069C	spermine (0.25 mM)				
YPR074C	phenylalanine				
YPR074C	tryptophan				
YPR074C	tyrosine				
YPR139C	myo-inositol				
YPR161C	myo-inositol				
YPR161C	myo-inositol				
YPR167C	myo-inositol				
YPR173C	ethanolamine				
YPR179C	myo-inositol				
YPR187W	inositol				
YPR187W					
YPR201W	myo-inositol				

Notes:

The selected auxotroph list removes the following ORFS: those annotated "unverified" in SGD, those which are also on the YKO essential list, those annotated in SGD as temperature-sensitive ino auxotrophs, and those not auxotrophs in the genome functional profiling study doi:10.1038/nature00935

Auxotrophy annotation was downloaded from the Saccharomyces Genome Database. Chemical auxotrophies including concentration information indicate the concentration of that compound in media, as described by White, Gunyuzlu, and Toyn (2001) doi: 10.1074/jbc.M009804200

Supplementary Table 4.3: Biomass Definition

rxn	substrate		product	
	stoichiometric coefficient	metabolite name	stoichiometric coefficient	metabolite name
growth	-1	biomass		
biomass pseudoreaction	-1.14	(1->3)-beta-D-glucan [cytoplasm]	59.3	ADP [cytoplasm]
	-0.051	AMP [cytoplasm]	1	biomass [cytoplasm]
	-59.3	ATP [cytoplasm]	58.7	H+ [cytoplasm]
	-0.05	CMP [cytoplasm]	59.3	phosphate [cytoplasm]
	-0.00359	dAMP [cytoplasm]		
	-0.00243	dCMP [cytoplasm]		
	-0.00243	dGMP [cytoplasm]		
	-0.00359	dTMP [cytoplasm]		
	-0.519	glycogen [cytoplasm]		
	-0.051	GMP [cytoplasm]		
	-59.3	H2O [cytoplasm]		
	-0.357	L-alanine [cytoplasm]		
	-0.136	L-arginine [cytoplasm]		
	-0.172	L-asparagine [cytoplasm]		
	-0.172	L-aspartate [cytoplasm]		
	-0.0429	L-cysteine [cytoplasm]		
	-0.0268	L-glutamate [cytoplasm]		
	-0.0268	L-glutamine [cytoplasm]		
	-0.325	L-glycine [cytoplasm]		
	-0.075	L-histidine [cytoplasm]		
	-0.0172	L-isoleucine [cytoplasm]		
	-0.25	L-leucine [cytoplasm]		
	-0.239	L-lysine [cytoplasm]		
	-0.05	L-methionine [cytoplasm]		
	-0.114	L-phenylalanine [cytoplasm]		
	-0.129	L-proline [cytoplasm]		
	-0.254	L-serine [cytoplasm]		
	-0.197	L-threonine [cytoplasm]		
	-0.028	L-tryptophan [cytoplasm]		
	-0.0965	L-tyrosine [cytoplasm]		
	-0.0257	L-valine [cytoplasm]		
	-1	lipid [cytoplasm]		
	-0.821	mannan [cytoplasm]		
	-0.0009	riboflavin [cytoplasm]		
	-0.02	sulphate [cytoplasm]		
	-0.0234	trehalose [cytoplasm]		

	-0.067	UMP [cytoplasm]		
lipid pseudoreaction	-0.00153	1-phosphatidyl-1D-myo-inositol [cytoplasm]	1	lipid [cytoplasm]
	-0.000056	14-demethyllanosterol [cytoplasm]		
(aerobic)	-0.000417	complex sphingolipid [cytoplasm]		
	-0.000096	episterol [cytoplasm]		
	-0.000125	ergosta-5,7,22,24(28)-tetraen-3beta-ol [cytoplasm]		
	-0.0056	ergosterol [cytoplasm]		
	-0.000812	ergosterol ester [cytoplasm]		
	-0.000206	fatty acid [cytoplasm]		
	-0.000114	fecosterol [cytoplasm]		
	-0.000032	lanosterol [cytoplasm]		
	-0.000373	phosphatidyl-L-serine [cytoplasm]		
	-0.00288	phosphatidylcholine [cytoplasm]		
	-0.000697	phosphatidylethanolamine [cytoplasm]		
	-0.000781	triglyceride [cytoplasm]		
	-0.000015	zymosterol [cytoplasm]		
lipid pseudoreaction	-0.00153	1-phosphatidyl-1D-myo-inositol [cytoplasm]	1	lipid [cytoplasm]
	-0.000417	complex sphingolipid [cytoplasm]		
(anaerobic)	-0.000096	episterol [cytoplasm]		
	-0.0056	ergosterol [cytoplasm]		
	-0.000812	ergosterol ester [cytoplasm]		
	-0.000206	fatty acid [cytoplasm]		
	-0.000114	fecosterol [cytoplasm]		
	-0.000032	lanosterol [cytoplasm]		
	-0.000373	phosphatidyl-L-serine [cytoplasm]		
	-0.00288	phosphatidylcholine [cytoplasm]		
	-0.000697	phosphatidylethanolamine [cytoplasm]		
	-0.000781	triglyceride [cytoplasm]		
	-0.000015	zymosterol [cytoplasm]		

Supplementary Table 4.4: Epistasis Analysis

Multiplicative Epistasis cutoff threshold = 0.01

Yeast 5 Gene-Based Analysis

Negative epistasis

Flux Restriction	0 (KO)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0 (KO)	5566	5435	5313	5111	4843	4700	4466	4197	3800	3512
0.1		6152	6190	5947	5518	5558	5421	4855	4688	4369
0.2			5614	5515	5433	5356	5371	4726	4587	3629
0.3				5396	5363	5243	5067	4682	4514	3505
0.4					5226	4998	4772	4457	4366	2874
0.5						4672	4469	4219	3755	2081
0.6							4122	4098	3560	1414
0.7								3892	2213	966
0.8									1198	452
0.9										242

% Positive Epistasis

Flux Restriction	0 (KO)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0 (KO)	42.63	75.83	76.39	77.58	79.67	80.2	81.12	82.6	84.28	84.34
0.1		91.05	91.07	91.03	91.77	91.23	91.29	92.13	92.13	90.36
0.2			91.94	91.97	92.2	92.2	92.26	92.83	92.46	93.55
0.3				92.04	92.13	92.37	92.5	93.12	92.96	93.57
0.4					92.21	92.5	92.92	93.31	93.4	94.67
0.5						92.89	93.15	93.6	94.13	96.04
0.6							93.55	93.56	94.19	97.2
0.7								93.8	96.3	98.2
0.8									97.93	99.18
0.9										99.54

iMM904 Gene-Based Analysis

Negative epistasis

Flux Restriction	0 (KO)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0 (KO)	3490	4422	4349	4180	4033	3944	4020	3867	3525	3241
0.1		6910	7079	6938	6762	6500	6704	6446	6003	5592
0.2			6614	6623	6376	6240	6573	6010	5830	5521
0.3				6390	6111	6024	6003	5754	5647	4899
0.4					5762	5855	5897	5666	5627	4200
0.5						5554	5654	5519	5482	3311
0.6							5374	5485	5279	2867
0.7								5096	4619	2371
0.8									2642	1960
0.9										1656

% Positive Epistasis

Flux Restriction	0 (KO)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0 (KO)	47.53	77.44	78.2	79.12	79.81	80.82	80.13	80.66	81.97	82.81
0.1		89.57	89.38	89.53	89.74	89.44	88.57	88.81	89.36	86.5
0.2			90.04	89.98	90.38	90.41	89.88	90.67	89.86	90.03
0.3				90.31	90.74	90.67	90.57	90.96	91.03	91.08
0.4					91.17	90.95	90.7	91.04	91.03	92.04
0.5						91.21	90.92	91.1	91.15	94.02
0.6							91.19	90.97	91.28	94.91
0.7								91.47	92.17	95.74
0.8									95.34	96.43
0.9										96.91

iND750 Gene-Based Analysis

Negative epistasis

Flux Restriction	0 (KO)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0 (KO)	2836	3709	3613	3581	3575	3529	3487	3403	3373	3270
0.1		4668	4750	4704	4727	4436	4411	4353	4325	4373
0.2			4350	4403	4405	4353	4338	4298	4217	3991
0.3				4276	4352	4319	4282	4254	4187	3211
0.4					4246	4296	4282	4231	4189	2397
0.5						4162	4229	4197	4146	1879
0.6							4118	4185	3787	1626
0.7								4010	2147	702
0.8									1006	350
0.9										136

% Positive Epistasis

Flux Restriction	0 (KO)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0 (KO)	46.51	77.4	78.71	79.4	79.3	79.69	79.8	80.17	80.46	80.73
0.1		91.9	91.85	91.96	91.52	91.4	91.41	91.57	91.6	84.71
0.2			92.53	92.44	92.43	92.51	92.51	92.18	91.74	91.88
0.3				92.68	92.57	92.56	92.62	92.64	92.3	93.51
0.4					92.67	92.65	92.66	92.67	92.68	95.02
0.5						92.8	92.7	92.72	92.75	96.03
0.6							92.86	92.73	93.26	96.76
0.7								92.99	96.04	98.58
0.8									98.11	99.32
0.9										99.73

Yeast 5 Reaction-Based Analysis

Negative epistasis

Flux Restriction	0 (KO)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0 (KO)	862	909	894	862	802	761	743	696	667	641
0.1		874	900	857	814	792	784	734	695	669
0.2			830	835	792	763	769	713	690	664
0.3				764	763	723	705	677	646	614
0.4					672	650	640	618	599	566
0.5						574	595	566	543	510
0.6							510	524	508	451
0.7								458	473	406
0.8									378	373
0.9										316

% Positive Epistasis

Flux Restriction	0 (KO)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0 (KO)	27.56	73.58	73.95	77.69	81.41	83.57	83.68	86.36	87.35	87.33
0.1		96.6	96.55	96.69	96.9	96.94	96.96	97.17	97.26	96.52
0.2			96.82	96.76	96.98	97.03	97.08	97.25	97.32	97.33
0.3				96.99	97.03	97.19	97.23	97.39	97.47	97.45
0.4					97.33	97.41	97.47	97.56	97.66	97.65
0.5						97.68	97.58	97.73	97.81	97.83
0.6							97.88	97.83	97.87	98
0.7								98.09	98.01	98.2
0.8									98.39	98.35
0.9										98.54

iMM904 Reaction-Based Analysis

Negative epistasis

Flux Restriction	0 (KO)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0 (KO)	262	321	316	307	303	288	280	260	244	218
0.1		346	360	355	350	329	329	302	286	256
0.2			346	345	341	328	325	282	270	251
0.3				324	320	317	298	269	254	232
0.4					302	302	295	252	243	212
0.5						270	273	252	237	196
0.6							246	239	229	184
0.7								200	199	159
0.8									146	142
0.9										110

% Positive Epistasis

Flux Restriction	0 (KO)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0 (KO)	22.02	75.15	76.7	78.33	79.6	80.39	80.54	81.72	82.6	83.68
0.1		95.63	95.45	95.51	95.61	95.71	95.49	95.84	95.97	94.91
0.2			95.54	95.59	95.68	95.91	95.9	96.32	96.26	96.42
0.3				95.88	95.97	95.97	96.19	96.58	96.67	96.75
0.4					96.16	96.19	96.2	96.76	96.78	96.99
0.5						96.55	96.5	96.73	96.94	97.3
0.6							96.77	96.88	96.98	97.43
0.7								97.35	97.34	97.75
0.8									98.02	98.03
0.9										98.4

iND750 Reaction-Based Analysis

Negative epistasis

Flux Restriction	0 (KO)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0 (KO)	276	347	340	333	334	337	335	333	317	301
0.1		374	388	375	380	376	374	376	357	342
0.2			366	366	369	365	361	352	338	327
0.3				334	349	353	346	341	324	299
0.4					334	344	342	336	320	282
0.5						316	333	333	311	257
0.6							310	323	308	232
0.7								294	254	213
0.8									194	179
0.9										136

% Positive Epistasis

Flux Restriction	0 (KO)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0 (KO)	25.41	77.63	78.91	81.15	80.95	81.45	81.49	81.57	82.96	83.47
0.1		96.33	96.19	96.34	96.22	96.13	96.12	96.12	96.21	91.62
0.2			96.39	96.39	96.35	96.35	96.36	96.42	96.44	96.35
0.3				96.7	96.55	96.47	96.54	96.58	96.66	96.77
0.4					96.64	96.55	96.57	96.6	96.73	96.98
0.5						96.74	96.6	96.6	96.81	97.23
0.6							96.8	96.67	96.81	97.54
0.7								96.96	97.34	97.71
0.8									97.97	98.09
0.9										98.54

Supplementary Table 4.5: ORFs in Y4 not in Y5

ORFs in Yeast 4 not in Yeast 5	Gene name (from SGD)	ORF type (from SGD)	ORF function (from SGD)	Comment
YAL014C	SYN8	ORF, Verified	Endosomal SNARE related to mammalian syntaxin 8	Y4 annotates a palmitoyl transferase reaction with this ORF
YAL030W	SNC1	ORF, Verified	Vesicle membrane receptor protein (v-SNARE) involved in the fusion between Golgi-derived secretory vesicles with the plasma membrane	Y4 annotates a palmitoyl transferase reaction with this ORF
YAR042W	SWH1	ORF, Verified	Protein similar to mammalian oxysterol-binding protein	Y4 annotates a sterol intercompartmental transport reaction with this ORF
YCR073W-A	SOL2	ORF, Verified	Protein with a possible role in tRNA export; shows similarity to 6-phosphogluconolactonase non-catalytic domains but does not exhibit this enzymatic activity	Y4 annotates a 6-phosphogluconolactonase reaction with this ORF
YDL019C	OSH2	ORF, Verified	Member of an oxysterol-binding protein family with seven members in S. cerevisiae;	Y4 annotates a sterol intercompartmental transport reaction with this ORF; metabolic function unclear
YDR313C	PIB1	ORF, Verified	RING-type ubiquitin ligase of the endosomal and vacuolar membranes, binds phosphatidylinositol(3)-phosphate;	Y4 includes this binding; out of Y5 scope

YDR331W	GPI8	ORF, Verified	ER membrane glycoprotein subunit of the glycosylphosphatidylinositol transamidase complex	GPI-anchor assembly lumped reaction in Y4
YDR468C	TLG1	ORF, Verified	Essential t-SNARE that forms a complex with Tlg2p and Vti1p and mediates fusion of endosome-derived vesicles with the late Golgi	Y4 annotates a palmitoyl transferase reaction with this ORF
YEL011W	GLC3	ORF, Verified	Glycogen branching enzyme, involved in glycogen accumulation	Y4 annotates a reaction using the polysaccharide amylose as a substrate; a general reaction
YEL013W	VAC8	ORF, Verified	Phosphorylated and palmitoylated vacuolar membrane protein that interacts with Atg13p	Y4 annotates a palmitoylation reaction with this ORF; metabolic function unknown
YER093C	TSC11	ORF, Verified	Subunit of TORC2 (Tor2p-Lst8p-Avo1-Avo2-Tsc11p-Bit61p), a membrane-associated complex that regulates actin cytoskeletal dynamics during polarized growth and cell wall integrity	Y4 annotates a reaction which only uses boundary species with this ORF
YFR055W	IRC7	ORF, Uncharacterized	Beta-lyase involved in the production of thiols	Y4 annotates lyase with this ORF or YGL184C; YGL184C annotation retained in Y5
YGR199W	PMT6	ORF, Verified	Protein O-mannosyltransferase	PMT6 wasn't among the genes analyzed in the reference paper linked from SGD; Y4 annotates this reaction with the other genes in this family
YHR005C	GPA1	ORF, Verified	GTP-binding alpha subunit of the heterotrimeric G protein that couples to pheromone receptors	Y4 annotates a YHR005C myristoylation reaction using boundary species with this ORF

YHR073W	OSH3	ORF, Verified	Member of an oxysterol-binding protein family with seven members in S. cerevisiae	Y4 annotates an oxysterol transport and binding reaction with this ORF
YMR068W	AVO2	ORF, Verified	Component of a complex containing the Tor2p kinase and other proteins, which may have a role in regulation of cell growth	Y4 annotates a reaction which only uses boundary species with this ORF
YNL006W	LST8	ORF, Verified	Protein required for the transport of amino acid permease Gap1p from the Golgi to the cell surface;	Y4 annotates a reaction which only uses boundary species with this ORF
YNR034W	SOL1	ORF, Verified	Protein with a possible role in tRNA export; shows similarity to 6-phosphogluconolactonase non-catalytic domains but does not exhibit this enzymatic activity	Y4 annotates a 6-phosphogluconolactonase reaction with this ORF
YOL078W	AVO1	ORF, Verified	Component of a membrane-bound complex containing the Tor2p kinase and other proteins, which may have a role in regulation of cell growth	Y4 annotates a reaction which only uses boundary species with this ORF
YPL145C	KES1	ORF, Verified	Member of the oxysterol binding protein family	Y4 annotates two oxysterol transport and binding reactions with this ORF
YPL275W	FDH2	pseudogene	NAD(+)-dependent formate dehydrogenase, may protect cells from exogenous formate; YPL275W and YPL276W comprise a continuous open reading frame in some S. cerevisiae strains but not in the genomic reference strain S288C	Y4 annotates a formate dehydrogenase reaction with this ORF

YPL276W	FDH2	pseudogene	NAD(+)-dependent formate dehydrogenase, may protect cells from exogenous formate; YPL275W and YPL276W comprise a continuous open reading frame in some S. cerevisiae strains but not in the genomic reference strain S288C	Y4 annotates a formate dehydrogenase reaction with this ORF
YOR237W	HES1	ORF, Verified	Protein implicated in the regulation of ergosterol biosynthesis	Y4 annotates an oxysterol transport and binding reaction with this ORF
YIL105C	SLM1	ORF, Verified	Phosphoinositide PI4,5P(2) binding protein, forms a complex with Slm2p	Y4 includes this binding; out of Y5 scope
YJL058C	BIT61	ORF, Verified	Subunit of TORC2 (Tor2p-Lst8p-Avo1-Avo2-Tsc11p-Bit61p-Slm1p-Slm2p), a membrane-associated complex that regulates	Y4 annotates a reaction which only uses boundary species with this ORF
YJR160C	MPH3	ORF, Verified	Alpha-glucoside permease, transports maltose, maltotriose, alpha-methylglucoside, and turanose;	Replaced by YDL247W in Y5 (YJR160C reintroduced in subsequent updates)
YKL203C	TOR2	ORF, Verified	PIK-related protein kinase and rapamycin target	Y4 annotates a reaction which only uses boundary species with this ORF
YML082W	-	ORF, Uncharacterized	Putative protein predicted to have carbon-sulfur lyase activity	Y4 annotates a cystathionine gamma-synthase reaction and a O-succinylhomoserine lyase with this ORF
YNL047C	SLM2	ORF, Verified	Phosphoinositide PI4,5P(2) binding protein	Y4 includes this binding; out of Y5 scope

Supplementary Table 4.6: ORFs in Y5 not in Y4

ORFs in Yeast 5 not in Yeast 4	Gene name (from SGD)	ORF type (from SGD)	ORF function (from SGD)
YBR001C	NTH2	ORF, Verified	Putative neutral trehalase, required for thermotolerance and may mediate resistance to other cellular stresses
YBR058C-A	TSC3	ORF, Verified	Protein that stimulates the activity of serine palmitoyltransferase (Lcb1p, Lcb2p) several-fold; involved in sphingolipid biosynthesis
YBR161W	CSH1	ORF, Verified	Probable catalytic subunit of a mannosylinositol phosphorylceramide (MIPC) synthase, forms a complex with probable regulatory subunit Csg2p; function in sphingolipid biosynthesis is overlapping with that of Sur1p
YBR199W	KTR4	ORF, Verified	Putative mannosyltransferase involved in protein glycosylation; member of the KRE2/MNT1 mannosyltransferase family
YDR196C	CAB5	ORF, Verified	Probable dephospho-CoA kinase (DPCK) that catalyzes the last step in coenzyme A biosynthesis; null mutant lethality is complemented by E. coli coaE (encoding DPCK); detected in purified mitochondria in high-throughput studies
YDR367W	KEI1	ORF, Verified	Component of inositol phosphorylceramide (IPC) synthase; forms a complex with Aur1p and regulates its activity; required for IPC synthase complex localization to the Golgi; post-translationally processed by Kex2p; KEI1 is an essential gene
YGL084C	GUP1	ORF, Verified	Plasma membrane protein involved in remodeling GPI anchors; member of the MBOAT family of putative membrane-bound O-acyltransferases; proposed to be involved in glycerol transport
YGR138C	TPO2	ORF, Verified	Polyamine transport protein specific for spermine; localizes to the plasma membrane; transcription of TPO2 is regulated by Haa1p; member of the major facilitator superfamily
YGR277C	CAB4	ORF, Verified	Probable pantetheine-phosphate adenylyltransferase (PPAT), which catalyzes the fourth step in the biosynthesis of coenzyme A from pantothenate; null mutant lethality is complemented by E. coli coaD (encoding PPAT); widely conserved

YIL083C	CAB2	ORF, Verified	Probable phosphopantothenoylcysteine synthetase (PPCS), which catalyzes the second step of coenzyme A biosynthesis from pantothenate; null mutant lethality is complemented by E. coli coaBC (encoding a bifunctional enzyme with PPCS activity)
YJL200C	ACO2	ORF, Verified	Putative mitochondrial aconitase isozyme; similarity to Aco1p, an aconitase required for the TCA cycle; expression induced during growth on glucose, by amino acid starvation via Gcn4p, and repressed on ethanol
YKL088W	CAB3	ORF, Verified	Subunit of a phosphopantothenoylcysteine decarboxylase (PPCDC; Cab3p, Sis2p, Vhs3p) complex, which catalyzes the third step of coenzyme A biosynthesis; null mutant lethality is complemented by E. coli coaBC
YKL132C	RMA1	ORF, Verified	Putative dihydrofolate synthetase; has similarity to Fol3p and to E. coli folylpolyglutamate synthetase/dihydrofolate synthetase; the authentic, non-tagged protein is detected in highly purified mitochondria in high-throughput studies
YML056C	IMD4	ORF, Verified	Inosine monophosphate dehydrogenase, catalyzes the first step of GMP biosynthesis, member of a four-gene family in S. cerevisiae, constitutively expressed
YMR241W	YHM2	ORF, Verified	Carrier protein that exports citrate from and imports oxoglutarate into the mitochondrion, causing net export of NADPH reducing equivalents; also associates with mt nucleoids and has a role in replication and segregation of the mt genome
YMR278W	PGM3	ORF, Verified	Phosphoglucomutase, catalyzes interconversion of glucose-1-phosphate and glucose-6-phospate; transcription induced in response to stress; green fluorescent protein (GFP)-fusion protein localizes to the cytoplasm and nucleus; non-essential
YMR298W	LIP1	ORF, Verified	Ceramide synthase subunit; single-span ER membrane protein associated with Lag1p and Lac1p and required for ceramide synthase activity, null mutant grows extremely slowly and is defective in ceramide synthesis
YNL029C	KTR5	ORF, Verified	Putative mannosyltransferase involved in protein glycosylation; member of the KRE2/MNT1 mannosyltransferase family
YOR175C	ALE1	ORF, Verified	Broad-specificity lysophospholipid acyltransferase, part of MBOAT family of membrane-bound O-acyltransferases; key component of Lands cycle; may have role in fatty acid exchange at sn-2 position of mature glycerophospholipids

YPL023C	MET12	ORF, Verified	Protein with methylenetetrahydrofolate reductase (MTHFR) activity in vitro; null mutant has no phenotype and is prototrophic for methionine; MET13 encodes major isozyme of MTHFR
YPL053C	KTR6	ORF, Verified	Probable mannosylphosphate transferase involved in the synthesis of core oligosaccharides in protein glycosylation pathway; member of the KRE2/MNT1 mannosyltransferase family
YPL189W	GUP2	ORF, Verified	Probable membrane protein with a possible role in proton symport of glycerol; member of the MBOAT family of putative membrane-bound O-acyltransferases; Gup1p homolog
YPR156C	TPO3	ORF, Verified	Polyamine transport protein specific for spermine; localizes to the plasma membrane; member of the major facilitator superfamily

Supplementary File 5.1.1: three_knockout1.m script

```
% FILE NAME: three-knockout1
%
% DATE CREATED: 18 January 2012
%
% PROGRAMMER: B. Heavner
% Department of Biological
% and Environmental Engineering
% Cornell University
% Ithaca, NY 14853
%
% LAST REVISED: 10 July 2012
% revisions:
% 10 July 2012 - BDH - clean up for publication
% 11 May 2012 - BDH - fix strain background
%
% REVISED BY: B. Heavner
%
% PURPOSE:
% To attempt to reproduce the model results reported by Kennedy et al. -
% that is, to apply constraints to iND750 to demonstrate the prediction of
% formic acid secretion with the mutants reported in Kennedy et al. Table
% 1.
%
% REFERENCE:
% Kennedy CJ, Boyle PM, Waks Z, Silver PA. "Systems-Level Engineering of
% Nonfermentative Metabolism in Yeast" Genetics (2009) 183:385-397. DOI:
% 10.1534/genetics.109.105254
%
% VARIABLES:
% none - it's a script. But you need iND.mat in the working directory (or
% load iND750 using readcbmodel, and comment out the "load" line of the
% script).
%
% EXPECTED OUTPUT:
% growth rates and fluxes for the first 5 mutants listed in Table 1 of
% Kennedy et al. I can't do mutants 6-8 because this version of iND750
% doesn't include reactions annotated with YAT2
%
% REQUIRED SOFTWARE:
% this script calls functions from the COBRA Toolbox and uses the gurobi
% solver
%
%-------------------------------------------------------------

%%
% first, load the iND750 model. This matlab variable was made by running
% readcbmodel to load a .sbml file downloaded from the bigg database on
% 1/13/12

load iND.mat;
model=iND;
```

```
fprintf(' \r');
fprintf('iND750 model loaded.\n');
fprintf(' \r');

%make sure we're using the gurobi solver (glpk may give different answers!)
changeCobraSolver('gurobi');

fprintf('LP solver changed to gurobi.\n');
fprintf(' \r');

%%
%next modify the model to add some strain-specific changes.

%paper wt background is fdh1 fdh2 deletion
[model,hasEffect,constrRxnNames,deletedGenes] = ...
        deleteModelGenes(model,{'YOR388C','YPL275W','YPL276W'});

%cytosolic fdh rxn (592) isn't annotated with fdh1 or fdh2. It needs to be
%constrained, too
FDH=strmatch('FDNG',model.rxns);
model.lb(FDH)=0; %make fdh1/2 mutant, remove mito FDH activity (593)
model.ub(FDH)=0; %make fdh1/2 mutant, remove mito FDH activity (593)

%%
%next, make the iND media the same as that in Kennedy et al. (based on
%their supplemental table 1).

%NOTE: the COBRA exchange reaction convention means that I need to use the
%opposite sign than Kennedy et al.
%
%NOTE 2: Table S1 lists "pydxn", "4abz", "nac", and "dhf" as having an
%upper bound of 0.5. I think these are: Pyridoxine (met 875),
%4-Aminobenzoate (met 98), Nicotinate (met 714), and 7,8-Dihydrofolate (met
%383). They don't have exchange reactions in iND750. Kennedy et al. may
%have added them, but they don't have fluxes in Table S2, so they're not
%essential.
%
%NOTE 3: I'm guessing rxn names from their met names.

%start with a clean slate: set all exchange reactions to upper bound = 1000
%and lower bound = 0 (ie, unconstrained excretion, no uptake)

exchangeRxns = findExcRxns(model);
model.lb(exchangeRxns)=0;
model.ub(exchangeRxns)=1000;

%the following exchange rxns have an upper bound and lower bound of 20 in
%Table S1 (ie, a fixed rate of uptake). I'm switching signs, so do bounds
%of -20.

ub_lb_20={'D-Glucose exchange'};

for i=1:length(ub_lb_20)
    index=find(strcmp(ub_lb_20(i),model.rxnNames));
    model.lb(index)=-20;
    model.ub(index)=-20;
```

```
end

%the following have upper and lower bounds of inf in Table S1.
%
%NOTE: oxygen not in table S1, but added to do aerobic, based on
%text of paper

ub_inf={'H2O exchange', 'Ammonia exchange', 'Phosphate exchange', ...
    'Sulfate exchange', 'Sodium exchange', 'K+ exchange', ...
    'CO2 exchange', 'O2 exchange'};

for i=1:length(ub_inf)
    index=find(strcmp(ub_inf(i),model.rxnNames));
    model.lb(index)=-1000;
end

%the following have an upper bound of 0.5 in Table S1 (but I'm switching
%signs, so do a lower bound of -0.5).

ub_05 = {'L-Asparagine exchange', 'L-Aspartate exchange', ...
    'L-Valine exchange', 'L-Tyrosine exchange', ...
    'L-Tryptophan exchange', 'L-Threonine exchange', ...
    'L-Serine exchange', 'L-Proline exchange', ...
    'L-Phenylalanine exchange', 'L-Methionine exchange', ...
    'L-Lysine exchange', 'L-Leucine exchange', ...
    'L-Isoleucine exchange', 'L-Histidine exchange', ...
    'Glycine exchange', 'L-Glutamate exchange', ...
    'L-Glutamine exchange', 'L-Cysteine exchange', ...
    'L-Arginine exchange', 'L-Alanine exchange' ...
    'Riboflavin exchange', 'Thiamin exchange', 'zymosterol exchange', ...
    'Uracil exchange', '(R)-Pantothenate exchange', ...
    'octadecynoate (n-C18:2) exchange', ...
    'octadecenoate (n-C18:1) exchange', 'myo-Inositol exchange', ...
    'hexadecenoate (n-C16:1) exchange', 'Ergosterol exchange', ...
    'Biotin exchange', 'Adenine exchange'};


for i=1:length(ub_05)
    index=find(strcmp(ub_05(i),model.rxnNames));
    model.lb(index)=-0.5; %constrained uptake
end

%the following have a lower bound of -inf in Table S1 (but I'm switching
%signs, so do an upper bound of 1000)
%
% Note: 'Deoxycytidine exchange' not included in Table S1, but has a flux
% in Table S2, so I added it here.

lb_inf = {' 1,3-beta-D-Glucan exchange', ...
    ' 4-Aminobutanoate exchange', ' 5-Amino-4-oxopentanoate exchange', ...
    ' 8-Amino-7-oxononanoate exchange', 'L-Arabinitol exchange', ...
    'Acetaldehyde exchange''Acetate exchange', 'Adenosine exchange', ...
    ' 2-Oxoglutarate exchange', 'Allantoin exchange', ...
    'S-Adenosyl-L-methionine exchange', 'D-Arabinose exchange', ...
    'L-Arabinose exchange', 'Choline exchange', 'Citrate exchange', ...
    'L-Carnitine exchange', 'Cytosine exchange', 'Cytidine exchange', ...
    'Deoxyadenosine exchange', ' 7,8-Diaminononanoate exchange', ...
```

```
    'Deoxyguanosine exchange', 'Deoxyinosine exchange', ...
    'dTTP exchange', 'Deoxyuridine exchange', 'Ethanol exchange', ...
    'FMN exchange', 'Hexadecanoate (n-C16:0) exchange', 'H+ exchange', ...
    'Hypoxanthine exchange', 'Inosine exchange', 'L-Lactate exchange', ...
    'L-Malate exchange', 'Maltose exchange', 'D-Mannose exchange', ...
    'Melibiose exchange', 'S-Methyl-L-methionine exchange', ...
    'NMN exchange', 'octadecanoate (n-C18:0) exchange', ...
    'Ornithine exchange', 'Adenosine 3'',5''-bisphosphate exchange', ...
    'peptide exchange', 'Putrescine exchange', 'Pyruvate exchange', ...
    'D-Ribose exchange', 'D-Sorbitol exchange', 'L-Sorbitol exchange', ...
    'Spermidine exchange', 'Spermine exchange', 'L-Sorbose exchange', ...
    'Succinate exchange', 'Sucrose exchange', ...
    'Thiamin monophosphate exchange', 'Thiamine diphosphate exchange', ...
    'Thymidine exchange', 'Thymine exchange', 'Trehalose exchange', ...
    'tetradecanoate (n-C14:0) exchange', 'Urea exchange', ...
    'Uridine exchange', 'Xanthine exchange', 'Xanthosine exchange', ...
    'D-Xylose exchange', 'Xylitol exchange', 'Deoxycytidine exchange'};

for i=1:length(lb_inf)
    index=find(strcmp(lb_inf(i),model.rxnNames));
    model.lb(index)=0;
    model.ub(index)=1000;
end

fprintf('Exhange reaction bounds set.\n');
fprintf(' \r');

%%
%apply extra constraints (see below) and check that FBA works on the model
model.lb([876])=0; %rxn 876 is an iND750 error. removing it affects results.
model.ub([876])=0; %rxn 876 is an iND750 error. removing it affects results.

%Kennedy S2 lists no flux for rxn 693, but it leads to very low fluxes if
%constrained. There should be a flux, but the rxn 876 model error (above)
%redirects the flux away from it. So, with the fix above, constraining rxn
%693 leads to mistakes - so it shouldn't be constrained.
%model.lb([693])=0; model.ub([693])=0;

%now, with media set, check that iND750 growth simulation still predicts
%growth.

wt_sln=optimizeCbModel(model,[],'one');
formate_ex=find(strcmp('Formate exchange',model.rxnNames));

fprintf('The WT flux is: %6.4f.\n', wt_sln.f); %2.7257
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', wt_sln.x(formate_ex)); %0
fprintf(' \r');

%%
%next step: demonstrate the results of Table 1 (strains 1-5, b/c yat2 isn't
%in iND750 - I'm not sure how strains 6-8 were generated)
%

%Kennedy reported in silico knockouts:
%1 fdh1=YOR388C fdh2=YPL275W/YPL276W alt2=YDR111C fum1=YPL262W zwf1=YNL241C
%2 fdh1 fdh2 aat2=YLR027C fum1=YPL262W zwf1=YNL241C
```

```
%3 fdh1 fdh2 cat2=YML042W fum1=YPL262W zwf1=YNL241C
%4 fdh1 fdh2 cat2=YML042W fum1=YPL262W rpe1=YJL121C
%5 fdh1 fdh2 cat2=YML042W fbp1=YLR377C fum1=YPL262W
%6 fdh1 fdh2 cat2=YML042W yat2=YER024W slc1=YDL052C
%7 fdh1 fdh2 cat2=YML042W yat2=YER024W cho1=YER026C
%8 fdh1 fdh2 cat2=YML042W yat2=YER024W alt2=YDR111C

%these genes (except YAT2) are in iND750:
%fdh1 = gene 692, rxns: 592
%fdh2 = genes 725, 726 rxns: 592
%alt2 = gene 145 rxns: 122
%fum1 = gene 718 rxns: 610 612
%zwf1 = gene 609 rxns: 627
%aat2 = gene 482 rxns: 182 184 373 375 1223 1226
%cat2 = gene 544 rxns: 266 267
%rpe1 = gene 396 rxns: 1126
%fbp1 = gene 532 rxns: 589

%make the knockouts and test them
[KO1,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YDR111C','YPL262W','YNL241C'});
KO1_sln=optimizeCbModel(KO1,[],'one');

formate_ex=find(strcmp('Formate exchange',KO1.rxnNames));
fprintf('The KO1 growth flux is: %6.4f.\n', KO1_sln.f); %1.9257
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO1_sln.x(formate_ex));
%57.7067
fprintf(' \r');


[KO2,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YLR027C', 'YPL262W', 'YNL241C'});
KO2_sln=optimizeCbModel(KO2,[],'one');

formate_ex=find(strcmp('Formate exchange',KO2.rxnNames));
fprintf('The KO2 growth flux is: %6.4f.\n', KO2_sln.f); %1.9402
fprintf('Formate exchange is reaction %d.\n', formate_ex); %
fprintf('The formate exchange flux is: %6.4f.\n', KO2_sln.x(formate_ex));
%57.5842
fprintf(' \r');


[KO3,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YPL262W', 'YNL241C'});
KO3_sln=optimizeCbModel(KO3,[],'one');

formate_ex=find(strcmp('Formate exchange',KO3.rxnNames));
fprintf('The KO3 growth flux is: %6.4f.\n', KO3_sln.f); %1.9464
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO3_sln.x(formate_ex));
%57.4775
fprintf(' \r');

[KO4,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YPL262W', 'YJL121C'});
KO4_sln=optimizeCbModel(KO4,[],'one');
```

```
formate_ex=find(strcmp('Formate exchange',KO4.rxnNames));
fprintf('The KO4 growth flux is: %6.4f.\n', KO4_sln.f); %1.9492
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO4_sln.x(formate_ex));
%57.1581
fprintf(' \r');

KO5=model;
[KO5,hasEffect,constrRxnNames,deletedGenes] = ...
        deleteModelGenes(model,{'YML042W', 'YLR377C', 'YPL262W'});
KO5_sln=optimizeCbModel(KO5,[],'one');

formate_ex=find(strcmp('Formate exchange',KO5.rxnNames));
fprintf('The KO5 growth flux is: %6.4f.\n', KO5_sln.f); %2.2378
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO5_sln.x(formate_ex));
%24.3062
fprintf(' \r');

% KO6=model;
% [KO6,hasEffect,constrRxnNames,deletedGenes] = ...
%         deleteModelGenes(model,{'YML042W', 'YER024W', 'YDL052C'});
% KO6_sln=optimizeCbModel(KO6,[],'one');
%
% fprintf('The KO6 flux is: %6.4f.\n', KO6_sln.f);
% fprintf('Formate exchange is reaction %d.\n', formate_ex);
% fprintf('The formate exchange flux is: %6.4f.\n', KO6_sln.x(formate_ex));
% fprintf(' \r');
%
% KO7=model;
% [KO7,hasEffect,constrRxnNames,deletedGenes] = ...
%         deleteModelGenes(model,{'YML042W', 'YER024W', 'YER026C'});
% KO7_sln=optimizeCbModel(KO7,[],'one');
%
% fprintf('The KO7 flux is: %6.4f.\n', KO7_sln.f);
% fprintf('Formate exchange is reaction %d.\n', formate_ex);
% fprintf('The formate exchange flux is: %6.4f.\n', KO7_sln.x(formate_ex));
% fprintf(' \r');
%
% KO8=model;
% [KO8,hasEffect,constrRxnNames,deletedGenes] = ...
%         deleteModelGenes(model,{'YML042W', 'YER024W', 'YDR111C'});
% KO8_sln=optimizeCbModel(KO8,[],'one');
%
% fprintf('The KO8 flux is: %6.4f.\n', KO8_sln.f);
% fprintf('Formate exchange is reaction %d.\n', formate_ex);
% fprintf('The formate exchange flux is: %6.4f.\n', KO8_sln.x(formate_ex));
% fprintf(' \r');

save('model','model');
```

Supplementary File 5.1.2: two-knockout2.m script

```
% FILE NAME: two-knockout2
%
% DATE CREATED: 2 February 2012
%
% PROGRAMMER: B. Heavner
% Department of Biological
% and Environmental Engineering
% Cornell University
% Ithaca, NY 14853
%
% LAST REVISED: 10 July 2012
%
% REVISIONS:
% 10 July 2012 - BDH - cleaned up for publication
% 11 May 2012 - BDH - modified candidate doubles list to add alt2 as
% possible
%
% PURPOSE:
% A variation on the model results reported by Kennedy et al. which
% compares all viable 2-knockout mutants in iND750 to find one that will
% cause formic acid secretion.
%
% In this second script, I apply a "global" search for double deletions
% that excrete formic acid to the the constrained iND750 I made in the
% three-knockout1 script (the variable "model", which has the same
% media as Kennedy et al.)
%
%
% REFERENCE:
% Kennedy CJ, Boyle PM, Waks Z, Silver PA. "Systems-Level Engineering of
% Nonfermentative Metabolism in Yeast" Genetics (2009) 183:385-397. DOI:
% 10.1534/genetics.109.105254
%
% VARIABLES:
% this script requires the model.mat variable built by three_knockout1.m in
% the working directory, and uses functions from the COBRA toolbox and the
% gurobi solver.
%
% EXPECTED OUTPUT:
% varibles listing double-gene deleted mutants that are predicted to
% excrete formate.
%
%------------------------------------------------------------

%%
%first, load the modified iND750 model with the Kennedy constraints. This
%variable generated from iND750 by the three_knockout1.m script.

load model.mat;

fprintf(' \r');
fprintf('Constrained iND750 model loaded.\n');
fprintf(' \r');

%%
```

```
%So, now that I have confidence that my iND750 constraints allow
%predictions similar to what Kennedy et al. found, I'll work on searching
%to find these mutants.

%Since iND750 has 750 ORFs, there are 70,031,500 possible triple deletions.
%To get to the ">4 million" mentioned in the Kennedy paper, we need to
%restrict the list to about 300 possible deletions. I emailed Patrick Boyle
%(one of the authors), and he sent me a mapping of 293 possible genes and
%reactions, but it's not obvious how he built it. So, I'll build my own
%list of candidate genes which can be deleted.

%%
%First, find reactions that are annotated with just 1 gene. Any of these
%reactions can be blocked with a single gene deletion, so combinations of
%three are feasible. Then, I'll remove duplicates and essential genes, to
%have a first list of candidate genes.

%Start by finding all reactions annotated with only 1 gene in
%model.rxnGeneMat

number_of_genes=sum(model.rxnGeneMat');
%number_of_genes is a row vector whose entries correspond to the number of
%genes annotating each reaction (Ex: there are 1266 reactions.
%number_of_genes(1265) is 1 b/c rxn 1265 is catalyzed by YOR011W.)

single_gene_reaction_genes=model.grRules(number_of_genes == 1);
%single_gene_reaction_genes is a list of genes which catalyze reactions,
%when the reaction is annotated as requiring only a single gene. There are
%579 such genes.

unique_single_gene_reaction_genes=unique(single_gene_reaction_genes);
%There are 379 unique genes which catalyze reactions annotated as only
%requiring a single gene (this is a nice reduction from 750 already!). How
%many of them are essential?

%next, let's remove known essential genes and dubious ORFs from our list.

%
% list of 1191 unique inviable ORFs taken from the Yeast Deletion Project (14
aug
% 11)
% http://www-
sequence.stanford.edu/group/yeast_deletion_project/downloads.html

sgd_essential_genes =
{'YAL001C';'YAL003W';'YAL025C';'YAL032C';'YAL033W';'YAL034W-a';'YAL035C-
A';'YAL038W';'YAL041W';'YAL043C';'YAR007C';'YAR008W';'YAR019C';'YBL004W';'YBL
014C';'YBL018C';'YBL020W';'YBL023C';'YBL026W';'YBL030C';'YBL034C';'YBL035C';'
YBL040C';'YBL041W';'YBL050W';'YBL073W';'YBL074C';'YBL076C';'YBL077W';'YBL084C
';'YBL092W';'YBL097W';'YBL105C';'YBR002C';'YBR004C';'YBR011C';'YBR029C';'YBR0
38W';'YBR049C';'YBR055C';'YBR060C';'YBR070C';'YBR079C';'YBR080C';'YBR087W';'Y
BR087W';'YBR088C';'YBR089W';'YBR091C';'YBR102C';'YBR109C';'YBR110W';'YBR123C'
;'YBR124W';'YBR135W';'YBR136W';'YBR140C';'YBR142W';'YBR143C';'YBR152W';'YBR15
3W';'YBR154C';'YBR155W';'YBR160W';'YBR167C';'YBR190W';'YBR192W';'YBR193C';'YB
R196C';'YBR198C';'YBR202W';'YBR211C';'YBR233W-A';'YBR233W-
A';'YBR234C';'YBR236C';'YBR237W';'YBR243C';'YBR247C';'YBR252W';'YBR253W';'YBR
254C';'YBR256C';'YBR257W';'YBR265W';'YCL003W';'YCL004W';'YCL017C';'YCL031C';'
```

YCL031C';'YCL031C';'YCL043C';'YCL052C';'YCL053C';'YCL054W';'YCL059C';'YCR012W
';'YCR012W';'YCR013C';'YCR013C';'YCR035C';'YCR052W';'YCR054C';'YCR057C';'YCR0
72C';'YCR093W';'YDL003W';'YDL004W';'YDL007W';'YDL008W';'YDL014W';'YDL015C';'Y
DL016C';'YDL017W';'YDL028C';'YDL029W';'YDL030W';'YDL031W';'YDL043C';'YDL045C'
;'YDL055C';'YDL058W';'YDL060W';'YDL064W';'YDL084W';'YDL087C';'YDL092W';'YDL09
7C';'YDL098C';'YDL102W';'YDL103C';'YDL105W';'YDL108W';'YDL111C';'YDL120W';'YD
L126C';'YDL132W';'YDL139C';'YDL140C';'YDL141W';'YDL143W';'YDL145C';'YDL147W';
'YDL148C';'YDL150W';'YDL152W';'YDL153C';'YDL163W';'YDL164C';'YDL165W';'YDL166
C';'YDL193W';'YDL195W';'YDL196W';'YDL205C';'YDL207W';'YDL208W';'YDL209C';'YDL
212W';'YDL217C';'YDL220C';'YDL221W';'YDL235C';'YDR002W';'YDR013W';'YDR016C';'
YDR021W';'YDR023W';'YDR037W';'YDR041W';'YDR044W';'YDR045C';'YDR047W';'YDR050C
';'YDR050C';'YDR052C';'YDR053W';'YDR054C';'YDR060W';'YDR062W';'YDR064W';'YDR0
81C';'YDR082W';'YDR086C';'YDR087C';'YDR088C';'YDR091C';'YDR113C';'YDR118W';'Y
DR141C';'YDR145W';'YDR160W';'YDR164C';'YDR166C';'YDR167W';'YDR168W';'YDR170C'
;'YDR172W';'YDR177W';'YDR180W';'YDR182W';'YDR187C';'YDR188W';'YDR189W';'YDR19
0C';'YDR196C';'YDR201W';'YDR208W';'YDR211W';'YDR212W';'YDR224C';'YDR224C';'YD
R224C';'YDR228C';'YDR232W';'YDR235W';'YDR236C';'YDR238C';'YDR240C';'YDR243C';
'YDR246W';'YDR267C';'YDR280W';'YDR288W';'YDR292C';'YDR299W';'YDR301W';'YDR302
W';'YDR303C';'YDR308C';'YDR311W';'YDR320C-A';'YDR320C-
A';'YDR324C';'YDR325W';'YDR327W';'YDR328C';'YDR331W';'YDR339C';'YDR341C';'YDR
353W';'YDR355C';'YDR356W';'YDR361C';'YDR362C';'YDR365C';'YDR367W';'YDR373W';'
YDR376W';'YDR381W';'YDR390C';'YDR394W';'YDR396W';'YDR397C';'YDR398W';'YDR404C
';'YDR407C';'YDR412W';'YDR413C';'YDR416W';'YDR427W';'YDR427W';'YDR429C';'YDR4
34W';'YDR437W';'YDR449C';'YDR454C';'YDR460W';'YDR464W';'YDR468C';'YDR472W';'Y
DR473C';'YDR478W';'YDR487C';'YDR489W';'YDR498C';'YDR499W';'YDR510W';'YDR526C'
;'YDR527W';'YDR531W';'YEL002C';'YEL019C';'YEL026W';'YEL032W';'YEL034W';'YEL03
5C';'YEL055C';'YEL058W';'YER003C';'YER006W';'YER008C';'YER009W';'YER012W';'YE
R013W';'YER018C';'YER021W';'YER022W';'YER023W';'YER025W';'YER029C';'YER029C';
'YER036C';'YER038C';'YER043C';'YER048W-A';'YER074W-A';'YER074W-
A';'YER082C';'YER093C';'YER094C';'YER104W';'YER112W';'YER125W';'YER126C';'YER
127W';'YER133W';'YER136W';'YER146W';'YER147C';'YER148W';'YER157W';'YER159C';'
YER165W';'YER168C';'YER171W';'YER172C';'YFL002C';'YFL005W';'YFL008W';'YFL009W
';'YFL017C';'YFL018W-A';'YFL022C';'YFL024C';'YFL029C';'YFL035C';'YFL035C-
A';'YFL037W';'YFL038C';'YFL039C';'YFL045C';'YFR002W';'YFR003C';'YFR004W';'YFR
005C';'YFR027W';'YFR028C';'YFR029W';'YFR031C';'YFR037C';'YFR042W';'YFR050C';'
YFR051C';'YFR052W';'YGL001C';'YGL008C';'YGL011C';'YGL018C';'YGL022W';'YGL030W
';'YGL040C';'YGL044C';'YGL044C';'YGL044C';'YGL047W';'YGL048C';'YGL055W';'YGL0
61C';'YGL065C';'YGL068W';'YGL069C';'YGL073W';'YGL074C';'YGL075C';'YGL091C';'Y
GL092W';'YGL093W';'YGL097W';'YGL097W';'YGL098W';'YGL099W';'YGL102C';'YGL103W'
;'YGL111W';'YGL112C';'YGL113W';'YGL116W';'YGL120C';'YGL122C';'YGL123W';'YGL12
8C';'YGL130W';'YGL137W';'YGL142C';'YGL145W';'YGL145W';'YGL145W';'YGL150C';'YG
L155W';'YGL169W';'YGL171W';'YGL172W';'YGL201C';'YGL207W';'YGL225W';'YGL233W';
'YGL238W';'YGL239C';'YGL239C';'YGL245W';'YGL247W';'YGR002C';'YGR005C';'YGR009
C';'YGR013W';'YGR024C';'YGR029W';'YGR029W';'YGR029W';'YGR030C';'YGR046W';'YGR
047C';'YGR048W';'YGR060W';'YGR065C';'YGR073C';'YGR074W';'YGR075C';'YGR082W';'
YGR083C';'YGR090W';'YGR091W';'YGR094W';'YGR095C';'YGR098C';'YGR099W';'YGR103W
';'YGR113W';'YGR114C';'YGR115C';'YGR116W';'YGR119C';'YGR120C';'YGR128C';'YGR1
40W';'YGR145W';'YGR147C';'YGR156W';'YGR158C';'YGR172C';'YGR175C';'YGR179C';'Y
GR185C';'YGR186W';'YGR190C';'YGR191W';'YGR195W';'YGR198W';'YGR211W';'YGR216C'
;'YGR218W';'YGR245C';'YGR246C';'YGR251W';'YGR253C';'YGR264C';'YGR265W';'YGR26
7C';'YGR274C';'YGR277C';'YGR278W';'YGR280C';'YHL015W';'YHR005C-
A';'YHR007C';'YHR019C';'YHR020W';'YHR023W';'YHR024C';'YHR036W';'YHR040W';'YHR
042W';'YHR058C';'YHR062C';'YHR065C';'YHR068W';'YHR069C';'YHR070W';'YHR072W';'
YHR072W-A';'YHR072W-
A';'YHR074W';'YHR083W';'YHR085W';'YHR088W';'YHR088W';'YHR088W';'YHR089C';'YHR
089C';'YHR101C';'YHR102W';'YHR102W';'YHR107C';'YHR118C';'YHR122W';'YHR128W';'
YHR128W';'YHR143W-

A';'YHR148W';'YHR164C';'YHR165C';'YHR165C';'YHR166C';'YHR169W';'YHR169W';'YHR
170W';'YHR172W';'YHR186C';'YHR188C';'YHR188C';'YHR188C';'YHR190W';'YHR196W';'
YHR197W';'YHR197W';'YHR197W';'YHR199C-A';'YHR199C-
A';'YIL003W';'YIL004C';'YIL019W';'YIL021W';'YIL022W';'YIL026C';'YIL031W';'YIL
046W';'YIL048W';'YIL051C';'YIL061C';'YIL062C';'YIL063C';'YIL068C';'YIL075C';'
YIL078W';'YIL083C';'YIL091C';'YIL104C';'YIL106W';'YIL106W';'YIL106W';'YIL109C
';'YIL115C';'YIL118W';'YIL126W';'YIL129C';'YIL142W';'YIL143C';'YIL144W';'YIL1
47C';'YIL150C';'YIL171W';'YIR006C';'YIR008C';'YIR010W';'YIR011C';'YIR012W';'Y
IR015W';'YIR022W';'YJL001W';'YJL002C';'YJL005W';'YJL008C';'YJL008C';'YJL008C'
;'YJL009W';'YJL010C';'YJL011C';'YJL014W';'YJL015C';'YJL018W';'YJL019W';'YJL02
5W';'YJL026W';'YJL031C';'YJL032W';'YJL033W';'YJL034W';'YJL035C';'YJL039C';'YJ
L041W';'YJL050W';'YJL054W';'YJL061W';'YJL069C';'YJL072C';'YJL074C';'YJL076W';
'YJL081C';'YJL085W';'YJL086C';'YJL087C';'YJL090C';'YJL091C';'YJL097W';'YJL104
W';'YJL109C';'YJL111W';'YJL125C';'YJL143W';'YJL156C';'YJL167W';'YJL173C';'YJL
174W';'YJL194W';'YJL195C';'YJL195C';'YJL202C';'YJL202C';'YJL203W';'YJR002W';'
YJR006W';'YJR007W';'YJR012C';'YJR013W';'YJR016C';'YJR017C';'YJR022W';'YJR023C
';'YJR041C';'YJR042W';'YJR045C';'YJR046W';'YJR046W';'YJR046W';'YJR057W';'YJR0
64W';'YJR065C';'YJR067C';'YJR068W';'YJR072C';'YJR076C';'YJR089W';'YJR089W';'Y
JR093C';'YJR112W';'YJR123W';'YJR141W';'YKL004W';'YKL006C-
A';'YKL012W';'YKL013C';'YKL014C';'YKL018W';'YKL019W';'YKL021C';'YKL022C';'YKL
024C';'YKL028W';'YKL033W';'YKL035W';'YKL036C';'YKL042W';'YKL045W';'YKL049C';'
YKL049C';'YKL049C';'YKL052C';'YKL058W';'YKL059C';'YKL060C';'YKL078W';'YKL082C
';'YKL083W';'YKL088W';'YKL089W';'YKL095W';'YKL099C';'YKL104C';'YKL108W';'YKL1
11C';'YKL112W';'YKL122C';'YKL125W';'YKL138C-A';'YKL138C-
A';'YKL141W';'YKL144C';'YKL145W';'YKL152C';'YKL153W';'YKL154W';'YKL165C';'YKL
172W';'YKL172W';'YKL172W';'YKL173W';'YKL180W';'YKL182W';'YKL186C';'YKL189W';'
YKL192C';'YKL193C';'YKL195W';'YKL196C';'YKL203C';'YKL210W';'YKR002W';'YKR004C
';'YKR008W';'YKR022C';'YKR025W';'YKR037C';'YKR038C';'YKR062W';'YKR063C';'YKR0
68C';'YKR071C';'YKR079C';'YKR081C';'YKR083C';'YKR086W';'YLL003W';'YLL004W';'Y
LL008W';'YLL011W';'YLL018C';'YLL031C';'YLL034C';'YLL035W';'YLL036C';'YLL037W'
;'YLL050C';'YLR002C';'YLR005W';'YLR007W';'YLR008C';'YLR009W';'YLR010C';'YLR02
2C';'YLR026C';'YLR029C';'YLR033W';'YLR045C';'YLR051C';'YLR060W';'YLR066W';'YL
R071C';'YLR075W';'YLR076C';'YLR078C';'YLR086W';'YLR088W';'YLR099W-
A';'YLR099W-
A';'YLR100W';'YLR101C';'YLR103C';'YLR105C';'YLR106C';'YLR115W';'YLR116W';'YLR
117C';'YLR127C';'YLR129W';'YLR132C';'YLR140W';'YLR141W';'YLR145W';'YLR147C';'
YLR153C';'YLR163C';'YLR166C';'YLR167W';'YLR175W';'YLR186W';'YLR186W';'YLR195C
';'YLR196W';'YLR197W';'YLR198C';'YLR208W';'YLR212C';'YLR215C';'YLR222C';'YLR2
23C';'YLR229C';'YLR230W';'YLR243W';'YLR249W';'YLR259C';'YLR272C';'YLR274W';'Y
LR275W';'YLR276C';'YLR277C';'YLR291C';'YLR293C';'YLR298C';'YLR305C';'YLR310C'
;'YLR314C';'YLR316C';'YLR316C';'YLR316C';'YLR317W';'YLR321C';'YLR323C';'YLR33
6C';'YLR339C';'YLR340W';'YLR347C';'YLR355C';'YLR359W';'YLR378C';'YLR379W';'YL
R383W';'YLR397C';'YLR409C';'YLR424W';'YLR430W';'YLR438C-
A';'YLR440C';'YLR457C';'YLR458W';'YLR459W';'YML010W';'YML015C';'YML015C';'YML
023C';'YML023C';'YML025C';'YML031W';'YML043C';'YML046W';'YML049C';'YML064C';'
YML065W';'YML069W';'YML077W';'YML085C';'YML091C';'YML092C';'YML092C';'YML092C
';'YML093W';'YML098W';'YML105C';'YML114C';'YML125C';'YML126C';'YML127W';'YML1
30C';'YMR001C';'YMR005W';'YMR005W';'YMR013C';'YMR028W';'YMR033W';'YMR033W';'Y
MR043W';'YMR047C';'YMR047C';'YMR049C';'YMR059W';'YMR059W';'YMR059W';'YMR061W'
;'YMR076C';'YMR079W';'YMR093W';'YMR094W';'YMR108W';'YMR108W';'YMR112C';'YMR11
3W';'YMR117C';'YMR128W';'YMR131C';'YMR134W';'YMR146C';'YMR149W';'YMR168C';'YM
R197C';'YMR200W';'YMR203W';'YMR208W';'YMR211W';'YMR213W';'YMR218C';'YMR220W';
'YMR227C';'YMR229C';'YMR235C';'YMR236W';'YMR239C';'YMR240C';'YMR260C';'YMR268
C';'YMR270C';'YMR277W';'YMR281W';'YMR288W';'YMR290C';'YMR290W-
A';'YMR296C';'YMR298W';'YMR301C';'YMR308C';'YMR309C';'YMR314W';'YNL002C';'YNL
006W';'YNL007C';'YNL024C-A';'YNL024C-
A';'YNL026W';'YNL036W';'YNL036W';'YNL038W';'YNL039W';'YNL061W';'YNL062C';'YNL

```
075W';'YNL088W';'YNL102W';'YNL103W';'YNL110C';'YNL112W';'YNL112W';'YNL112W';'
YNL113W';'YNL114C';'YNL118C';'YNL124W';'YNL126W';'YNL126W';'YNL131W';'YNL132W
';'YNL137C';'YNL138W-A';'YNL138W-
A';'YNL149C';'YNL150W';'YNL151C';'YNL152W';'YNL158W';'YNL161W';'YNL163C';'YNL
172W';'YNL178W';'YNL181W';'YNL182C';'YNL188W';'YNL189W';'YNL207W';'YNL216W';'
YNL221C';'YNL222W';'YNL232W';'YNL240C';'YNL244C';'YNL245C';'YNL247W';'YNL251C
';'YNL256W';'YNL258C';'YNL260C';'YNL261W';'YNL262W';'YNL263C';'YNL267W';'YNL2
72C';'YNL282W';'YNL287W';'YNL290W';'YNL306W';'YNL308C';'YNL310C';'YNL312W';'Y
NL313C';'YNL317W';'YNR003C';'YNR011C';'YNR016C';'YNR017W';'YNR026C';'YNR035C'
;'YNR038W';'YNR043W';'YNR046W';'YNR053C';'YNR054C';'YOL005C';'YOL010W';'YOL02
1C';'YOL022C';'YOL026C';'YOL034W';'YOL038W';'YOL040C';'YOL066C';'YOL069W';'YO
L077C';'YOL078W';'YOL094C';'YOL097C';'YOL102C';'YOL120C';'YOL123W';'YOL127W';
'YOL130W';'YOL133W';'YOL134C';'YOL135C';'YOL139C';'YOL142W';'YOL142W';'YOL144
W';'YOL146W';'YOL149W';'YOR004W';'YOR020C';'YOR046C';'YOR048C';'YOR056C';'YOR
057W';'YOR060C';'YOR063W';'YOR074C';'YOR075W';'YOR077W';'YOR095C';'YOR098C';'
YOR102W';'YOR103C';'YOR110W';'YOR116C';'YOR117W';'YOR119C';'YOR122C';'YOR143C
';'YOR145C';'YOR146W';'YOR148C';'YOR149C';'YOR151C';'YOR157C';'YOR159C';'YOR1
60W';'YOR168W';'YOR169C';'YOR174W';'YOR176W';'YOR181W';'YOR194C';'YOR203W';'Y
OR204W';'YOR206W';'YOR207C';'YOR210W';'YOR217W';'YOR218C';'YOR224C';'YOR232W'
;'YOR236W';'YOR244W';'YOR249C';'YOR250C';'YOR254C';'YOR256C';'YOR257W';'YOR25
9C';'YOR260W';'YOR261C';'YOR262W';'YOR272W';'YOR278W';'YOR281C';'YOR282W';'YO
R287C';'YOR294W';'YOR310C';'YOR319W';'YOR326W';'YOR329C';'YOR335C';'YOR336W';
'YOR340C';'YOR341W';'YOR353C';'YOR361C';'YOR362C';'YOR370C';'YOR372C';'YOR373
W';'YPL007C';'YPL010W';'YPL011C';'YPL012W';'YPL016W';'YPL020C';'YPL028W';'YPL
043W';'YPL044C';'YPL063W';'YPL076W';'YPL082C';'YPL083C';'YPL085W';'YPL093W';'
YPL094C';'YPL117C';'YPL122C';'YPL124W';'YPL126W';'YPL128C';'YPL131W';'YPL142C
';'YPL143W';'YPL146C';'YPL151C';'YPL153C';'YPL160W';'YPL169C';'YPL175W';'YPL1
90C';'YPL204W';'YPL209C';'YPL210C';'YPL211W';'YPL217C';'YPL218W';'YPL228W';'Y
PL231W';'YPL233W';'YPL235W';'YPL237W';'YPL238C';'YPL242C';'YPL243W';'YPL251W'
;'YPL252C';'YPL255W';'YPL266W';'YPR010C';'YPR016C';'YPR019W';'YPR025C';'YPR03
3C';'YPR034W';'YPR035W';'YPR041W';'YPR048W';'YPR055W';'YPR056W';'YPR082C';'YP
R085C';'YPR086W';'YPR088C';'YPR094W';'YPR103W';'YPR104C';'YPR105C';'YPR107C';
'YPR108W';'YPR110C';'YPR112C';'YPR113W';'YPR133C';'YPR136C';'YPR137W';'YPR142
C';'YPR143W';'YPR144C';'YPR161C';'YPR162C';'YPR165W';'YPR168W';'YPR169W';'YPR
175W';'YPR176C';'YPR177C';'YPR178W';'YPR180W';'YPR181C';'YPR182W';'YPR183W';'
YPR186C';'YPR187W';'YPR190C';};

% list of 4941 verified ORFs taken from SGD (14 aug 11)
% http://www.yeastgenome.org/cgi-
bin/search/featureSearch?featuretype=ORF&qualifier=Verified

verified_genes =
{'Q0045';'Q0050';'Q0055';'Q0060';'Q0065';'Q0070';'Q0080';'Q0085';'Q0105';'Q01
10';'Q0115';'Q0120';'Q0130';'Q0140';'Q0160';'Q0250';'Q0275';'R0010W';'R0020C'
;'R0030W';'R0040C';'YAL001C';'YAL002W';'YAL003W';'YAL005C';'YAL007C';'YAL008W
';'YAL009W';'YAL010C';'YAL011W';'YAL012W';'YAL013W';'YAL014C';'YAL015C';'YAL0
16W';'YAL017W';'YAL019W';'YAL020C';'YAL021C';'YAL022C';'YAL023C';'YAL024C';'Y
AL025C';'YAL026C';'YAL027W';'YAL028W';'YAL029C';'YAL030W';'YAL031C';'YAL032C'
;'YAL033W';'YAL034C';'YAL034W-
A';'YAL035W';'YAL036C';'YAL038W';'YAL039C';'YAL040C';'YAL041W';'YAL042W';'YAL
043C';'YAL044C';'YAL046C';'YAL047C';'YAL048C';'YAL049C';'YAL051W';'YAL053W';'
YAL054C';'YAL055W';'YAL056W';'YAL058W';'YAL059W';'YAL060W';'YAL062W';'YAL063C
';'YAL064W';'YAL067C';'YAL068C';'YAR002C-
A';'YAR002W';'YAR003W';'YAR007C';'YAR008W';'YAR014C';'YAR015W';'YAR018C';'YAR
019C';'YAR020C';'YAR027W';'YAR031W';'YAR033W';'YAR035W';'YAR042W';'YAR050W';'
YAR071W';'YBL001C';'YBL002W';'YBL003C';'YBL004W';'YBL005W';'YBL006C';'YBL007C
';'YBL008W';'YBL009W';'YBL011W';'YBL013W';'YBL014C';'YBL015W';'YBL016W';'YBL0
```

17C';'YBL018C';'YBL019W';'YBL020W';'YBL021C';'YBL022C';'YBL023C';'YBL024W';'Y
BL025W';'YBL026W';'YBL027W';'YBL028C';'YBL030C';'YBL031W';'YBL032W';'YBL033C'
;'YBL034C';'YBL035C';'YBL036C';'YBL037W';'YBL038W';'YBL039C';'YBL040C';'YBL04
1W';'YBL042C';'YBL043W';'YBL045C';'YBL046W';'YBL047C';'YBL049W';'YBL050W';'YB
L051C';'YBL052C';'YBL054W';'YBL055C';'YBL056W';'YBL057C';'YBL058W';'YBL059C-
A';'YBL060W';'YBL061C';'YBL063W';'YBL064C';'YBL066C';'YBL067C';'YBL068W';'YBL
069W';'YBL071W-
A';'YBL072C';'YBL074C';'YBL075C';'YBL076C';'YBL078C';'YBL079W';'YBL080C';'YBL
082C';'YBL084C';'YBL085W';'YBL087C';'YBL088C';'YBL089W';'YBL090W';'YBL091C';'
YBL091C-
A';'YBL092W';'YBL093C';'YBL097W';'YBL098W';'YBL099W';'YBL101C';'YBL102W';'YBL
103C';'YBL104C';'YBL105C';'YBL106C';'YBL108C-
A';'YBR001C';'YBR002C';'YBR003W';'YBR004C';'YBR005W';'YBR006W';'YBR008C';'YBR
009C';'YBR010W';'YBR011C';'YBR014C';'YBR015C';'YBR016W';'YBR017C';'YBR018C';'
YBR019C';'YBR020W';'YBR021W';'YBR022W';'YBR023C';'YBR024W';'YBR025C';'YBR026C
';'YBR028C';'YBR029C';'YBR030W';'YBR031W';'YBR034C';'YBR035C';'YBR036C';'YBR0
37C';'YBR038W';'YBR039W';'YBR040W';'YBR041W';'YBR042C';'YBR043C';'YBR044C';'Y
BR045C';'YBR046C';'YBR048W';'YBR049C';'YBR050C';'YBR052C';'YBR054W';'YBR055C'
;'YBR057C';'YBR058C';'YBR058C-
A';'YBR059C';'YBR060C';'YBR061C';'YBR065C';'YBR066C';'YBR067C';'YBR068C';'YBR
069C';'YBR070C';'YBR071W';'YBR072W';'YBR073W';'YBR076W';'YBR077C';'YBR078W';'
YBR079C';'YBR080C';'YBR081C';'YBR082C';'YBR083W';'YBR084C-
A';'YBR084W';'YBR085W';'YBR086C';'YBR087W';'YBR088C';'YBR089C-
A';'YBR091C';'YBR092C';'YBR093C';'YBR094W';'YBR095C';'YBR097W';'YBR098W';'YBR
101C';'YBR102C';'YBR103W';'YBR104W';'YBR105C';'YBR106W';'YBR107C';'YBR108W';'
YBR109C';'YBR110W';'YBR111C';'YBR111W-
A';'YBR112C';'YBR114W';'YBR115C';'YBR117C';'YBR118W';'YBR119W';'YBR120C';'YBR
121C';'YBR122C';'YBR123C';'YBR125C';'YBR126C';'YBR127C';'YBR128C';'YBR129C';'
YBR130C';'YBR131W';'YBR132C';'YBR133C';'YBR135W';'YBR136W';'YBR137W';'YBR139W
';'YBR140C';'YBR142W';'YBR143C';'YBR145W';'YBR146W';'YBR147W';'YBR148W';'YBR1
49W';'YBR150C';'YBR151W';'YBR152W';'YBR153W';'YBR154C';'YBR155W';'YBR156C';'Y
BR157C';'YBR158W';'YBR159W';'YBR160W';'YBR161W';'YBR162C';'YBR162W-
A';'YBR163W';'YBR164C';'YBR165W';'YBR166C';'YBR167C';'YBR168W';'YBR169C';'YBR
170C';'YBR171W';'YBR172C';'YBR173C';'YBR175W';'YBR176W';'YBR177C';'YBR179C';'
YBR180W';'YBR181C';'YBR182C';'YBR183W';'YBR185C';'YBR186W';'YBR188C';'YBR189W
';'YBR191W';'YBR192W';'YBR193C';'YBR194W';'YBR195C';'YBR196C';'YBR198C';'YBR1
99W';'YBR200W';'YBR201W';'YBR202W';'YBR203W';'YBR204C';'YBR205W';'YBR207W';'Y
BR208C';'YBR210W';'YBR211C';'YBR212W';'YBR213W';'YBR214W';'YBR215W';'YBR216C'
;'YBR217W';'YBR218C';'YBR221C';'YBR222C';'YBR223C';'YBR227C';'YBR228W';'YBR22
9C';'YBR230C';'YBR231C';'YBR233W';'YBR233W-
A';'YBR234C';'YBR236C';'YBR237W';'YBR238C';'YBR240C';'YBR243C';'YBR244W';'YBR
245C';'YBR247C';'YBR248C';'YBR249C';'YBR250W';'YBR251W';'YBR252W';'YBR253W';'
YBR254C';'YBR255W';'YBR256C';'YBR257W';'YBR258C';'YBR260C';'YBR261C';'YBR262C
';'YBR263W';'YBR264C';'YBR265W';'YBR267W';'YBR268W';'YBR272C';'YBR273C';'YBR2
74W';'YBR275C';'YBR276C';'YBR278W';'YBR279W';'YBR280C';'YBR281C';'YBR282W';'Y
BR283C';'YBR286W';'YBR288C';'YBR289W';'YBR290W';'YBR291C';'YBR293W';'YBR294W'
;'YBR295W';'YBR296C';'YBR297W';'YBR298C';'YBR299W';'YBR301W';'YBR302C';'YCL00
1W';'YCL004W';'YCL005W';'YCL005W-
A';'YCL008C';'YCL009C';'YCL010C';'YCL011C';'YCL012C';'YCL014W';'YCL016C';'YCL
017C';'YCL018W';'YCL024W';'YCL025C';'YCL026C-
A';'YCL027W';'YCL028W';'YCL029C';'YCL030C';'YCL031C';'YCL032W';'YCL033C';'YCL
034W';'YCL035C';'YCL036W';'YCL037C';'YCL038C';'YCL039W';'YCL040W';'YCL043C';'
YCL044C';'YCL045C';'YCL047C';'YCL048W';'YCL050C';'YCL051W';'YCL052C';'YCL054W
';'YCL055W';'YCL056C';'YCL057W';'YCL058C';'YCL058W-
A';'YCL059C';'YCL061C';'YCL063W';'YCL064C';'YCL066W';'YCL067C';'YCL069W';'YCL
073C';'YCR002C';'YCR003W';'YCR004C';'YCR005C';'YCR008W';'YCR009C';'YCR010C';'
YCR011C';'YCR012W';'YCR014C';'YCR017C';'YCR018C';'YCR019W';'YCR020C';'YCR020C

-A';'YCR020W-B';'YCR021C';'YCR023C';'YCR024C';'YCR024C-
A';'YCR026C';'YCR027C';'YCR028C';'YCR028C-
A';'YCR030C';'YCR031C';'YCR032W';'YCR033W';'YCR034W';'YCR035C';'YCR036W';'YCR
037C';'YCR038C';'YCR039C';'YCR040W';'YCR042C';'YCR044C';'YCR045C';'YCR046C';'
YCR047C';'YCR048W';'YCR052W';'YCR053W';'YCR054C';'YCR057C';'YCR059C';'YCR060W
';'YCR063W';'YCR065W';'YCR066W';'YCR067C';'YCR068W';'YCR069W';'YCR071C';'YCR0
72C';'YCR073C';'YCR073W-
A';'YCR075C';'YCR077C';'YCR079W';'YCR081W';'YCR082W';'YCR083W';'YCR084C';'YCR
086W';'YCR088W';'YCR089W';'YCR091W';'YCR092C';'YCR093W';'YCR094W';'YCR096C';'
YCR097W';'YCR098C';'YCR104W';'YCR105W';'YCR106W';'YCR107W';'YDL001W';'YDL002C
';'YDL003W';'YDL004W';'YDL005C';'YDL006W';'YDL007W';'YDL008W';'YDL010W';'YDL0
12C';'YDL013W';'YDL014W';'YDL015C';'YDL017W';'YDL018C';'YDL019C';'YDL020C';'Y
DL021W';'YDL022W';'YDL024C';'YDL025C';'YDL028C';'YDL029W';'YDL030W';'YDL031W'
;'YDL033C';'YDL035C';'YDL036C';'YDL037C';'YDL039C';'YDL040C';'YDL042C';'YDL04
3C';'YDL044C';'YDL045C';'YDL045W-
A';'YDL046W';'YDL047W';'YDL048C';'YDL049C';'YDL051W';'YDL052C';'YDL053C';'YDL
054C';'YDL055C';'YDL056W';'YDL058W';'YDL059C';'YDL060W';'YDL061C';'YDL063C';'
YDL064W';'YDL065C';'YDL066W';'YDL067C';'YDL069C';'YDL070W';'YDL072C';'YDL074C
';'YDL075W';'YDL076C';'YDL077C';'YDL078C';'YDL079C';'YDL080C';'YDL081C';'YDL0
82W';'YDL083C';'YDL084W';'YDL085W';'YDL087C';'YDL088C';'YDL089W';'YDL090C';'Y
DL091C';'YDL092W';'YDL093W';'YDL095W';'YDL097C';'YDL098C';'YDL099W';'YDL100C'
;'YDL101C';'YDL102W';'YDL103C';'YDL104C';'YDL105W';'YDL106C';'YDL107W';'YDL10
8W';'YDL110C';'YDL111C';'YDL112W';'YDL113C';'YDL115C';'YDL116W';'YDL117W';'YD
L120W';'YDL122W';'YDL123W';'YDL124W';'YDL125C';'YDL126C';'YDL127W';'YDL128W';
'YDL130W';'YDL130W-A';'YDL131W';'YDL132W';'YDL133C-
A';'YDL133W';'YDL134C';'YDL135C';'YDL136W';'YDL137W';'YDL138W';'YDL139C';'YDL
140C';'YDL141W';'YDL142C';'YDL143W';'YDL145C';'YDL146W';'YDL147W';'YDL148C';'
YDL149W';'YDL150W';'YDL153C';'YDL154W';'YDL155W';'YDL159W';'YDL160C';'YDL161W
';'YDL164C';'YDL165W';'YDL166C';'YDL167C';'YDL168W';'YDL169C';'YDL170W';'YDL1
71C';'YDL173W';'YDL174C';'YDL175C';'YDL176W';'YDL178W';'YDL179W';'YDL181W';'Y
DL182W';'YDL183C';'YDL184C';'YDL185W';'YDL188C';'YDL189W';'YDL190C';'YDL191W'
;'YDL192W';'YDL193W';'YDL194W';'YDL195W';'YDL197C';'YDL198C';'YDL200C';'YDL20
1W';'YDL202W';'YDL203C';'YDL204W';'YDL205C';'YDL207W';'YDL208W';'YDL209C';'YD
L210W';'YDL212W';'YDL213C';'YDL214C';'YDL215C';'YDL216C';'YDL217C';'YDL219W';
'YDL220C';'YDL222C';'YDL223C';'YDL224C';'YDL225W';'YDL226C';'YDL227C';'YDL229
W';'YDL230W';'YDL231C';'YDL232W';'YDL234C';'YDL235C';'YDL236W';'YDL237W';'YDL
238C';'YDL239C';'YDL240W';'YDL243C';'YDL244W';'YDL245C';'YDL247W';'YDL248W';'
YDR001C';'YDR002W';'YDR003W';'YDR004W';'YDR005C';'YDR006C';'YDR007W';'YDR009W
';'YDR011W';'YDR012W';'YDR013W';'YDR014W';'YDR014W-
A';'YDR016C';'YDR017C';'YDR019C';'YDR021W';'YDR022C';'YDR023W';'YDR025W';'YDR
026C';'YDR027C';'YDR028C';'YDR030C';'YDR031W';'YDR032C';'YDR033W';'YDR034C';'
YDR035W';'YDR036C';'YDR037W';'YDR038C';'YDR039C';'YDR040C';'YDR041W';'YDR043C
';'YDR044W';'YDR045C';'YDR046C';'YDR047W';'YDR049W';'YDR050C';'YDR051C';'YDR0
52C';'YDR054C';'YDR055W';'YDR057W';'YDR058C';'YDR059C';'YDR060W';'YDR062W';'Y
DR063W';'YDR064W';'YDR065W';'YDR068W';'YDR069C';'YDR071C';'YDR072C';'YDR073W'
;'YDR074W';'YDR075W';'YDR076W';'YDR077W';'YDR078C';'YDR079C-
A';'YDR079W';'YDR080W';'YDR081C';'YDR082W';'YDR083W';'YDR084C';'YDR085C';'YDR
086C';'YDR087C';'YDR088C';'YDR091C';'YDR092W';'YDR093W';'YDR096W';'YDR097C';'
YDR098C';'YDR099W';'YDR100W';'YDR101C';'YDR103W';'YDR104C';'YDR105C';'YDR106W
';'YDR107C';'YDR108W';'YDR110W';'YDR113C';'YDR116C';'YDR117C';'YDR118W';'YDR1
20C';'YDR121W';'YDR122W';'YDR123C';'YDR125C';'YDR126W';'YDR127W';'YDR128W';'Y
DR129C';'YDR130C';'YDR135C';'YDR137W';'YDR138W';'YDR139C';'YDR140W';'YDR141C'
;'YDR142C';'YDR143C';'YDR144C';'YDR145W';'YDR146C';'YDR147W';'YDR148C';'YDR15
0W';'YDR151C';'YDR152W';'YDR153C';'YDR155C';'YDR156W';'YDR158W';'YDR159W';'YD
R160W';'YDR162C';'YDR163W';'YDR164C';'YDR165W';'YDR166C';'YDR167W';'YDR168W';
'YDR169C';'YDR170C';'YDR171W';'YDR172W';'YDR173C';'YDR174W';'YDR175C';'YDR176
W';'YDR177W';'YDR178W';'YDR179C';'YDR180W';'YDR181C';'YDR182W';'YDR183W';'YDR

184C';'YDR185C';'YDR186C';'YDR188W';'YDR189W';'YDR190C';'YDR191W';'YDR192C';'
YDR194C';'YDR195W';'YDR196C';'YDR197W';'YDR198C';'YDR200C';'YDR201W';'YDR202C
';'YDR204W';'YDR205W';'YDR206W';'YDR207C';'YDR208W';'YDR211W';'YDR212W';'YDR2
13W';'YDR214W';'YDR216W';'YDR217C';'YDR218C';'YDR219C';'YDR221W';'YDR223W';'Y
DR224C';'YDR225W';'YDR226W';'YDR227W';'YDR228C';'YDR229W';'YDR231C';'YDR232W'
;'YDR233C';'YDR234W';'YDR235W';'YDR236C';'YDR237W';'YDR238C';'YDR239C';'YDR24
0C';'YDR242W';'YDR243C';'YDR244W';'YDR245W';'YDR246W';'YDR247W';'YDR251W';'YD
R252W';'YDR253C';'YDR254W';'YDR255C';'YDR256C';'YDR257C';'YDR258C';'YDR259C';
'YDR260C';'YDR261C';'YDR263C';'YDR264C';'YDR265W';'YDR266C';'YDR267C';'YDR268
W';'YDR270W';'YDR272W';'YDR273W';'YDR275W';'YDR276C';'YDR277C';'YDR279W';'YDR
280W';'YDR281C';'YDR283C';'YDR284C';'YDR285W';'YDR287W';'YDR288W';'YDR289C';'
YDR292C';'YDR293C';'YDR294C';'YDR295C';'YDR296W';'YDR297W';'YDR298C';'YDR299W
';'YDR300C';'YDR301W';'YDR302W';'YDR303C';'YDR304C';'YDR305C';'YDR308C';'YDR3
09C';'YDR310C';'YDR311W';'YDR312W';'YDR313C';'YDR314C';'YDR315C';'YDR316W';'Y
DR317W';'YDR318W';'YDR320C';'YDR320C-A';'YDR321W';'YDR322C-
A';'YDR322W';'YDR323C';'YDR324C';'YDR325W';'YDR326C';'YDR328C';'YDR329C';'YDR
330W';'YDR331W';'YDR332W';'YDR334W';'YDR335W';'YDR337W';'YDR339C';'YDR341C';'
YDR342C';'YDR343C';'YDR345C';'YDR346C';'YDR347W';'YDR349C';'YDR350C';'YDR351W
';'YDR353W';'YDR354W';'YDR356W';'YDR358W';'YDR359C';'YDR361C';'YDR362C';'YDR3
63W';'YDR363W-
A';'YDR364C';'YDR365C';'YDR367W';'YDR368W';'YDR369C';'YDR372C';'YDR373W';'YDR
375C';'YDR376W';'YDR377W';'YDR378C';'YDR379C-A';'YDR379W';'YDR380W';'YDR381C-
A';'YDR381W';'YDR382W';'YDR383C';'YDR384C';'YDR385W';'YDR386W';'YDR388W';'YDR
389W';'YDR390C';'YDR392W';'YDR393W';'YDR394W';'YDR395W';'YDR397C';'YDR398W';'
YDR399W';'YDR400W';'YDR402C';'YDR403W';'YDR404C';'YDR405W';'YDR406W';'YDR407C
';'YDR408C';'YDR409W';'YDR410C';'YDR411C';'YDR412W';'YDR414C';'YDR416W';'YDR4
18W';'YDR419W';'YDR420W';'YDR421W';'YDR422C';'YDR423C';'YDR424C';'YDR425W';'Y
DR427W';'YDR428C';'YDR429C';'YDR430C';'YDR432W';'YDR434W';'YDR435C';'YDR436W'
;'YDR437W';'YDR438W';'YDR439W';'YDR440W';'YDR441C';'YDR443C';'YDR446W';'YDR44
7C';'YDR448W';'YDR449C';'YDR450W';'YDR451C';'YDR452W';'YDR453C';'YDR454C';'YD
R456W';'YDR457W';'YDR458C';'YDR459C';'YDR460W';'YDR461W';'YDR462W';'YDR463W';
'YDR464W';'YDR465C';'YDR466W';'YDR468C';'YDR469W';'YDR470C';'YDR471W';'YDR472
W';'YDR473C';'YDR475C';'YDR477W';'YDR478W';'YDR479C';'YDR480W';'YDR481C';'YDR
482C';'YDR483W';'YDR484W';'YDR485C';'YDR486C';'YDR487C';'YDR488C';'YDR489W';'
YDR490C';'YDR492W';'YDR493W';'YDR494W';'YDR495C';'YDR496C';'YDR497C';'YDR498C
';'YDR499W';'YDR500C';'YDR501W';'YDR502C';'YDR503C';'YDR504C';'YDR505C';'YDR5
07C';'YDR508C';'YDR510W';'YDR511W';'YDR512C';'YDR513W';'YDR514C';'YDR515W';'Y
DR516C';'YDR517W';'YDR518W';'YDR519W';'YDR522C';'YDR523C';'YDR524C';'YDR525W-
A';'YDR527W';'YDR528W';'YDR529C';'YDR530C';'YDR531W';'YDR532C';'YDR533C';'YDR
534C';'YDR536W';'YDR538W';'YDR539W';'YDR540C';'YDR542W';'YDR545W';'YEL001C';'
YEL002C';'YEL003W';'YEL004W';'YEL005C';'YEL006W';'YEL009C';'YEL011W';'YEL012W
';'YEL013W';'YEL015W';'YEL016C';'YEL017C-
A';'YEL017W';'YEL018W';'YEL019C';'YEL020W-
A';'YEL021W';'YEL022W';'YEL024W';'YEL026W';'YEL027W';'YEL029C';'YEL030W';'YEL
031W';'YEL032W';'YEL034W';'YEL036C';'YEL037C';'YEL038W';'YEL039C';'YEL040W';'
YEL041W';'YEL042W';'YEL043W';'YEL044W';'YEL046C';'YEL047C';'YEL048C';'YEL049W
';'YEL050C';'YEL051W';'YEL052W';'YEL053C';'YEL054C';'YEL055C';'YEL056W';'YEL0
58W';'YEL059C-
A';'YEL060C';'YEL061C';'YEL062W';'YEL063C';'YEL064C';'YEL065W';'YEL066W';'YEL
069C';'YEL071W';'YEL072W';'YER001W';'YER002W';'YER003C';'YER004W';'YER005W';'
YER006W';'YER007C-
A';'YER007W';'YER008C';'YER009W';'YER010C';'YER011W';'YER012W';'YER013W';'YER
014C-A';'YER014W';'YER015W';'YER016W';'YER017C';'YER018C';'YER019C-
A';'YER019W';'YER020W';'YER021W';'YER022W';'YER023W';'YER024W';'YER025W';'YER
026C';'YER027C';'YER028C';'YER029C';'YER030W';'YER031C';'YER032W';'YER033C';'
YER035W';'YER036C';'YER037W';'YER038C';'YER039C';'YER040W';'YER041W';'YER042W
';'YER043C';'YER044C';'YER044C-

A';'YER045C';'YER046W';'YER047C';'YER048C';'YER048W-
A';'YER049W';'YER050C';'YER051W';'YER052C';'YER053C';'YER054C';'YER055C';'YER
056C';'YER056C-A';'YER057C';'YER058W';'YER059W';'YER060W';'YER060W-
A';'YER061C';'YER062C';'YER063W';'YER065C';'YER067W';'YER068W';'YER069W';'YER
070W';'YER072W';'YER073W';'YER074W';'YER074W-
A';'YER075C';'YER078C';'YER080W';'YER081W';'YER082C';'YER083C';'YER086W';'YER
087C-
B';'YER087W';'YER088C';'YER089C';'YER090W';'YER091C';'YER092W';'YER093C';'YER
093C-
A';'YER094C';'YER095W';'YER096W';'YER098W';'YER099C';'YER100W';'YER101C';'YER
102W';'YER103W';'YER104W';'YER105C';'YER106W';'YER107C';'YER109C';'YER110C';'
YER111C';'YER112W';'YER113C';'YER114C';'YER115C';'YER116C';'YER117W';'YER118C
';'YER119C';'YER120W';'YER122C';'YER123W';'YER124C';'YER125W';'YER126C';'YER1
27W';'YER128W';'YER129W';'YER131W';'YER132C';'YER133W';'YER134C';'YER136W';'Y
ER139C';'YER141W';'YER142C';'YER143W';'YER144C';'YER145C';'YER146W';'YER147C'
;'YER148W';'YER149C';'YER150W';'YER151C';'YER152C';'YER153C';'YER154W';'YER15
5C';'YER157W';'YER159C';'YER161C';'YER162C';'YER164W';'YER165W';'YER166W';'YE
R167W';'YER168C';'YER169W';'YER170W';'YER171W';'YER172C';'YER173W';'YER174C';
'YER175C';'YER176W';'YER177W';'YER178W';'YER179W';'YER180C';'YER180C-
A';'YER183C';'YER185W';'YER190W';'YFL001W';'YFL002C';'YFL003C';'YFL004W';'YFL
005W';'YFL007W';'YFL008W';'YFL009W';'YFL010C';'YFL010W-
A';'YFL011W';'YFL013C';'YFL014W';'YFL016C';'YFL017C';'YFL017W-
A';'YFL018C';'YFL020C';'YFL021W';'YFL022C';'YFL023W';'YFL024C';'YFL025C';'YFL
026W';'YFL027C';'YFL028C';'YFL029C';'YFL030W';'YFL031W';'YFL033C';'YFL034C-
A';'YFL034C-
B';'YFL036W';'YFL037W';'YFL038C';'YFL039C';'YFL041W';'YFL044C';'YFL045C';'YFL
047W';'YFL048C';'YFL049W';'YFL050C';'YFL053W';'YFL055W';'YFL056C';'YFL057C';'
YFL058W';'YFL059W';'YFL060C';'YFL062W';'YFR001W';'YFR002W';'YFR003C';'YFR004W
';'YFR005C';'YFR007W';'YFR008W';'YFR009W';'YFR010W';'YFR011C';'YFR013W';'YFR0
14C';'YFR015C';'YFR016C';'YFR017C';'YFR019W';'YFR021W';'YFR022W';'YFR023W';'Y
FR024C-
A';'YFR025C';'YFR026C';'YFR027W';'YFR028C';'YFR029W';'YFR030W';'YFR031C';'YFR
031C-A';'YFR032C-
A';'YFR033C';'YFR034C';'YFR036W';'YFR037C';'YFR038W';'YFR040W';'YFR041C';'YFR
042W';'YFR043C';'YFR044C';'YFR046C';'YFR047C';'YFR048W';'YFR049W';'YFR050C';'
YFR051C';'YFR052W';'YFR053C';'YGL001C';'YGL002W';'YGL003C';'YGL004C';'YGL005C
';'YGL006W';'YGL008C';'YGL009C';'YGL011C';'YGL012W';'YGL013C';'YGL014W';'YGL0
16W';'YGL017W';'YGL018C';'YGL019W';'YGL020C';'YGL021W';'YGL022W';'YGL023C';'Y
GL025C';'YGL026C';'YGL027C';'YGL028C';'YGL029W';'YGL030W';'YGL031C';'YGL032C'
;'YGL033W';'YGL035C';'YGL037C';'YGL038C';'YGL039W';'YGL040C';'YGL043W';'YGL04
4C';'YGL045W';'YGL047W';'YGL048C';'YGL049C';'YGL050W';'YGL051W';'YGL053W';'YG
L054C';'YGL055W';'YGL056C';'YGL057C';'YGL058W';'YGL059W';'YGL060W';'YGL061C';
'YGL062W';'YGL063W';'YGL064C';'YGL065C';'YGL066W';'YGL067W';'YGL068W';'YGL070
C';'YGL071W';'YGL073W';'YGL075C';'YGL076C';'YGL077C';'YGL078C';'YGL083W';'YGL
084C';'YGL086W';'YGL087C';'YGL089C';'YGL090W';'YGL091C';'YGL092W';'YGL093W';'
YGL094C';'YGL095C';'YGL096W';'YGL097W';'YGL098W';'YGL099W';'YGL100W';'YGL103W
';'YGL104C';'YGL105W';'YGL106W';'YGL107C';'YGL110C';'YGL111W';'YGL112C';'YGL1
13W';'YGL115W';'YGL116W';'YGL119W';'YGL120C';'YGL121C';'YGL122C';'YGL123W';'Y
GL124C';'YGL125W';'YGL126W';'YGL127C';'YGL128C';'YGL129C';'YGL130W';'YGL131C'
;'YGL133W';'YGL134W';'YGL135W';'YGL136C';'YGL137W';'YGL139W';'YGL141W';'YGL14
2C';'YGL143C';'YGL144C';'YGL145W';'YGL147C';'YGL148W';'YGL150C';'YGL151W';'YG
L153W';'YGL154C';'YGL155W';'YGL156W';'YGL157W';'YGL158W';'YGL160W';'YGL161C';
'YGL162W';'YGL163C';'YGL164C';'YGL166W';'YGL167C';'YGL168W';'YGL169W';'YGL170
C';'YGL171W';'YGL172W';'YGL173C';'YGL174W';'YGL175C';'YGL178W';'YGL179C';'YGL
180W';'YGL181W';'YGL183C';'YGL184C';'YGL186C';'YGL187C';'YGL189C';'YGL190C';'
YGL191W';'YGL192W';'YGL194C';'YGL195W';'YGL196W';'YGL197W';'YGL198W';'YGL200C
';'YGL201C';'YGL202W';'YGL203C';'YGL205W';'YGL206C';'YGL207W';'YGL208W';'YGL2

```
09W';'YGL210W';'YGL211W';'YGL212W';'YGL213C';'YGL215W';'YGL216W';'YGL219C';'Y
GL220W';'YGL221C';'YGL222C';'YGL223C';'YGL224C';'YGL225W';'YGL226C-
A';'YGL226W';'YGL227W';'YGL228W';'YGL229C';'YGL231C';'YGL232W';'YGL233W';'YGL
234W';'YGL236C';'YGL237C';'YGL238W';'YGL240W';'YGL241W';'YGL243W';'YGL244W';'
YGL245W';'YGL246C';'YGL247W';'YGL248W';'YGL249W';'YGL250W';'YGL251C';'YGL252C
';'YGL253W';'YGL254W';'YGL255W';'YGL256W';'YGL257C';'YGL258W';'YGL263W';'YGR0
02C';'YGR003W';'YGR004W';'YGR005C';'YGR006W';'YGR007W';'YGR008C';'YGR009C';'Y
GR010W';'YGR012W';'YGR013W';'YGR014W';'YGR019W';'YGR020C';'YGR023W';'YGR024C'
;'YGR027C';'YGR028W';'YGR029W';'YGR030C';'YGR031C-
A';'YGR031W';'YGR032W';'YGR033C';'YGR034W';'YGR036C';'YGR037C';'YGR038W';'YGR
040W';'YGR041W';'YGR043C';'YGR044C';'YGR046W';'YGR047C';'YGR048W';'YGR049W';'
YGR054W';'YGR055W';'YGR056W';'YGR057C';'YGR058W';'YGR059W';'YGR060W';'YGR061C
';'YGR062C';'YGR063C';'YGR065C';'YGR068C';'YGR070W';'YGR072W';'YGR074W';'YGR0
75C';'YGR076C';'YGR077C';'YGR078C';'YGR080W';'YGR081C';'YGR082W';'YGR083C';'Y
GR084C';'YGR085C';'YGR086C';'YGR087C';'YGR088W';'YGR089W';'YGR090W';'YGR091W'
;'YGR092W';'YGR094W';'YGR095C';'YGR096W';'YGR097W';'YGR098C';'YGR099W';'YGR10
0W';'YGR101W';'YGR102C';'YGR103W';'YGR104C';'YGR105W';'YGR106C';'YGR108W';'YG
R109C';'YGR110W';'YGR112W';'YGR113W';'YGR116W';'YGR118W';'YGR119C';'YGR120C';
'YGR121C';'YGR122W';'YGR123C';'YGR124W';'YGR128C';'YGR129W';'YGR130C';'YGR131
W';'YGR132C';'YGR133W';'YGR134W';'YGR135W';'YGR136W';'YGR138C';'YGR140W';'YGR
141W';'YGR142W';'YGR143W';'YGR144W';'YGR145W';'YGR146C';'YGR147C';'YGR148C';'
YGR150C';'YGR152C';'YGR154C';'YGR155W';'YGR156W';'YGR157W';'YGR158C';'YGR159C
';'YGR162W';'YGR163W';'YGR165W';'YGR166W';'YGR167W';'YGR169C';'YGR170W';'YGR1
71C';'YGR172C';'YGR173W';'YGR174C';'YGR175C';'YGR177C';'YGR178C';'YGR179C';'Y
GR180C';'YGR181W';'YGR183C';'YGR184C';'YGR185C';'YGR186W';'YGR187C';'YGR188C'
;'YGR189C';'YGR191W';'YGR192C';'YGR193C';'YGR194C';'YGR195W';'YGR196C';'YGR19
7C';'YGR198W';'YGR199W';'YGR200C';'YGR202C';'YGR203W';'YGR204W';'YGR205W';'YG
R206W';'YGR207C';'YGR208W';'YGR209C';'YGR211W';'YGR212W';'YGR213C';'YGR214W';
'YGR215W';'YGR216C';'YGR217W';'YGR218W';'YGR220C';'YGR221C';'YGR222W';'YGR223
C';'YGR224W';'YGR225W';'YGR227W';'YGR229C';'YGR230W';'YGR231C';'YGR232W';'YGR
233C';'YGR234W';'YGR236C';'YGR238C';'YGR239C';'YGR240C';'YGR241C';'YGR244C';'
YGR245C';'YGR246C';'YGR247W';'YGR248W';'YGR249W';'YGR250C';'YGR251W';'YGR252W
';'YGR253C';'YGR254W';'YGR255C';'YGR256W';'YGR257C';'YGR258C';'YGR260W';'YGR2
61C';'YGR262C';'YGR263C';'YGR264C';'YGR266W';'YGR267C';'YGR268C';'YGR270W';'Y
GR271C-
A';'YGR271W';'YGR274C';'YGR275W';'YGR276C';'YGR277C';'YGR278W';'YGR279C';'YGR
280C';'YGR281W';'YGR282C';'YGR283C';'YGR284C';'YGR285C';'YGR286C';'YGR287C';'
YGR288W';'YGR289C';'YGR292W';'YGR294W';'YGR295C';'YGR296W';'YHL001W';'YHL002W
';'YHL003C';'YHL004W';'YHL006C';'YHL007C';'YHL009C';'YHL010C';'YHL011C';'YHL0
13C';'YHL014C';'YHL015W';'YHL016C';'YHL019C';'YHL020C';'YHL021C';'YHL022C';'Y
HL023C';'YHL024W';'YHL025W';'YHL027W';'YHL028W';'YHL030W';'YHL031C';'YHL032C'
;'YHL033C';'YHL034C';'YHL035C';'YHL036W';'YHL038C';'YHL039W';'YHL040C';'YHL04
3W';'YHL046C';'YHL047C';'YHL048W';'YHR001W';'YHR001W-
A';'YHR002W';'YHR003C';'YHR004C';'YHR005C';'YHR005C-
A';'YHR006W';'YHR007C';'YHR008C';'YHR010W';'YHR011W';'YHR012W';'YHR013C';'YHR
014W';'YHR015W';'YHR016C';'YHR017W';'YHR018C';'YHR019C';'YHR020W';'YHR021C';'
YHR023W';'YHR024C';'YHR025W';'YHR026W';'YHR027C';'YHR028C';'YHR029C';'YHR030C
';'YHR031C';'YHR032W';'YHR034C';'YHR036W';'YHR037W';'YHR038W';'YHR039C';'YHR0
39C-
A';'YHR040W';'YHR041C';'YHR042W';'YHR043C';'YHR044C';'YHR046C';'YHR047C';'YHR
049W';'YHR050W';'YHR051W';'YHR052W';'YHR053C';'YHR055C';'YHR056C';'YHR057C';'
YHR058C';'YHR059W';'YHR060W';'YHR061C';'YHR062C';'YHR063C';'YHR064C';'YHR065C
';'YHR066W';'YHR067W';'YHR068W';'YHR069C';'YHR070W';'YHR071W';'YHR072W';'YHR0
72W-A';'YHR073W';'YHR074W';'YHR075C';'YHR076W';'YHR077C';'YHR079C';'YHR079C-
A';'YHR080C';'YHR081W';'YHR082C';'YHR083W';'YHR084W';'YHR085W';'YHR086W';'YHR
087W';'YHR088W';'YHR089C';'YHR090C';'YHR091C';'YHR092C';'YHR094C';'YHR096C';'
YHR098C';'YHR099W';'YHR100C';'YHR101C';'YHR102W';'YHR103W';'YHR104W';'YHR105W
```

';'YHR106W';'YHR107C';'YHR108W';'YHR109W';'YHR110W';'YHR111W';'YHR112C';'YHR1
13W';'YHR114W';'YHR115C';'YHR116W';'YHR117W';'YHR118C';'YHR119W';'YHR120W';'Y
HR121W';'YHR122W';'YHR123W';'YHR124W';'YHR127W';'YHR128W';'YHR129C';'YHR132C'
;'YHR132W-
A';'YHR133C';'YHR134W';'YHR135C';'YHR136C';'YHR137W';'YHR139C';'YHR141C';'YHR
142W';'YHR143W';'YHR143W-
A';'YHR144C';'YHR146W';'YHR147C';'YHR148W';'YHR149C';'YHR150W';'YHR151C';'YHR
152W';'YHR153C';'YHR154W';'YHR155W';'YHR156C';'YHR157W';'YHR158C';'YHR160C';'
YHR161C';'YHR163W';'YHR164C';'YHR165C';'YHR166C';'YHR167W';'YHR168W';'YHR169W
';'YHR170W';'YHR171W';'YHR172W';'YHR174W';'YHR175W';'YHR176W';'YHR178W';'YHR1
79W';'YHR181W';'YHR183W';'YHR184W';'YHR185C';'YHR186C';'YHR187W';'YHR188C';'Y
HR189W';'YHR190W';'YHR191C';'YHR193C';'YHR194W';'YHR195W';'YHR196W';'YHR197W'
;'YHR198C';'YHR199C';'YHR199C-
A';'YHR200W';'YHR201C';'YHR203C';'YHR204W';'YHR205W';'YHR206W';'YHR207C';'YHR
208W';'YHR209W';'YHR211W';'YHR215W';'YHR216W';'YIL002C';'YIL003W';'YIL004C';'
YIL005W';'YIL006W';'YIL007C';'YIL008W';'YIL009C-
A';'YIL009W';'YIL010W';'YIL011W';'YIL013C';'YIL014W';'YIL015W';'YIL016W';'YIL
017C';'YIL018W';'YIL019W';'YIL020C';'YIL021W';'YIL022W';'YIL023C';'YIL026C';'
YIL027C';'YIL030C';'YIL031W';'YIL033C';'YIL034C';'YIL035C';'YIL036W';'YIL037C
';'YIL038C';'YIL039W';'YIL040W';'YIL041W';'YIL042C';'YIL043C';'YIL044C';'YIL0
45W';'YIL046W';'YIL047C';'YIL048W';'YIL049W';'YIL050W';'YIL051C';'YIL052C';'Y
IL053W';'YIL056W';'YIL057C';'YIL061C';'YIL062C';'YIL063C';'YIL064W';'YIL065C'
;'YIL066C';'YIL068C';'YIL069C';'YIL070C';'YIL071C';'YIL072W';'YIL073C';'YIL07
4C';'YIL075C';'YIL076W';'YIL078W';'YIL079C';'YIL083C';'YIL084C';'YIL085C';'YI
L087C';'YIL088C';'YIL089W';'YIL090W';'YIL091C';'YIL093C';'YIL094C';'YIL095W';
'YIL097W';'YIL098C';'YIL099W';'YIL101C';'YIL103W';'YIL104C';'YIL105C';'YIL106
W';'YIL107C';'YIL109C';'YIL110W';'YIL111W';'YIL112W';'YIL113W';'YIL114C';'YIL
115C';'YIL116W';'YIL117C';'YIL118W';'YIL119C';'YIL120W';'YIL121W';'YIL122W';'
YIL123W';'YIL124W';'YIL125W';'YIL126W';'YIL128W';'YIL129C';'YIL130W';'YIL131C
';'YIL132C';'YIL133C';'YIL134W';'YIL135C';'YIL136W';'YIL137C';'YIL138C';'YIL1
39C';'YIL140W';'YIL142W';'YIL143C';'YIL144W';'YIL145C';'YIL146C';'YIL147C';'Y
IL148W';'YIL149C';'YIL150C';'YIL153W';'YIL154C';'YIL155C';'YIL156W';'YIL157C'
;'YIL158W';'YIL159W';'YIL160C';'YIL162W';'YIL164C';'YIL172C';'YIL173W';'YIR00
1C';'YIR002C';'YIR003W';'YIR004W';'YIR005W';'YIR006C';'YIR008C';'YIR009W';'YI
R010W';'YIR011C';'YIR012W';'YIR013C';'YIR015W';'YIR017C';'YIR018W';'YIR019C';
'YIR021W';'YIR022W';'YIR023W';'YIR024C';'YIR025W';'YIR026C';'YIR027C';'YIR028
W';'YIR029W';'YIR030C';'YIR031C';'YIR032C';'YIR033W';'YIR034C';'YIR037W';'YIR
038C';'YIR039C';'YIR041W';'YJL001W';'YJL002C';'YJL003W';'YJL004C';'YJL005W';'
YJL006C';'YJL008C';'YJL010C';'YJL011C';'YJL012C';'YJL013C';'YJL014W';'YJL019W
';'YJL020C';'YJL023C';'YJL024C';'YJL025W';'YJL026W';'YJL028W';'YJL029C';'YJL0
30W';'YJL031C';'YJL033W';'YJL034W';'YJL035C';'YJL036W';'YJL037W';'YJL038C';'Y
JL039C';'YJL041W';'YJL042W';'YJL044C';'YJL045W';'YJL046W';'YJL047C';'YJL048C'
;'YJL050W';'YJL051W';'YJL052W';'YJL053W';'YJL054W';'YJL056C';'YJL057C';'YJL05
8C';'YJL059W';'YJL060W';'YJL061W';'YJL062W';'YJL062W-
A';'YJL063C';'YJL065C';'YJL066C';'YJL068C';'YJL069C';'YJL071W';'YJL072C';'YJL
073W';'YJL074C';'YJL076W';'YJL077C';'YJL078C';'YJL079C';'YJL080C';'YJL081C';'
YJL082W';'YJL083W';'YJL084C';'YJL085W';'YJL087C';'YJL088W';'YJL089W';'YJL090C
';'YJL091C';'YJL092W';'YJL093C';'YJL094C';'YJL095W';'YJL096W';'YJL097W';'YJL0
98W';'YJL099W';'YJL100W';'YJL101C';'YJL102W';'YJL103C';'YJL104W';'YJL105W';'Y
JL106W';'YJL108C';'YJL109C';'YJL110C';'YJL111W';'YJL112W';'YJL115W';'YJL116C'
;'YJL117W';'YJL118W';'YJL121C';'YJL122W';'YJL123C';'YJL124C';'YJL125C';'YJL12
6W';'YJL127C';'YJL128C';'YJL129C';'YJL130C';'YJL131C';'YJL133W';'YJL134W';'YJ
L136C';'YJL137C';'YJL138C';'YJL139C';'YJL140W';'YJL141C';'YJL143W';'YJL144W';
'YJL145W';'YJL146W';'YJL148W';'YJL149W';'YJL151C';'YJL153C';'YJL154C';'YJL155
C';'YJL156C';'YJL157C';'YJL158C';'YJL159W';'YJL162C';'YJL164C';'YJL165C';'YJL
166W';'YJL167W';'YJL168C';'YJL170C';'YJL171C';'YJL172W';'YJL173C';'YJL174W';'
YJL176C';'YJL177W';'YJL178C';'YJL179W';'YJL180C';'YJL183W';'YJL184W';'YJL186W

';'YJL187C';'YJL189W';'YJL190C';'YJL191W';'YJL192C';'YJL194W';'YJL196C';'YJL1
97W';'YJL198W';'YJL200C';'YJL201W';'YJL203W';'YJL204C';'YJL205C';'YJL207C';'Y
JL208C';'YJL209W';'YJL210W';'YJL212C';'YJL213W';'YJL214W';'YJL216C';'YJL217W'
;'YJL219W';'YJL221C';'YJL222W';'YJL223C';'YJR001W';'YJR002W';'YJR004C';'YJR00
5W';'YJR006W';'YJR007W';'YJR009C';'YJR010C-
A';'YJR010W';'YJR013W';'YJR014W';'YJR016C';'YJR017C';'YJR019C';'YJR021C';'YJR
022W';'YJR024C';'YJR025C';'YJR031C';'YJR032W';'YJR033C';'YJR034W';'YJR035W';'
YJR036C';'YJR040W';'YJR041C';'YJR042W';'YJR043C';'YJR044C';'YJR045C';'YJR046W
';'YJR047C';'YJR048W';'YJR049C';'YJR050W';'YJR051W';'YJR052W';'YJR053W';'YJR0
55W';'YJR057W';'YJR058C';'YJR059W';'YJR060W';'YJR062C';'YJR063W';'YJR064W';'Y
JR065C';'YJR066W';'YJR067C';'YJR068W';'YJR069C';'YJR070C';'YJR072C';'YJR073C'
;'YJR074W';'YJR075W';'YJR076C';'YJR077C';'YJR078W';'YJR080C';'YJR082C';'YJR08
3C';'YJR084W';'YJR086W';'YJR088C';'YJR089W';'YJR090C';'YJR091C';'YJR092W';'YJ
R093C';'YJR094C';'YJR094W-
A';'YJR095W';'YJR096W';'YJR097W';'YJR099W';'YJR100C';'YJR101W';'YJR102C';'YJR
103W';'YJR104C';'YJR105W';'YJR106W';'YJR108W';'YJR109C';'YJR110W';'YJR112W';'
YJR113C';'YJR117W';'YJR118C';'YJR119C';'YJR120W';'YJR121W';'YJR122W';'YJR123W
';'YJR125C';'YJR126C';'YJR127C';'YJR130C';'YJR131W';'YJR132W';'YJR133W';'YJR1
34C';'YJR135C';'YJR135W-
A';'YJR136C';'YJR137C';'YJR138W';'YJR139C';'YJR140C';'YJR143C';'YJR144W';'YJR
145C';'YJR147W';'YJR148W';'YJR150C';'YJR151C';'YJR152W';'YJR153W';'YJR155W';'
YJR156C';'YJR158W';'YJR159W';'YJR160C';'YJR161C';'YKL001C';'YKL002W';'YKL003C
';'YKL004W';'YKL005C';'YKL006C-
A';'YKL006W';'YKL007W';'YKL008C';'YKL009W';'YKL010C';'YKL011C';'YKL012W';'YKL
013C';'YKL014C';'YKL015W';'YKL016C';'YKL017C';'YKL018W';'YKL019W';'YKL020C';'
YKL021C';'YKL022C';'YKL024C';'YKL025C';'YKL026C';'YKL027W';'YKL028W';'YKL029C
';'YKL032C';'YKL033W';'YKL034W';'YKL035W';'YKL037W';'YKL038W';'YKL039W';'YKL0
40C';'YKL041W';'YKL042W';'YKL043W';'YKL045W';'YKL046C';'YKL048C';'YKL049C';'Y
KL050C';'YKL051W';'YKL052C';'YKL053C-
A';'YKL054C';'YKL055C';'YKL056C';'YKL057C';'YKL058W';'YKL059C';'YKL060C';'YKL
062W';'YKL064W';'YKL065C';'YKL067W';'YKL068W';'YKL069W';'YKL072W';'YKL073W';'
YKL074C';'YKL078W';'YKL079W';'YKL080W';'YKL081W';'YKL082C';'YKL084W';'YKL085W
';'YKL086W';'YKL087C';'YKL088W';'YKL089W';'YKL090W';'YKL091C';'YKL092C';'YKL0
93W';'YKL094W';'YKL095W';'YKL096W';'YKL096W-
A';'YKL098W';'YKL099C';'YKL101W';'YKL103C';'YKL104C';'YKL106W';'YKL108W';'YKL
109W';'YKL110C';'YKL112W';'YKL113C';'YKL114C';'YKL116C';'YKL117W';'YKL119C';'
YKL120W';'YKL122C';'YKL124W';'YKL125W';'YKL126W';'YKL127W';'YKL128C';'YKL129C
';'YKL130C';'YKL132C';'YKL134C';'YKL135C';'YKL137W';'YKL138C';'YKL138C-
A';'YKL139W';'YKL140W';'YKL141W';'YKL142W';'YKL143W';'YKL144C';'YKL145W';'YKL
146W';'YKL148C';'YKL149C';'YKL150W';'YKL152C';'YKL154W';'YKL155C';'YKL156W';'
YKL157W';'YKL159C';'YKL160W';'YKL161C';'YKL163W';'YKL164C';'YKL165C';'YKL166C
';'YKL167C';'YKL168C';'YKL170W';'YKL171W';'YKL172W';'YKL173W';'YKL174C';'YKL1
75W';'YKL176C';'YKL178C';'YKL179C';'YKL180W';'YKL181W';'YKL182W';'YKL183W';'Y
KL184W';'YKL185W';'YKL186C';'YKL188C';'YKL189W';'YKL190W';'YKL191W';'YKL192C'
;'YKL193C';'YKL194C';'YKL195W';'YKL196C';'YKL197C';'YKL198C';'YKL201C';'YKL20
3C';'YKL204W';'YKL205W';'YKL206C';'YKL207W';'YKL208W';'YKL209C';'YKL210W';'YK
L211C';'YKL212W';'YKL213C';'YKL214C';'YKL215C';'YKL216W';'YKL217W';'YKL218C';
'YKL219W';'YKL220C';'YKL221W';'YKL222C';'YKL224C';'YKR001C';'YKR002W';'YKR003
W';'YKR004C';'YKR006C';'YKR007W';'YKR008W';'YKR009C';'YKR010C';'YKR013W';'YKR
014C';'YKR016W';'YKR019C';'YKR020W';'YKR021W';'YKR022C';'YKR024C';'YKR025W';'
YKR026C';'YKR027W';'YKR028W';'YKR029C';'YKR030W';'YKR031C';'YKR034W';'YKR035W
-
A';'YKR036C';'YKR037C';'YKR038C';'YKR039W';'YKR041W';'YKR042W';'YKR043C';'YKR
044W';'YKR046C';'YKR048C';'YKR049C';'YKR050W';'YKR052C';'YKR053C';'YKR054C';'
YKR055W';'YKR056W';'YKR057W';'YKR058W';'YKR059W';'YKR060W';'YKR061W';'YKR062W
';'YKR063C';'YKR064W';'YKR065C';'YKR066C';'YKR067W';'YKR068C';'YKR069W';'YKR0
71C';'YKR072C';'YKR074W';'YKR076W';'YKR077W';'YKR078W';'YKR079C';'YKR080W';'Y

KR081C';'YKR082W';'YKR083C';'YKR084C';'YKR085C';'YKR086W';'YKR087C';'YKR088C'
;'YKR089C';'YKR090W';'YKR091W';'YKR092C';'YKR093W';'YKR094C';'YKR095W';'YKR09
5W-
A';'YKR096W';'YKR097W';'YKR098C';'YKR099W';'YKR100C';'YKR101W';'YKR102W';'YKR
103W';'YKR104W';'YKR106W';'YLL001W';'YLL002W';'YLL003W';'YLL004W';'YLL005C';'
YLL006W';'YLL008W';'YLL009C';'YLL010C';'YLL011W';'YLL012W';'YLL013C';'YLL014W
';'YLL015W';'YLL018C';'YLL018C-
A';'YLL019C';'YLL021W';'YLL022C';'YLL023C';'YLL024C';'YLL025W';'YLL026W';'YLL
027W';'YLL028W';'YLL029W';'YLL031C';'YLL032C';'YLL033W';'YLL034C';'YLL035W';'
YLL036C';'YLL038C';'YLL039C';'YLL040C';'YLL041C';'YLL042C';'YLL043W';'YLL045C
';'YLL046C';'YLL048C';'YLL049W';'YLL050C';'YLL051C';'YLL052C';'YLL055W';'YLL0
57C';'YLL060C';'YLL061W';'YLL062C';'YLL063C';'YLL064C';'YLR002C';'YLR003C';'Y
LR004C';'YLR005W';'YLR006C';'YLR007W';'YLR008C';'YLR009W';'YLR010C';'YLR011W'
;'YLR013W';'YLR014C';'YLR015W';'YLR016C';'YLR017W';'YLR018C';'YLR019W';'YLR02
0C';'YLR021W';'YLR022C';'YLR023C';'YLR024C';'YLR025W';'YLR026C';'YLR027C';'YL
R028C';'YLR029C';'YLR032W';'YLR033W';'YLR034C';'YLR035C';'YLR037C';'YLR038C';
'YLR039C';'YLR043C';'YLR044C';'YLR045C';'YLR047C';'YLR048W';'YLR051C';'YLR052
W';'YLR054C';'YLR055C';'YLR056W';'YLR058C';'YLR059C';'YLR060W';'YLR061W';'YLR
064W';'YLR066W';'YLR067C';'YLR068W';'YLR069C';'YLR070C';'YLR071C';'YLR073C';'
YLR074C';'YLR075W';'YLR077W';'YLR078C';'YLR079W';'YLR080W';'YLR081W';'YLR082C
';'YLR083C';'YLR084C';'YLR085C';'YLR086W';'YLR087C';'YLR088W';'YLR089C';'YLR0
90W';'YLR091W';'YLR092W';'YLR093C';'YLR094C';'YLR095C';'YLR096W';'YLR097C';'Y
LR098C';'YLR099C';'YLR100W';'YLR102C';'YLR103C';'YLR105C';'YLR106C';'YLR107W'
;'YLR109W';'YLR110C';'YLR113W';'YLR114C';'YLR115W';'YLR116W';'YLR117C';'YLR11
8C';'YLR119W';'YLR120C';'YLR121C';'YLR127C';'YLR128W';'YLR129W';'YLR130C';'YL
R131C';'YLR133W';'YLR134W';'YLR135W';'YLR136C';'YLR137W';'YLR138W';'YLR139C';
'YLR141W';'YLR142W';'YLR144C';'YLR145W';'YLR146C';'YLR147C';'YLR148W';'YLR150
W';'YLR151C';'YLR153C';'YLR154C';'YLR154W-
C';'YLR155C';'YLR157C';'YLR158C';'YLR160C';'YLR162W';'YLR163C';'YLR164W';'YLR
165C';'YLR166C';'YLR167W';'YLR168C';'YLR170C';'YLR172C';'YLR174W';'YLR175W';'
YLR176C';'YLR178C';'YLR179C';'YLR180W';'YLR181C';'YLR182W';'YLR183C';'YLR185W
';'YLR186W';'YLR188W';'YLR189C';'YLR190W';'YLR191W';'YLR192C';'YLR193C';'YLR1
94C';'YLR195C';'YLR196W';'YLR197W';'YLR199C';'YLR200W';'YLR201C';'YLR203C';'Y
LR204W';'YLR205C';'YLR206W';'YLR207W';'YLR208W';'YLR209C';'YLR210W';'YLR212C'
;'YLR213C';'YLR214W';'YLR215C';'YLR216C';'YLR218C';'YLR219W';'YLR220W';'YLR22
1C';'YLR222C';'YLR223C';'YLR226W';'YLR227C';'YLR228C';'YLR229C';'YLR231C';'YL
R233C';'YLR234W';'YLR237W';'YLR238W';'YLR239C';'YLR240W';'YLR242C';'YLR244C';
'YLR245C';'YLR246W';'YLR247C';'YLR248W';'YLR249W';'YLR250W';'YLR251W';'YLR254
C';'YLR256W';'YLR258W';'YLR259C';'YLR260W';'YLR262C';'YLR262C-
A';'YLR263W';'YLR264W';'YLR265C';'YLR266C';'YLR268W';'YLR270W';'YLR272C';'YLR
273C';'YLR274W';'YLR275W';'YLR276C';'YLR277C';'YLR284C';'YLR285W';'YLR286C';'
YLR287C-
A';'YLR288C';'YLR289W';'YLR291C';'YLR292C';'YLR293C';'YLR295C';'YLR298C';'YLR
299W';'YLR300W';'YLR301W';'YLR303W';'YLR304C';'YLR305C';'YLR306W';'YLR307W';'
YLR308W';'YLR309C';'YLR310C';'YLR312W-
A';'YLR313C';'YLR314C';'YLR315W';'YLR316C';'YLR318W';'YLR319C';'YLR320W';'YLR
321C';'YLR323C';'YLR324W';'YLR325C';'YLR327C';'YLR328W';'YLR329W';'YLR330W';'
YLR332W';'YLR333C';'YLR335W';'YLR336C';'YLR337C';'YLR340W';'YLR341W';'YLR342W
';'YLR343W';'YLR344W';'YLR347C';'YLR348C';'YLR350W';'YLR351C';'YLR353W';'YLR3
54C';'YLR355C';'YLR356W';'YLR357W';'YLR359W';'YLR360W';'YLR361C';'YLR362W';'Y
LR363C';'YLR364W';'YLR367W';'YLR368W';'YLR369W';'YLR370C';'YLR371W';'YLR372W'
;'YLR373C';'YLR375W';'YLR376C';'YLR377C';'YLR378C';'YLR380W';'YLR381W';'YLR38
2C';'YLR383W';'YLR384C';'YLR385C';'YLR386W';'YLR387C';'YLR388W';'YLR389C';'YL
R390W';'YLR390W-
A';'YLR392C';'YLR393W';'YLR394W';'YLR395C';'YLR396C';'YLR397C';'YLR398C';'YLR
399C';'YLR401C';'YLR403W';'YLR404W';'YLR405W';'YLR406C';'YLR409C';'YLR410W';'
YLR411W';'YLR412W';'YLR414C';'YLR417W';'YLR418C';'YLR420W';'YLR421C';'YLR423C

';'YLR424W';'YLR425W';'YLR427W';'YLR429W';'YLR430W';'YLR431C';'YLR432W';'YLR4
33C';'YLR435W';'YLR436C';'YLR437C';'YLR438C-
A';'YLR438W';'YLR439W';'YLR440C';'YLR441C';'YLR442C';'YLR443W';'YLR447C';'YLR
448W';'YLR449W';'YLR450W';'YLR451W';'YLR452C';'YLR453C';'YLR457C';'YLR459W';'
YLR461W';'YLR466W';'YLR467W';'YML001W';'YML004C';'YML005W';'YML006C';'YML007W
';'YML008C';'YML009C';'YML010W';'YML011C';'YML012W';'YML013W';'YML014W';'YML0
15C';'YML016C';'YML017W';'YML019W';'YML021C';'YML022W';'YML023C';'YML024W';'Y
ML025C';'YML026C';'YML027W';'YML028W';'YML029W';'YML030W';'YML031W';'YML032C'
;'YML034W';'YML035C';'YML036W';'YML038C';'YML041C';'YML042W';'YML043C';'YML04
6W';'YML047C';'YML048W';'YML049C';'YML050W';'YML051W';'YML052W';'YML054C';'YM
L055W';'YML056C';'YML057W';'YML058W';'YML058W-
A';'YML059C';'YML060W';'YML061C';'YML062C';'YML063W';'YML064C';'YML065W';'YML
066C';'YML067C';'YML068W';'YML069W';'YML070W';'YML071C';'YML072C';'YML073C';'
YML074C';'YML075C';'YML076C';'YML077W';'YML078W';'YML080W';'YML081C-
A';'YML081W';'YML085C';'YML086C';'YML087C';'YML088W';'YML091C';'YML092C';'YML
093W';'YML094W';'YML095C';'YML097C';'YML098W';'YML099C';'YML100W';'YML101C';'
YML102W';'YML103C';'YML104C';'YML105C';'YML106W';'YML107C';'YML109W';'YML110C
';'YML111W';'YML112W';'YML113W';'YML114C';'YML115C';'YML116W';'YML117W';'YML1
18W';'YML120C';'YML121W';'YML123C';'YML124C';'YML125C';'YML126C';'YML127W';'Y
ML128C';'YML129C';'YML130C';'YML132W';'YMR001C';'YMR002W';'YMR003W';'YMR004W'
;'YMR005W';'YMR006C';'YMR008C';'YMR009W';'YMR011W';'YMR012W';'YMR013C';'YMR01
4W';'YMR015C';'YMR016C';'YMR017W';'YMR019W';'YMR020W';'YMR021C';'YMR022W';'YM
R023C';'YMR024W';'YMR025W';'YMR026C';'YMR028W';'YMR029C';'YMR030W';'YMR031C';
'YMR032W';'YMR033W';'YMR035W';'YMR036C';'YMR037C';'YMR038C';'YMR039C';'YMR040
W';'YMR041C';'YMR042W';'YMR043W';'YMR044W';'YMR047C';'YMR048W';'YMR049C';'YMR
052W';'YMR053C';'YMR054W';'YMR055C';'YMR056C';'YMR058W';'YMR059W';'YMR060C';'
YMR061W';'YMR062C';'YMR063W';'YMR064W';'YMR065W';'YMR066W';'YMR067C';'YMR068W
';'YMR069W';'YMR070W';'YMR071C';'YMR072W';'YMR073C';'YMR074C';'YMR075W';'YMR0
76C';'YMR077C';'YMR078C';'YMR079W';'YMR080C';'YMR081C';'YMR083W';'YMR086W';'Y
MR087W';'YMR088C';'YMR089C';'YMR091C';'YMR092C';'YMR093W';'YMR094W';'YMR095C'
;'YMR096W';'YMR097C';'YMR098C';'YMR099C';'YMR100W';'YMR101C';'YMR104C';'YMR10
5C';'YMR106C';'YMR107W';'YMR108W';'YMR109W';'YMR110C';'YMR112C';'YMR113W';'YM
R114C';'YMR115W';'YMR116C';'YMR117C';'YMR119W';'YMR120C';'YMR121C';'YMR123W';
'YMR125W';'YMR127C';'YMR128W';'YMR129W';'YMR131C';'YMR133W';'YMR135C';'YMR136
W';'YMR137C';'YMR138W';'YMR139W';'YMR140W';'YMR142C';'YMR143W';'YMR145C';'YMR
146C';'YMR148W';'YMR149W';'YMR150C';'YMR152W';'YMR153W';'YMR154C';'YMR156C';'
YMR157C';'YMR158W';'YMR159C';'YMR161W';'YMR162C';'YMR163C';'YMR164C';'YMR165C
';'YMR167W';'YMR168C';'YMR169C';'YMR170C';'YMR171C';'YMR172W';'YMR173W';'YMR1
74C';'YMR175W';'YMR176W';'YMR177W';'YMR179W';'YMR180C';'YMR182C';'YMR183C';'Y
MR184W';'YMR186W';'YMR188C';'YMR189W';'YMR190C';'YMR191W';'YMR192W';'YMR193W'
;'YMR194C-
B';'YMR194W';'YMR195W';'YMR197C';'YMR198W';'YMR199W';'YMR200W';'YMR201C';'YMR
202W';'YMR203W';'YMR204C';'YMR205C';'YMR207C';'YMR208W';'YMR210W';'YMR211W';'
YMR212C';'YMR213W';'YMR214W';'YMR215W';'YMR216C';'YMR217W';'YMR218C';'YMR219W
';'YMR220W';'YMR222C';'YMR223W';'YMR224C';'YMR225C';'YMR226C';'YMR227C';'YMR2
28W';'YMR229C';'YMR230W';'YMR231W';'YMR232W';'YMR233W';'YMR234W';'YMR235C';'Y
MR236W';'YMR237W';'YMR238W';'YMR239C';'YMR240C';'YMR241W';'YMR242C';'YMR243C'
;'YMR246W';'YMR247C';'YMR250W';'YMR251W';'YMR251W-
A';'YMR255W';'YMR256C';'YMR257C';'YMR258C';'YMR260C';'YMR261C';'YMR263W';'YMR
264W';'YMR266W';'YMR267W';'YMR268C';'YMR269W';'YMR270C';'YMR271C';'YMR272C';'
YMR273C';'YMR274C';'YMR275C';'YMR276W';'YMR277W';'YMR278W';'YMR279C';'YMR280C
';'YMR281W';'YMR282C';'YMR283C';'YMR284W';'YMR285C';'YMR286W';'YMR287C';'YMR2
88W';'YMR289W';'YMR290C';'YMR291W';'YMR292W';'YMR293C';'YMR294W';'YMR295C';'Y
MR296C';'YMR297W';'YMR298W';'YMR299C';'YMR300C';'YMR301C';'YMR302C';'YMR303C'
;'YMR304W';'YMR305C';'YMR306W';'YMR307W';'YMR308C';'YMR309C';'YMR311C';'YMR31
2W';'YMR313C';'YMR314W';'YMR315W';'YMR316W';'YMR318C';'YMR319C';'YMR323W';'YM
R325W';'YNL001W';'YNL002C';'YNL003C';'YNL004W';'YNL005C';'YNL006W';'YNL007C';

'YNL008C';'YNL009W';'YNL012W';'YNL014W';'YNL015W';'YNL016W';'YNL020C';'YNL021
W';'YNL023C';'YNL024C-
A';'YNL025C';'YNL026W';'YNL027W';'YNL029C';'YNL030W';'YNL031C';'YNL032W';'YNL
036W';'YNL037C';'YNL038W';'YNL039W';'YNL041C';'YNL042W';'YNL044W';'YNL045W';'
YNL047C';'YNL048W';'YNL049C';'YNL051W';'YNL052W';'YNL053W';'YNL054W';'YNL055C
';'YNL056W';'YNL059C';'YNL061W';'YNL062C';'YNL063W';'YNL064C';'YNL065W';'YNL0
66W';'YNL067W';'YNL068C';'YNL069C';'YNL070W';'YNL071W';'YNL072W';'YNL073W';'Y
NL074C';'YNL075W';'YNL076W';'YNL077W';'YNL078W';'YNL079C';'YNL080C';'YNL081C'
;'YNL082W';'YNL083W';'YNL084C';'YNL085W';'YNL087W';'YNL088W';'YNL090W';'YNL09
1W';'YNL093W';'YNL094W';'YNL096C';'YNL097C';'YNL098C';'YNL099C';'YNL100W';'YN
L101W';'YNL102W';'YNL103W';'YNL104C';'YNL106C';'YNL107W';'YNL110C';'YNL111C';
'YNL112W';'YNL113W';'YNL116W';'YNL117W';'YNL118C';'YNL119W';'YNL121C';'YNL123
W';'YNL124W';'YNL125C';'YNL126W';'YNL127W';'YNL128W';'YNL129W';'YNL130C';'YNL
131W';'YNL132W';'YNL133C';'YNL135C';'YNL136W';'YNL137C';'YNL138W';'YNL138W-
A';'YNL139C';'YNL141W';'YNL142W';'YNL145W';'YNL147W';'YNL148C';'YNL149C';'YNL
151C';'YNL152W';'YNL153C';'YNL154C';'YNL156C';'YNL157W';'YNL158W';'YNL159C';'
YNL160W';'YNL161W';'YNL162W';'YNL163C';'YNL164C';'YNL166C';'YNL167C';'YNL169C
';'YNL172W';'YNL173C';'YNL175C';'YNL177C';'YNL178W';'YNL180C';'YNL182C';'YNL1
83C';'YNL185C';'YNL186W';'YNL187W';'YNL188W';'YNL189W';'YNL191W';'YNL192W';'Y
NL194C';'YNL197C';'YNL199C';'YNL201C';'YNL202W';'YNL204C';'YNL206C';'YNL207W'
;'YNL208W';'YNL209W';'YNL210W';'YNL212W';'YNL213C';'YNL214W';'YNL215W';'YNL21
6W';'YNL218W';'YNL219C';'YNL220W';'YNL221C';'YNL222W';'YNL223W';'YNL224C';'YN
L225C';'YNL227C';'YNL229C';'YNL230C';'YNL231C';'YNL232W';'YNL233W';'YNL234W';
'YNL236W';'YNL237W';'YNL238W';'YNL239W';'YNL240C';'YNL241C';'YNL242W';'YNL243
W';'YNL244C';'YNL245C';'YNL246W';'YNL247W';'YNL248C';'YNL249C';'YNL250W';'YNL
251C';'YNL252C';'YNL253W';'YNL254C';'YNL255C';'YNL256W';'YNL257C';'YNL258C';'
YNL259C';'YNL261W';'YNL262W';'YNL263C';'YNL264C';'YNL265C';'YNL267W';'YNL268W
';'YNL269W';'YNL270C';'YNL271C';'YNL272C';'YNL273W';'YNL274C';'YNL275W';'YNL2
77W';'YNL278W';'YNL279W';'YNL280C';'YNL281W';'YNL282W';'YNL283C';'YNL284C';'Y
NL286W';'YNL287W';'YNL288W';'YNL289W';'YNL290W';'YNL291C';'YNL292W';'YNL293W'
;'YNL294C';'YNL297C';'YNL298W';'YNL299W';'YNL301C';'YNL302C';'YNL304W';'YNL30
5C';'YNL306W';'YNL307C';'YNL308C';'YNL309W';'YNL310C';'YNL311C';'YNL312W';'YN
L313C';'YNL314W';'YNL315C';'YNL316C';'YNL317W';'YNL318C';'YNL321W';'YNL322C';
'YNL323W';'YNL325C';'YNL326C';'YNL327W';'YNL328C';'YNL329C';'YNL330C';'YNL331
C';'YNL332W';'YNL333W';'YNL334C';'YNL336W';'YNL339C';'YNR001C';'YNR002C';'YNR
003C';'YNR006W';'YNR007C';'YNR008W';'YNR009W';'YNR010W';'YNR011C';'YNR012W';'
YNR013C';'YNR015W';'YNR016C';'YNR017W';'YNR018W';'YNR019W';'YNR020C';'YNR022C
';'YNR023W';'YNR024W';'YNR026C';'YNR027W';'YNR028W';'YNR030W';'YNR031C';'YNR0
32C-
A';'YNR032W';'YNR033W';'YNR034W';'YNR035C';'YNR036C';'YNR037C';'YNR038W';'YNR
039C';'YNR041C';'YNR043W';'YNR044W';'YNR045W';'YNR046W';'YNR047W';'YNR048W';'
YNR049C';'YNR050C';'YNR051C';'YNR052C';'YNR053C';'YNR054C';'YNR055C';'YNR056C
';'YNR057C';'YNR058W';'YNR059W';'YNR060W';'YNR064C';'YNR067C';'YNR069C';'YNR0
72W';'YNR074C';'YNR075W';'YNR076W';'YOL001W';'YOL002C';'YOL003C';'YOL004W';'Y
OL005C';'YOL006C';'YOL007C';'YOL008W';'YOL009C';'YOL010W';'YOL011W';'YOL012C'
;'YOL013C';'YOL015W';'YOL016C';'YOL017W';'YOL018C';'YOL020W';'YOL021C';'YOL02
2C';'YOL023W';'YOL025W';'YOL026C';'YOL027C';'YOL028C';'YOL030W';'YOL031C';'YO
L032W';'YOL033W';'YOL034W';'YOL038W';'YOL039W';'YOL040C';'YOL041C';'YOL042W';
'YOL043C';'YOL044W';'YOL045W';'YOL049W';'YOL051W';'YOL052C';'YOL052C-
A';'YOL053W';'YOL054W';'YOL055C';'YOL056W';'YOL057W';'YOL058W';'YOL059W';'YOL
060C';'YOL061W';'YOL062C';'YOL063C';'YOL064C';'YOL065C';'YOL066C';'YOL067C';'
YOL068C';'YOL069W';'YOL070C';'YOL071W';'YOL072W';'YOL076W';'YOL077C';'YOL077W
-
A';'YOL078W';'YOL080C';'YOL081W';'YOL082W';'YOL083W';'YOL084W';'YOL086C';'YOL
088C';'YOL089C';'YOL090W';'YOL091W';'YOL093W';'YOL094C';'YOL095C';'YOL096C';'
YOL097C';'YOL100W';'YOL101C';'YOL102C';'YOL103W';'YOL104C';'YOL105C';'YOL108C
';'YOL109W';'YOL110W';'YOL111C';'YOL112W';'YOL113W';'YOL115W';'YOL116W';'YOL1

17W';'YOL119C';'YOL120C';'YOL121C';'YOL122C';'YOL123W';'YOL124C';'YOL125W';'Y
OL126C';'YOL127W';'YOL128C';'YOL129W';'YOL130W';'YOL132W';'YOL133W';'YOL135C'
;'YOL136C';'YOL137W';'YOL138C';'YOL139C';'YOL140W';'YOL141W';'YOL142W';'YOL14
3C';'YOL144W';'YOL145C';'YOL146W';'YOL147C';'YOL148C';'YOL149W';'YOL151W';'YO
L152W';'YOL154W';'YOL155C';'YOL156W';'YOL157C';'YOL158C';'YOL159C';'YOL159C-
A';'YOL161C';'YOL164W';'YOL165C';'YOR001W';'YOR002W';'YOR003W';'YOR004W';'YOR
005C';'YOR006C';'YOR007C';'YOR008C';'YOR009W';'YOR010C';'YOR011W';'YOR014W';'
YOR016C';'YOR017W';'YOR018W';'YOR019W';'YOR020C';'YOR023C';'YOR025W';'YOR026W
';'YOR027W';'YOR028C';'YOR030W';'YOR031W';'YOR032C';'YOR033C';'YOR034C';'YOR0
35C';'YOR036W';'YOR037W';'YOR038C';'YOR039W';'YOR040W';'YOR042W';'YOR043W';'Y
OR044W';'YOR045W';'YOR046C';'YOR047C';'YOR048C';'YOR049C';'YOR051C';'YOR052C'
;'YOR054C';'YOR056C';'YOR057W';'YOR058C';'YOR060C';'YOR061W';'YOR063W';'YOR06
4C';'YOR065W';'YOR066W';'YOR067C';'YOR068C';'YOR069W';'YOR070C';'YOR071C';'YO
R073W';'YOR074C';'YOR075W';'YOR076C';'YOR077W';'YOR078W';'YOR079C';'YOR080W';
'YOR081C';'YOR083W';'YOR084W';'YOR085W';'YOR086C';'YOR087W';'YOR089C';'YOR090
C';'YOR091W';'YOR092W';'YOR094W';'YOR095C';'YOR096W';'YOR098C';'YOR099W';'YOR
100C';'YOR101W';'YOR103C';'YOR104W';'YOR106W';'YOR107W';'YOR108W';'YOR109W';'
YOR110W';'YOR112W';'YOR113W';'YOR115C';'YOR116C';'YOR117W';'YOR118W';'YOR119C
';'YOR120W';'YOR122C';'YOR123C';'YOR124C';'YOR125C';'YOR126C';'YOR127W';'YOR1
28C';'YOR129C';'YOR130C';'YOR132W';'YOR133W';'YOR134W';'YOR136W';'YOR137C';'Y
OR138C';'YOR140W';'YOR141C';'YOR142W';'YOR143C';'YOR144C';'YOR145C';'YOR147W'
;'YOR148C';'YOR149C';'YOR150W';'YOR151C';'YOR153W';'YOR155C';'YOR156C';'YOR15
7C';'YOR158W';'YOR159C';'YOR160W';'YOR161C';'YOR162C';'YOR163W';'YOR164C';'YO
R165W';'YOR166C';'YOR167C';'YOR168W';'YOR171C';'YOR172W';'YOR173W';'YOR174W';
'YOR175C';'YOR176W';'YOR177C';'YOR178C';'YOR179C';'YOR180C';'YOR181W';'YOR182
C';'YOR184W';'YOR185C';'YOR187W';'YOR188W';'YOR189W';'YOR190W';'YOR191W';'YOR
192C';'YOR193W';'YOR194C';'YOR195W';'YOR196C';'YOR197W';'YOR198C';'YOR201C';'
YOR202W';'YOR204W';'YOR205C';'YOR206W';'YOR207C';'YOR208W';'YOR209C';'YOR210W
';'YOR211C';'YOR212W';'YOR213C';'YOR215C';'YOR216C';'YOR217W';'YOR219C';'YOR2
21C';'YOR222W';'YOR223W';'YOR224C';'YOR226C';'YOR227W';'YOR228C';'YOR229W';'Y
OR230W';'YOR231W';'YOR232W';'YOR233W';'YOR234C';'YOR236W';'YOR237W';'YOR239W'
;'YOR241W';'YOR242C';'YOR243C';'YOR244W';'YOR245C';'YOR247W';'YOR249C';'YOR25
0C';'YOR251C';'YOR252W';'YOR253W';'YOR254C';'YOR255W';'YOR256C';'YOR257W';'YO
R258W';'YOR259C';'YOR260W';'YOR261C';'YOR262W';'YOR264W';'YOR265W';'YOR266W';
'YOR267C';'YOR269W';'YOR270C';'YOR272W';'YOR273C';'YOR274W';'YOR275C';'YOR276
W';'YOR278W';'YOR279C';'YOR280C';'YOR281C';'YOR283W';'YOR284W';'YOR285W';'YOR
286W';'YOR287C';'YOR288C';'YOR290C';'YOR291W';'YOR293W';'YOR294W';'YOR295W';'
YOR297C';'YOR298C-
A';'YOR298W';'YOR299W';'YOR301W';'YOR302W';'YOR303W';'YOR304W';'YOR305W';'YOR
306C';'YOR307C';'YOR308C';'YOR310C';'YOR311C';'YOR312C';'YOR313C';'YOR315W';'
YOR316C';'YOR317W';'YOR319W';'YOR320C';'YOR321W';'YOR322C';'YOR323C';'YOR324C
';'YOR326W';'YOR327C';'YOR328W';'YOR329C';'YOR330C';'YOR332W';'YOR334W';'YOR3
35C';'YOR336W';'YOR337W';'YOR339C';'YOR340C';'YOR341W';'YOR344C';'YOR346W';'Y
OR347C';'YOR348C';'YOR349W';'YOR350C';'YOR351C';'YOR353C';'YOR354C';'YOR355W'
;'YOR356W';'YOR357C';'YOR358W';'YOR359W';'YOR360C';'YOR361C';'YOR362C';'YOR36
3C';'YOR367W';'YOR368W';'YOR369C';'YOR370C';'YOR371C';'YOR372C';'YOR373W';'YO
R374W';'YOR375C';'YOR377W';'YOR380W';'YOR381W';'YOR382W';'YOR383C';'YOR384W';
'YOR386W';'YOR388C';'YOR391C';'YOR393W';'YOR394W';'YPL001W';'YPL002C';'YPL003
W';'YPL004C';'YPL005W';'YPL006W';'YPL007C';'YPL008W';'YPL009C';'YPL010W';'YPL
011C';'YPL012W';'YPL013C';'YPL015C';'YPL016W';'YPL017C';'YPL018W';'YPL019C';'
YPL020C';'YPL021W';'YPL022W';'YPL023C';'YPL024W';'YPL026C';'YPL027W';'YPL028W
';'YPL029W';'YPL030W';'YPL031C';'YPL032C';'YPL033C';'YPL036W';'YPL037C';'YPL0
38W';'YPL040C';'YPL042C';'YPL043W';'YPL045W';'YPL046C';'YPL047W';'YPL048W';'Y
PL049C';'YPL050C';'YPL051W';'YPL052W';'YPL053C';'YPL054W';'YPL055C';'YPL057C'
;'YPL058C';'YPL059W';'YPL060W';'YPL061W';'YPL063W';'YPL064C';'YPL065W';'YPL06
9C';'YPL070W';'YPL072W';'YPL074W';'YPL075W';'YPL076W';'YPL078C';'YPL079W';'YP
L081W';'YPL082C';'YPL083C';'YPL084W';'YPL085W';'YPL086C';'YPL087W';'YPL089C';

```
'YPL090C';'YPL091W';'YPL092W';'YPL093W';'YPL094C';'YPL095C';'YPL096C-
A';'YPL096W';'YPL097W';'YPL098C';'YPL099C';'YPL100W';'YPL101W';'YPL103C';'YPL
104W';'YPL105C';'YPL106C';'YPL110C';'YPL111W';'YPL112C';'YPL115C';'YPL116W';'
YPL117C';'YPL118W';'YPL119C';'YPL120W';'YPL121C';'YPL122C';'YPL123C';'YPL124W
';'YPL125W';'YPL126W';'YPL127C';'YPL128C';'YPL129W';'YPL130W';'YPL131W';'YPL1
32W';'YPL133C';'YPL134C';'YPL135W';'YPL137C';'YPL138C';'YPL139C';'YPL140C';'Y
PL141C';'YPL143W';'YPL144W';'YPL145C';'YPL146C';'YPL147W';'YPL148C';'YPL149W'
;'YPL151C';'YPL152W';'YPL153C';'YPL154C';'YPL155C';'YPL156C';'YPL157W';'YPL15
8C';'YPL159C';'YPL160W';'YPL161C';'YPL163C';'YPL164C';'YPL165C';'YPL166W';'YP
L167C';'YPL169C';'YPL170W';'YPL171C';'YPL172C';'YPL173W';'YPL174C';'YPL175W';
'YPL176C';'YPL177C';'YPL178W';'YPL179W';'YPL180W';'YPL181W';'YPL183C';'YPL183
W-A';'YPL184C';'YPL186C';'YPL187W';'YPL188W';'YPL189C-
A';'YPL189W';'YPL190C';'YPL192C';'YPL193W';'YPL194W';'YPL195W';'YPL196W';'YPL
198W';'YPL200W';'YPL201C';'YPL202C';'YPL203W';'YPL204W';'YPL206C';'YPL207W';'
YPL208W';'YPL209C';'YPL210C';'YPL211W';'YPL212C';'YPL213W';'YPL214C';'YPL215W
';'YPL217C';'YPL218W';'YPL219W';'YPL220W';'YPL221W';'YPL223C';'YPL224C';'YPL2
25W';'YPL226W';'YPL227C';'YPL228W';'YPL230W';'YPL231W';'YPL232W';'YPL233W';'Y
PL234C';'YPL235W';'YPL237W';'YPL239W';'YPL240C';'YPL241C';'YPL242C';'YPL243W'
;'YPL244C';'YPL246C';'YPL248C';'YPL249C';'YPL249C-
A';'YPL250C';'YPL252C';'YPL253C';'YPL254W';'YPL255W';'YPL256C';'YPL258C';'YPL
259C';'YPL262W';'YPL263C';'YPL265W';'YPL266W';'YPL267W';'YPL268W';'YPL269W';'
YPL270W';'YPL271W';'YPL273W';'YPL274W';'YPL281C';'YPL282C';'YPL283C';'YPR001W
';'YPR002W';'YPR004C';'YPR005C';'YPR006C';'YPR007C';'YPR008W';'YPR009W';'YPR0
10C';'YPR016C';'YPR017C';'YPR018W';'YPR019W';'YPR020W';'YPR021C';'YPR023C';'Y
PR024W';'YPR025C';'YPR026W';'YPR028W';'YPR029C';'YPR030W';'YPR031W';'YPR032W'
;'YPR033C';'YPR034W';'YPR035W';'YPR036W';'YPR036W-
A';'YPR037C';'YPR040W';'YPR041W';'YPR042C';'YPR043W';'YPR045C';'YPR046W';'YPR
047W';'YPR048W';'YPR049C';'YPR051W';'YPR052C';'YPR054W';'YPR055W';'YPR056W';'
YPR057W';'YPR058W';'YPR060C';'YPR061C';'YPR062W';'YPR065W';'YPR066W';'YPR067W
';'YPR068C';'YPR069C';'YPR070W';'YPR072W';'YPR073C';'YPR074C';'YPR075C';'YPR0
79W';'YPR080W';'YPR081C';'YPR082C';'YPR083W';'YPR085C';'YPR086W';'YPR088C';'Y
PR091C';'YPR093C';'YPR094W';'YPR095C';'YPR096C';'YPR097W';'YPR098C';'YPR100W'
;'YPR101W';'YPR102C';'YPR103W';'YPR104C';'YPR105C';'YPR106W';'YPR107C';'YPR10
8W';'YPR110C';'YPR111W';'YPR112C';'YPR113W';'YPR115W';'YPR116W';'YPR118W';'YP
R119W';'YPR120C';'YPR121W';'YPR122W';'YPR124W';'YPR125W';'YPR127W';'YPR128C';
'YPR129W';'YPR131C';'YPR132W';'YPR133C';'YPR133W-
A';'YPR134W';'YPR135W';'YPR137W';'YPR138C';'YPR139C';'YPR140W';'YPR141C';'YPR
143W';'YPR144C';'YPR145W';'YPR148C';'YPR149W';'YPR151C';'YPR152C';'YPR153W';'
YPR154W';'YPR155C';'YPR156C';'YPR158W';'YPR159W';'YPR160W';'YPR161C';'YPR162C
';'YPR163C';'YPR164W';'YPR165W';'YPR166C';'YPR167C';'YPR168W';'YPR169W';'YPR1
71W';'YPR173C';'YPR175W';'YPR176C';'YPR178W';'YPR179C';'YPR180W';'YPR181C';'Y
PR182W';'YPR183W';'YPR184W';'YPR185W';'YPR186C';'YPR187W';'YPR188C';'YPR189W'
;'YPR190C';'YPR191W';'YPR192W';'YPR193C';'YPR194C';'YPR198W';'YPR199C';'YPR20
0C';'YPR201W';'YPR204W';};

% To do set comparison stuff, we need to strip the parantheses from around
% the gene names in unique_single_gene_reaction_genes.

for i=1:length(unique_single_gene_reaction_genes)
    pretty_unique_single_gene_reaction_genes{i}= ...
        unique_single_gene_reaction_genes{i}(2:8);
end

included_dubiousORFs = setdiff(...
    pretty_unique_single_gene_reaction_genes, verified_genes);
%there are 4 dubious ORFs on the list of rxns catalyzed by 1 gene
```

```
included_essential=intersect(...
    pretty_unique_single_gene_reaction_genes,sgd_essential_genes);
%there are 88 essential genes on the list of rxns catalyzed by 1 gene

candidate_singles =
setxor(included_dubiousORFs,pretty_unique_single_gene_reaction_genes);
candidate_singles = setxor(included_essential,candidate_singles);

%So, now there are 287 genes on the list of possible candidates (a
%manageable number). Patrick had a list of 304 single gene possibilities on
%his list. What are the differences?

%that list includes:
%'YBR184W' - SGD says uncharacterized
%'YDR111C' - SGD says uncharacterized
%'YFR055W' - SGD says uncharacterized
%'YJL167W' - SGD says essential
%'YML082W' - SGD says uncharacterized
%'YNL247W' - on my essential list

%I'm okay leaving them out.

%As I work, I've found other genes which shouldn't be included
others_to_remove = {'YNL280C'}; %SGD annotates as inviable in aerobic

candidate_singles = setxor(others_to_remove,candidate_singles);

%and I want to add YDR111C back in to the list of candidates because alt2
%is in Kennedy's KO1 (it was removed b/c uncharacterized).

candidate_singles(end+1) = {'YDR111C'};
save('candidate_singles','candidate_singles');

%%
%So now I have a list of genes that annotate reactions that are only
%annotated with a single genes. I (think) I can knock out any three of
%these genes and look for increased formic acid secretion.

%Begin by generating a list of all possible double deletions of
%non-essential genes.
doubles = combntns((1:length(candidate_singles)),2); %there are 41041
length(doubles)
save('doubles','doubles');

formate_ex=find(strcmp('Formate exchange',model.rxnNames));

%make sure we're using the gurobi solver (glpk gives different answers!)
changeCobraSolver('gurobi');

fprintf(' \r');
fprintf('LP solver changed to gurobi.\n');
fprintf(' \r');

double_formate_fluxes=[];

%now loop over all possible double deletions, and look for formate
%production (there is probably a more clever way to do this - email and let
```

```
%me know how to do it better! bheavner@gmail.com)

h = waitbar(0,'Checking double knockouts ...');

for i=1:length(doubles)
    if mod(i,100) == 0
        waitbar(i/length(doubles),h);
    end
    clear mex;
    [newmodel,hasEffect,constrRxnNames,deletedGenes] = ...
        deleteModelGenes(model,candidate_singles(doubles(i,:)));
    if hasEffect
        flux=optimizeCbModel(newmodel,[],'one');
        if numel(flux.x)==0
            i
        elseif (flux.x(formate_ex)>0 && (flux.f>0.5518)) %formate flux + >20%
WT growth rate
            double_formate_fluxes = [double_formate_fluxes; i
flux.x(formate_ex) flux.f];
            sprintf('%12i %12.4f %12.4f\n',i, flux.x(formate_ex), flux.f)
        end
    end
end
close(h);

save('double_formate_fluxes','double_formate_fluxes');

% the double_formate_fluxes variable is now a 3-column vector. The first
% column is the "doubles" entry screened, the second is the formate flux,
% and the third is the growth rate of that strain. To find which deletion
% gives a particular formate flux, do something like this:
max_formate_flux=max(double_formate_fluxes(:,2));
index=find(double_formate_fluxes(:,2)==max_formate_flux);
max_double=double_formate_fluxes(index,1);
singles_index=doubles(max_double,:);
max_formate_genes=candidate_singles(singles_index)
```

Supplementary File 5.1.3: three_knockout2.m script

```
% FILE NAME: three-knockout2
%
% DATE CREATED: 19 January 2012
%
% PROGRAMMER: B. Heavner
% Department of Biological
% and Environmental Engineering
% Cornell University
% Ithaca, NY 14853
%
% LAST REVISED: 10 July 2012
%
% REVISIONS:
%
% 7/10/12 - BDH cleaned up code for publication
%
% 5/14/12 - BH changed to load list of common candidate singles. Previous
% screen on multiple machines led to incorrect flux predictions; perhaps
% because each machine generated own candidate singles. Also, added some
% java memory statements to try to fix memory leaks observed with Y5 screen.
%
% 5/11/12 - BH added YDR111C as possible deletion candidate, fixed
% genetic background
%
% 2/9/12 - BH removed YNL280C as possible single, standardize on gurobi
% solver for speed and consistency
%
% 2/7/12 - BH added fdh mutant, constraint on min growth rate, growth rate
% to formate_fluxes variable
%
% PURPOSE:
% To attempt to reproduce the model results reported by Kennedy et al. -
% that is, to compare all viable 3-knockout mutants in iND750 to find those
% that will cause formic acid secretion. Kennedy et al. report constructing
% a alt2, fum1, zwf1 mutant that does this.
%
% In this script, I apply a "global" search for triple deletions that
% excrete formic acid to the the constrained iND750 I made in the
% three-knockout1 script (the variable "model", which has the same media as
% Kennedy et al.)
%
% REFERENCE:
% Kennedy CJ, Boyle PM, Waks Z, Silver PA. "Systems-Level Engineering of
% Nonfermentative Metabolism in Yeast" Genetics (2009) 183:385-397. DOI:
% 10.1534/genetics.109.105254
%
% VARIABLES:
% this script requires the model.mat variable built by three-kncokout1.m in
% the working directory.
%
% EXPECTED OUTPUT:
% a list of triple-gene deleted mutants that are predicted to excrete
% formate.
%
```

```
%-------------------------------------------------------------

%%
% If previously run, just load the list of triple deletion strains, don't
% have to regenerate every time - can uncomment the following 4 lines. If
% you do this, be sure to comment out line 86.

%load('triplets');
%fprintf(' \r');
%fprintf('Triplets loaded.\n');
%fprintf(' \r');

%%
%first, load the modified iND750 model with the Kennedy constraints. This
%variable generated from iND750 by the three_knockout1.m script.

load model.mat;

fprintf(' \r');
fprintf('Constrained iND750 model loaded.\n');
fprintf(' \r');

%load list of genes that can be deleted, generated by two_knockout2.m code.
load('candidate_singles');

fprintf(' \r');
fprintf('Candidate singles loaded.\n');
fprintf(' \r');

%generating a list of all possible tripple deletions of non-essential
%genes. COMMENT OUT THE FOLLOWING LINE IF YOU LOAD TRIPLETS FROM FILE
triplets = combntns((1:length(candidate_singles)),3); %there are 3898895

%make sure we're using the gurobi solver (glpk gives different answers!)
changeCobraSolver('gurobi');

fprintf(' \r');
fprintf('LP solver changed to gurobi.\n');
fprintf(' \r');

h = waitbar(0,'Checking triple knockouts ...');
%this is not a fast process!

%to run on multiple machines, copy the necessary variables, and adjust
%lower and upper bounds (as demonstrated below for running on 7 computer)
%so they don't overlap.

%lower=1; upper=floor(length(triplets)/7);
%lower=floor(length(triplets)/7)+1; upper=floor(2/7*length(triplets));
%lower=floor(2/7*length(triplets))+1; upper=floor(3/7*length(triplets));
%lower=floor(3/7*length(triplets))+1; upper=floor(4/7*length(triplets));
%lower=floor(4/7*length(triplets))+1; upper=floor(5/7*length(triplets));
%lower=floor(5/7*length(triplets))+1; upper=floor(6/7*length(triplets));
%lower=floor(6/7*length(triplets))+1; upper=length(triplets);
lower=1; upper=length(triplets);

%save results as we go in case things crash
```

```
fileID = fopen('triple_formate_fluxes_1.txt','w');

for i=lower:upper
    if mod(i,100) == 0
        waitbar((i-lower+1)/(upper-lower),h);
    end
    clear mex; %an attempt to prevent crashes from gurobi_mex memory leaks
    [newmodel,hasEffect,constrRxnNames,deletedGenes] = ...
        deleteModelGenes(model,candidate_singles(triplets(i,:)));
    if hasEffect
        formate_ex=find(strcmp('Formate exchange',model.rxnNames));
        flux=optimizeCbModel(newmodel,[],'one');
        if numel(flux.x)==0
            i
        elseif (flux.f>.5518) %20% of WT
            if (flux.x(formate_ex)>0)
               fprintf(fileID,'%12i %12.4f %12.4f\n',i, flux.x(formate_ex),
flux.f);
                sprintf('%12i %12.4f %12.4f\n',i, flux.x(formate_ex), flux.f)
%print to screen, too, so we can see where things left off if there's a crash
            end
        end
    end
end

close(h);
fclose(fileID);

date=num2str(floor(now));
backup_name=strcat('triple_formate_fluxes_1',date,'.txt');
copyfile('triple_formate_fluxes_1.txt',backup_name)

%%
% the triple_formate_fluxes1.txt file is now a 3-column text file. The first
% column is the "triplets" entry screened, the second is the formate flux,
% and the third is the growth rate of that strain. To find which deletion
% gives a particular formate flux, do something like this:

% %First, load the screen results and print some very basic comparisons
% results=spcread('triple_formate_fluxes_1.txt',3);
% load('candidate_singles')
% load('triplets')
%
% %total number screened
% screened=length(triplets);
%
% %number of formate-producing mutants
% number_of_formate_producers=length(results);
%
% %percentage of screened mutants which produce formate
% percentage_of_formate_producers=
(number_of_formate_producers/screened)*100;
%
% %maximum formate flux
% maximum_formate=max(results(:,2));
%
% %list of mutants that have the max formate flux
```

```
% max_mutants=find(results(:,2)==maximum_formate);
% number_of_max_fluxes=length(max_mutants);
```

Supplementary File 5.2.1:Y5_ three_knockout1.m script

```
% FILE NAME: Y5_three_knockout1_unmodified
%
% DATE CREATED: 14 May 2012
%
% PROGRAMMER: B. Heavner
% Department of Biological
% and Environmental Engineering
% Cornell University
% Ithaca, NY 14853
%
% LAST REVISED: 11 July 2012
%
% REVISED BY: B. Heavner
%
% REVISIONS:
% 7/11/12 - BDH - cleaned up for publication
%
%
% PURPOSE:
% To attempt to reproduce the model results reported by Kennedy et al. with
% the unmodified Yeast 5.01 model.
%
% In this script, I begin by applying constraints to Y5 to duplicate the
% Kennedy et al. media and mutants to see if I still get a predicted formic
% acid secretion with the mutants reported in Kennedy et al. Table 1.
%
% REFERENCE:
% Kennedy CJ, Boyle PM, Waks Z, Silver PA. "Systems-Level Engineering of
% Nonfermentative Metabolism in Yeast" Genetics (2009) 183:385-397. DOI:
% 10.1534/genetics.109.105254
%
% VARIABLES:
% none - it's a script. But you need Y5.mat in the working directory (or
% load the Y5 GEM using readcbmodel, and comment out the "load" line of the
% script.
%
% EXPECTED OUTPUT:
% growth rates and fluxes for the first 5 mutants listed in Table 1 of
% Kennedy et al. Version 5.01 of the Yeast model does not predict formate
% production for these strains.
%
% REQUIRED SOFTWARE:
% this script calls functions from the COBRA Toolbox and uses the Gurobi
% solver.
%
%-------------------------------------------------------------

%%
%first, load the Y5.01 model.

load Y5.mat;
model=Y5;

fprintf(' \r');
```

```
fprintf('Yeast model v 5.01 loaded.\n');
fprintf(' \r');

%make sure we're using the gurobi solver (glpk gives different answers!)
changeCobraSolver('gurobi');

fprintf('LP solver changed to gurobi.\n');
fprintf(' \r');

%%
%next modify the changed model to add some strain-specific changes.

%add the alt2 gene, which is not in Y5.01 b/c it's a putative ORF in SGD.
model=addReaction(model,'NRXN4',{'s_0180', 's_0955', 's_0991', ...
    's_1399'}, [-1 -1 1 1], true, -1000, 1000, 0, '', 'YDR111C');

%add the fdh2 gene, which isn't annotated b/c it's not in the SGD
%background strain - rxn 445 can be catalyzed by FDH2 in CEN.PK 113-7D
%strain (see pmid: 11921099)
model=changeGeneAssociation(model,model.rxns(445),...
    '(YOR388C or (YPL275W AND YPL276W))');

%paper wt background is fdh1 fdh2 deletion
[model,hasEffect,constrRxnNames,deletedGenes] = ...
        deleteModelGenes(model,{'YOR388C','YPL275W','YPL276W'});

%%
%next, make the media the same as that in Kennedy et al. (based on their
%supplemental table 1).
%
%NOTE: the COBRA exchange reaction convention means that I need to use the
%opposite sign than Kennedy et al.
%
%NOTE 2: Table S1 lists "pydxn", "4abz", "nac", and "dhf" as having an
%upper bound of 0.5. I think these are: Pyridoxine (met 875),
%4-Aminobenzoate (met 98), Nicotinate (met 714), and 7,8-Dihydrofolate (met
%383). They don't have exchange reactions in iND750. Kennedy et al. may
%have added them, but they don't have fluxes in Table S2, so they're not
%essential.
%
%NOTE 3: I'm guessing rxn names from their met names.

%start with a clean slate: set all exchange reactions to upper bound = 1000
%and lower bound = 0 (ie, unconstrained excretion, no uptake)

exchangeRxns = findExcRxns(model);
model.lb(exchangeRxns)=0;
model.ub(exchangeRxns)=1000;

%the following exchange rxns have an upper bound and lower bound of 20 in
%Table S1 (ie, a fixed rate of uptake). I'm switching signs, so do bounds
%of -20.

ub_lb_20={'D-glucose exchange'};

for i=1:length(ub_lb_20)
    index=find(strcmp(ub_lb_20(i),model.rxnNames));
```

```
    model.lb(index)=-20;
    model.ub(index)=-20;
end

%the following have upper and lower bounds of inf in Table S1.
%
%NOTE: oxygen not in table S1, but added to do aerobic, based on
%text of paper

ub_inf={'water exchange', 'ammonium exchange', 'phosphate exchange', ...
    'sulphate exchange', 'sodium exchange', 'potassium exchange', ...
    'carbon dioxide exchange', 'oxygen exchange'};

for i=1:length(ub_inf)
    index=find(strcmp(ub_inf(i),model.rxnNames));
    model.lb(index)=-1000;
end

%the following have an upper bound of 0.5 in Table S1 (but I'm switching
%signs, so do a lower bound of -0.5).

ub_05 = {'L-asparagine exchange', 'L-aspartate exchange', ...
    'L-valine exchange', 'L-tyrosine exchange', ...
    'L-tryptophan exchange', 'L-threonine exchange', ...
    'L-serine exchange', 'L-proline exchange', ...
    'L-phenylalanine exchange', 'L-methionine exchange', ...
    'L-lysine exchange', 'L-leucine exchange', ...
    'L-isoleucine exchange', 'L-histidine exchange', ...
    'glycine exchange', 'L-glutamate exchange', ...
    'L-glutamine exchange', 'L-cysteine exchange', ...
    'L-arginine exchange', 'L-alanine exchange' ...
    'riboflavin exchange', 'thiamine(1+) exchange', ...
    'zymosterol exchange', 'uracil exchange', ...
    '(R)-pantothenate exchange', 'linoleic acid exchange', ...
    'myo-inositol exchange', 'palmitoleate exchange', ...
    'ergosterol exchange', 'biotin exchange', 'adenine exchange'};

%note: Y5 doesn't have octadecynoate (n-C18:2) or octadecenoate (n-C18:1)
%exchange reactions. I used linoleic (18:1). It also doesn't have
%hexadecenoate (n-C16:1). I used palmitoleate (16:1) for that. I don't
%think they should really matter for aerobic simulations, though.

for i=1:length(ub_05)
    index=find(strcmp(ub_05(i),model.rxnNames));
    model.lb(index)=-0.5; %constrained uptake
end

%the following have a lower bound of -inf in Table S1 (but I'm switching
%signs, so do an upper bound of 1000)
%
% Note: 'Deoxycytidine exchange' not included in Table S1, but has a flux
% in Table S2, so I added it here.

lb_inf = {'(1->3)-beta-D-glucan exchange', ...
    'gamma-aminobutyrate exchange', '(S)-3-methyl-2-oxopentanoate exchange',
...
    ' 8-amino-7-oxononanoate exchange', 'L-arabinitol exchange', ...
```

```
    'acetaldehyde exchange', 'acetate exchange', 'adenosine exchange', ...
    ' 2-oxoglutarate exchange', 'allantoin exchange', ...
    'S-adenosyl-L-methionine exchange', 'D-arabinose exchange', ...
    'L-arabinose exchange', 'choline exchange', 'citrate(3-) exchange', ...
    '(R)-carnitine exchange', 'cytosine exchange', 'cytidine exchange', ...
    '2''-deoxyadenosine exchange', ' 7,8-diaminononanoate exchange', ...
    '2''-deoxyguanosine exchange', '2''-deoxyinosine exchange', ...
    'dTTP exchange', '2''-deoxyuridine exchange', 'ethanol exchange', ...
    'FMN exchange', 'palmitate exchange', 'H+ exchange', ...
    'hypoxanthine exchange', 'inosine exchange', '(S)-lactate exchange', ...
    '(S)-malate exchange', 'maltose exchange', 'D-mannose exchange', ...
    'melibiose exchange', 'S-methyl-L-methionine exchange', ...
    'NMN exchange', 'stearate exchange', ...
    'ornithine exchange', 'adenosine 3'',5''-bismonophosphate exchange', ...
    'peptide exchange', 'putrescine exchange', 'pyruvate exchange', ...
    'D-ribose exchange', 'D-glucitol exchange', 'L-glucitol exchange', ...
    'spermidine exchange', 'spermine exchange', 'L-sorbose exchange', ...
    'succinate exchange', 'sucrose exchange', ...
    'thiamine(1+) monophosphate exchange', 'thiamine(1+) diphosphate(1-)
exchange', ...
    'thymidine exchange', 'thymine exchange', 'alpha,alpha-trehalose
exchange', ...
    'myristate exchange', 'urea exchange', ...
    'uridine exchange', '9H-xanthine exchange', 'xanthosine exchange', ...
    'D-xylose exchange', 'xylitol exchange', 'deoxycytidine exchange'};

%note: Y5 doesn't have 4-Aminobutanoate exchangeI used gamma-aminobutyrate.
%Y5 doesn't have 5-Amino-4-oxopentanoate exchange. I used
%'(S)-3-methyl-2-oxopentanoate exchange'.

for i=1:length(lb_inf)
    index=find(strcmp(lb_inf(i),model.rxnNames));
    model.lb(index)=0;
    model.ub(index)=1000;
end

fprintf('Exhange reaction bounds set.\n');
fprintf(' \r');

%%
%next step: simulate the strains of Kennedy Table 1

%Kennedy reported in silico knockouts:
%1 fdh1=YOR388C fdh2=YPL275W/YPL276W alt2=YDR111C fum1=YPL262W zwf1=YNL241C
%2 fdh1 fdh2 aat2=YLR027C fum1=YPL262W zwf1=YNL241C
%3 fdh1 fdh2 cat2=YML042W fum1=YPL262W zwf1=YNL241C
%4 fdh1 fdh2 cat2=YML042W fum1=YPL262W rpe1=YJL121C
%5 fdh1 fdh2 cat2=YML042W fbp1=YLR377C fum1=YPL262W
%6 fdh1 fdh2 cat2=YML042W yat2=YER024W slc1=YDL052C
%7 fdh1 fdh2 cat2=YML042W yat2=YER024W cho1=YER026C
%8 fdh1 fdh2 cat2=YML042W yat2=YER024W alt2=YDR111C

%looking at Yeast 5.01:
%fdh1 = gene 848, rxns: 445
%fdh2 = gene added by my chages, rxns: 445
%alt2 - added to model, above
%fum1 = gene 883 rxns: 451 452
```

```
%zwf1 = gene 743 rxns: 466
%aat2 = gene 588 rxns: 216 218 681 683 1065
%cat2 = gene 659 rxns: 253 254
%rpe1 = gene 476 rxns: 984
%fbp1 = gene 645 rxns: 449
%yat2 = gene 263 rxns: 252
%slc1 = gene 130 rxns: 9
%cho1 = gene 264 rxns: 880 881 - essential for model.

%%
%check that FBA works on the model
wt_sln=optimizeCbModel(model,[],'one');

formate_ex=find(strcmp('formate exchange',model.rxnNames));

fprintf('The WT flux is: %6.4f.\n', wt_sln.f); %2.5889
fprintf('Formate exchange is reaction %d.\n', formate_ex); %0
fprintf('The formate exchange flux is: %6.4f.\n', wt_sln.x(formate_ex));
fprintf(' \r');

%make the knockouts and test them
[KO1,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YDR111C','YPL262W','YNL241C'});
KO1_sln=optimizeCbModel(KO1,[],'one');

fprintf('The KO1 flux is: %6.4f.\n', KO1_sln.f); %1.6065
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO1_sln.x(formate_ex));
%0.0001
fprintf(' \r');

[KO2,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YLR027C', 'YPL262W', 'YNL241C'});
KO2_sln=optimizeCbModel(KO2,[],'one');

fprintf('The KO2 flux is: %6.4f.\n', KO2_sln.f); %1.6065
fprintf('Formate exchange is reaction %d.\n', formate_ex); %
fprintf('The formate exchange flux is: %6.4f.\n', KO2_sln.x(formate_ex));
%0.0001
fprintf(' \r');

[KO3,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YPL262W', 'YNL241C'});
KO3_sln=optimizeCbModel(KO3,[],'one');

fprintf('The KO3 flux is: %6.4f.\n', KO3_sln.f); %1.6065
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO3_sln.x(formate_ex));
%0.0001
fprintf(' \r');

[KO4,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YPL262W', 'YJL121C'});
KO4_sln=optimizeCbModel(KO4,[],'one');

fprintf('The KO4 flux is: %6.4f.\n', KO4_sln.f); %1.6065
fprintf('Formate exchange is reaction %d.\n', formate_ex);
```

```
fprintf('The formate exchange flux is: %6.4f.\n', KO4_sln.x(formate_ex));
%0.0001
fprintf(' \r');

KO5=model;
[KO5,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YLR377C', 'YPL262W'});
KO5_sln=optimizeCbModel(KO5,[],'one');

fprintf('The KO5 flux is: %6.4f.\n', KO5_sln.f);
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO5_sln.x(formate_ex));
fprintf(' \r');

KO6=model;
[KO6,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YER024W', 'YDL052C'});
KO6_sln=optimizeCbModel(KO6,[],'one');

fprintf('The KO6 flux is: %6.4f.\n', KO6_sln.f);
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO6_sln.x(formate_ex));
fprintf(' \r');

KO7=model;
[KO7,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YER024W', 'YER026C'});
KO7_sln=optimizeCbModel(KO7,[],'one');

fprintf('The KO7 flux is: %6.4f.\n', KO7_sln.f);
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO7_sln.x(formate_ex));
fprintf(' \r');

KO8=model;
[KO8,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YER024W', 'YDR111C'});
KO8_sln=optimizeCbModel(KO8,[],'one');

fprintf('The KO8 flux is: %6.4f.\n', KO8_sln.f);
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO8_sln.x(formate_ex));
fprintf(' \r');
```

Supplementary File 5.3.1: apply_y5_changes.m script

```
% FILE NAME: apply_y5_changes
%
% DATE CREATED: 27 Feb 2012
%
% PROGRAMMER: B. Heavner
% Department of Biological
% and Environmental Engineering
% Cornell University
% Ithaca, NY 14853
%
% LAST REVISED: 11 July 2012
%
% REVISIONS:
% 11 July 2012 - BH cleaned up for publication
% 3 May 2012 - BH added changes from comparing to YeastCyc pathways
%
% REVISED BY: B. Heavner
%
% PURPOSE:
% To apply all changes I find to correct the Y5.01 model.
%
% REFERENCES:
% Various, included as comments to changes
%
% VARIABLES:
% none - it's a script. But you need Y5.mat in the working directory (or
% load the Y5 GEM using readcbmodel, and comment out the "load" line of the
% script.
%
% EXPECTED OUTPUT:
% updated model based on Y5.01
%
% REQUIRED SOFTWARE:
% this script calls functions from the COBRA Toolbox.
%
%-------------------------------------------------------------

%%
%first, load the Y5.01 model. This matlab variable was made by running
% readcbmodel to load a .sbml file downloaded from yeast.sf.net on
% 2/9/12

load Y5.mat;
model=Y5;

fprintf(' \r');
fprintf('Yeast model v 5.01 loaded.\n');
fprintf(' \r');

%make sure we're using the gurobi solver (glpk gives different answers!)
changeCobraSolver('gurobi');

fprintf('LP solver changed to gurobi.\n');
fprintf(' \r');
```

```
%check the media:
exchangeRxns = findExcRxns(model);
temp=find(exchangeRxns);
temp2=find(model.lb(temp));
media_indexes=temp(temp2);

fprintf('Media rxns\n')
fprintf('rxnName \t lb\n')
for i=1:length(media_indexes)
    fprintf('%s\t%g\n',model.rxnNames{media_indexes(i)},
model.lb(media_indexes(i)));
end

fprintf(' \r');

%Check the solution prior to any changes
unchangedsolution = optimizeCbModel(model,[],'one');

%%
%model updates and fixes - note: add new to the end - numbers may change!

%fix Y5 over-constrained rxns
%refs: http://www.ncbi.nlm.nih.gov/pubmed/8852837
%http://www.ncbi.nlm.nih.gov/pubmed/10871621
%http://www.ncbi.nlm.nih.gov/pubmed/1916088

model=changeRxnBounds(model,model.rxns([733 447 446]),-1000,'l');
model=changeRxnBounds(model,model.rxns([733 447 446]),1000,'u');

%some changes in the NAD biosynthesis pathway - ref: PMID: 18205391
%"quinaldic acid (dehydroxylated KA, which is produced directly from KA)"
model=addReaction(model,'NRXN1','s_0147 -> s_1403 + s_0803');

% from PMID: 12771147
model=removeRxns(model,'r_0768');
%model=addReaction(s_0423 + s_0633 + s_0794 + s_1198 + s_0991');
model=addReaction(model,'r_0768',{'s_0999', 's_0434', 's_0591', ...
    's_0803', 's_0423', 's_0633', 's_0794', 's_1198', 's_0991'}, ...
    [-1 -1 -1 -1 1 1 1 1 1], false, 0, 1000,'',...
    'YHR074W');

%ref: http://www.yeastgenome.org/cgi-bin/locus.fpl?locus=YMR189W
%I'm not sure about these - BH 2/29/12
%model.lb([505 508])=-1000; model.ub([505 508])=1000;

%ref https://en.wikipedia.org/wiki/Malate_dehydrogenase
changeRxnBounds(model,model.rxns(713),-1000,'l');
changeRxnBounds(model,model.rxns(713),1000,'u');

%%
%corrections to mitochondrial NAD metabolism, 3/1/12

%see also http://dx.doi.org/10.1111/j.1574-6976.2001.tb00570.x (ref added
%3/27/12, after examining individual reactions)

%add CO2 as product to rxn 545, make irreversible
%ref:
```

```
%pathway.yeastgenome.org/YEAST/new-image?type=PATHWAY&object=LYSINE-AMINOAD-
PWY&detail-level=2
model.S(460,545)=1;
model=changeRxnBounds(model,model.rxns(545),0,'l');

%http://www.jbc.org/content/281/3/1524.long "The main role of Ndt1p and
%Ndt2p is to import NAD+ into mitochondria by unidirectional transport or
%by exchange with intramitochondrially generated (d)AMP and (d)GMP." First,
%make unidirectional.
model=changeRxnBounds(model,model.rxns(1228),0,'l');

%next make the function to be irreversible exchange, not symport
%dAMP
model.S(584,1229)=1; model.S(585,1229)=-1;
model=changeRxnBounds(model,model.rxns(1229),0,'l');

%AMP
model.S(423,1230)=1; model.S(424,1230)=-1;
model=changeRxnBounds(model,model.rxns(1230),0,'l');

%dGMP
model.S(615,1231)=1; model.S(616,1231)=-1;
model=changeRxnBounds(model,model.rxns(1231),0,'l');

%GMP
model.S(782,1232)=1; model.S(784,1232)=-1;
model=changeRxnBounds(model,model.rxns(1232),0,'l');

%http://www.jbc.org/content/281/3/1524.long "It transported NAD+ and, to a
%lesser extent, (d)AMP and (d)GMP but virtually not ?-NAD+, NADH, NADP+, or
%NADPH." - so no to rxn 1233 (NOTE: I'm NOT SURE OF THIS)
model=changeRxnBounds(model,model.rxns(1233),0,'l');
model=changeRxnBounds(model,model.rxns(1233),0,'u');

%rxn 12 looks like what PUT2 catalyzes (which is reversible).
%PUT2 is model.genes(407)
model=changeGeneAssociation(model,model.rxns(12),'YHR037W');
model=changeRxnBounds(model,model.rxns(12),-1000,'l');

%rxn 31 is http://www.genome.jp/dbget-bin/www_bget?rn:R01199 - per KEGG,
%not in yeast
model=changeRxnBounds(model,model.rxns(31),0,'l');
model=changeRxnBounds(model,model.rxns(31),0,'u');

%The Y5 reference for rxn 71 is a hypothetical pathway. Kegg says "poorly
%characterized". Axe it for now (but maybe need for modelling purposes?)
model=changeRxnBounds(model,model.rxns(71),0,'l');
model=changeRxnBounds(model,model.rxns(71),0,'u');

%Per SGD, rxn 165 can be catalyzed by ADH3 or ADH4.
%YMR083W is model.genes(684). YGL256W is model.genes(339)
model=changeGeneAssociation(model,model.rxns(165),'YMR083W or YGL256W');

%looking at rxn 201, SGD says YER073W uses NADP, not NAD:
%YOR374W is model.genes(845).
model=changeGeneAssociation(model,model.rxns(201),'YOR374W');
```

```
%and the activity of rxns 201, 174, and 176 is different than the SGD
%activity
%http://pathway.yeastgenome.org/YEAST/NEW-IMAGE?type=REACTION-IN-
PATHWAY&object=RXN3O-274&detail-level=2
%the sgd activity is rxn 175.

% On a second look, I don't see this, and rxn 174 has a flux in Kennedy KO1
KO2 KO3 KO4
%model.lb([201 174 176])=0; model.ub([201 174 176])=0;

%rxn 672 and 673 aren't in SGD for PUT2, but is in KEGG (which has them as
%reversible - see http://www.genome.jp/dbget-bin/www_bget?C04281). So I'll
%consider them low confidence, but leave them in and reversible for now.
model=changeRxnBounds(model,model.rxns([672 673]),-1000,'l');

%rxn 766 is right, but should also have NADH kinase activity, per SGD
%http://www.yeastgenome.org/cgi-bin/locus.fpl?locus=YPL188W
%this rxn is reaction 771 (r_0772), annotated YEL041W or YJR049C.
%perhaps both reactions should be annotated "YEL041W or YJR049C or YPL188W"
%YPL188W is model.genes(873)
model=changeGeneAssociation(model,model.rxns(766),...
    '(YEL041W or YJR049C or YPL188W)');
model=changeGeneAssociation(model,model.rxns(771),...
    '(YEL041W or YJR049C or YPL188W)');

%I'm not sure what to do with rxn 684 - no ref, mis-annotated as PUT2. For
%now, set to zero (it has no flux in WT anyway) If remove, may also remove
%rxn 2076 and 685. In fact, met 1501 and 1502 only participate in these, so
%they should be removed and 1501 should just be proline (see rxn 687 in
%cytosol, below).
model=changeRxnBounds(model,model.rxns([684 685 2076]),0,'l');
model=changeRxnBounds(model,model.rxns([684 685 2076]),0,'u');

%rxn 939 should use FAD as the cofactor, not NAD. ref: PMID: 387737
model.S(1200,939)=0; model.S(1205,939)=0; model.S(799,939)=0;
model.S(688,939)=-1; model.S(690,939)=1;

%KEGG has rxn 1593, but not in yeast.
%http://www.genome.jp/dbget-bin/www_bget?rn:R03337
model=changeRxnBounds(model,model.rxns(1593),0,'l');
model=changeRxnBounds(model,model.rxns(1593),0,'u');

%I haven't been able to confim that there is a yeast NADP phosphatase.
model=changeRxnBounds(model,model.rxns(1958),0,'l');
model=changeRxnBounds(model,model.rxns(1958),0,'u');

%I'm not sure what to do about rxns 167 170 180 183 187. The biochemistry
%isn't in SGD, though the description seems okay. For now, fix annotation,
%leave them and check to see if they have fluxes in WT or kennedy KO
%simulations.
%YGL256W is model.genes(339).
model=changeGeneAssociation(model,model.rxns(167), '(YMR083W or YGL256W)');
model=changeGeneAssociation(model,model.rxns(170), '(YMR083W or YGL256W)');
model=changeGeneAssociation(model,model.rxns(180), '(YMR083W or YGL256W)');
model=changeGeneAssociation(model,model.rxns(183), '(YMR083W or YGL256W)');
model=changeGeneAssociation(model,model.rxns(187), '(YMR083W or YGL256W)');
```

```
%%
%mitochondrial NADP rxns - started 3/5/12

%not sure of rxn 175 or 178 reversibility (or 178's biochem)

%NADP not in SGD, only NAD;
%http://pathway.yeastgenome.org/YEAST/NEW-IMAGE?type=ENZYME-IN-
PATHWAY&object=YHR037W-MONOMER&detail-level=2
model=changeRxnBounds(model,model.rxns(325),0,'l');
model=changeRxnBounds(model,model.rxns(325),0,'u');

%rxn 345 reversible in KEGG, not SGD; left irreversible for now.

%leave rxns 396 423:431 to fatty acid metabolism check

%rxn 719 - SGD only has NAD version, and there's no ref.
model=changeRxnBounds(model,model.rxns(719),0,'l');
model=changeRxnBounds(model,model.rxns(719),0,'u');

%rxn 759 is reversible in KEGG, not SGD. leaving irreversible.

%rxn 1038 seems outside scope of YN - protein modification? Additionally,
%its biochem doesn't agree with SGD. But, it has a flux in each Kennedy KO,
%so leave alone?
%model=changeRxnBounds(model,model.rxns(1038),0,'l');
%model=changeRxnBounds(model,model.rxns(1038),0,'u');

%%
%cytoplasmic reactions involving NAD - see
%http://www.ncbi.nlm.nih.gov/pubmed/14977171 for some test cases.

%rxn 61 biochem doesn't agree with SGD, but combined with rxn 29, has same
%effect. So, don't change it for now.

%per SGD, "In S. cerevisiae, there are five genes that encode alcohol
%dehydrogenases involved in  ethanol metabolism, ADH1 to ADH5. Four of
%these enzymes, Adh1p, Adh3p, Adh4p, and Adh5p,  reduce acetaldehyde to
%ethanol during glucose fermentation, while Adh2p catalyzes the reverse
%reaction of oxidizing ethanol to acetaldehyde." Combined with localization
%data, rxn 163 should be split into two rxns, and annotation should change.
%'YBR145W' 'YDL168W' 'YGL256W' 'YMR303C' and 'YOL086C' are model.genes([66
%145 339 718 785]).

model=changeRxnBounds(model,model.rxns(163),0,'l');
model=changeRxnBounds(model,model.rxns(163),1000,'u');
model=changeGeneAssociation(model,model.rxns(163),'YMR303C');

model=addReaction(model,'NRXN2',{'s_0359', 's_0794', 's_1203', ...
    's_0680', 's_1198'}, [-1 -1 -1 1 1], false, 0, 1000, 0, '',...
    '(YOL086C or YBR145W)');

%Like rxns 167 170 180 183 187, I'm not sure what to do about 166 169 179
%182 186. The biochemistry isn't in SGD, though the description seems okay.
%For now, fix annotation, leave them and check to see if they have fluxes
%in WT or kennedy KO simulations.
model=changeGeneAssociation(model,model.rxns(166),...
    '(YBR145W or YDL168W or YOL086C)');
```

```
model=changeGeneAssociation(model,model.rxns(169),...
    '(YBR145W or YDL168W or YOL086C)');
model=changeGeneAssociation(model,model.rxns(179),...
    '(YBR145W or YDL168W or YOL086C)');
model=changeGeneAssociation(model,model.rxns(182),...
    '(YBR145W or YDL168W or YOL086C)');
model=changeGeneAssociation(model,model.rxns(186),...
    '(YBR145W or YDL168W or YOL086C)');

%so are model.genes([66 145 339 718 785]) all still used? checking
%model.rxnGeneMat(:,[66 145 339 718 785]) says yes.

%rxns 172 is and 185 may be a functions of ALD2/ALD3, but that biochem
%isn't in SGD, and surprisingly, the acetaldehyde -> acetate rxn
%(acetaldehyde dehydrogenase) doesn't seem to be in the model (which would
%explain part of the trouble with growth on ethanol!
%YMR170C is model.genes(694)
model=addReaction(model,'NRXN3',{'s_0359', 's_0803', 's_1198', ...
    's_0362', 's_1203', 's_0794'}, [-1 -1 -1 1 1 1], false, 0, 1000, 0,...
    '','YMR170C');

%rxn 232 is annotated with genes that don't have that function in SGD, and
%are not in the cytosol compartment (a mix of ER and mitochondria). *not
%changed now b/c concerened that ER compartmentalisation isn't right, and
%may be needed for growth.
model=changeRxnBounds(model,model.rxns(232),0,'l');
model=changeRxnBounds(model,model.rxns(232),0,'u');

%rxn 235 should be in ER *note: this has a flux in WT, and blocking the
%reaction stops aerobic growth*
%model=changeRxnBounds(model,model.rxns(235),0,'l');
%model=changeRxnBounds(model,model.rxns(235),0,'u');

%rxn 316 appears miscompartmentalized, annotation doesn't agree with SGD.
%*note: this has a flux in WT, but blocking the reaction doesn't stop
%aerobic growth* but, 316 has flux in Kennedy KOs
model=changeRxnBounds(model,model.rxns(316),0,'l');
model=changeRxnBounds(model,model.rxns(316),0,'u');

%Basically, something seems odd with ergosterol metabolism in Y5.

%per SGD biochem, rxn 443 includes a spontaneous rxn - so likely not
%reversible
model=changeRxnBounds(model,model.rxns(443),0,'l');
model=changeRxnBounds(model,model.rxns(443),1000,'u');

%rxn 445 can be catalyzed by FDH2 in CEN.PK 113-7D strain (see pmid:
%11921099)
model=changeGeneAssociation(model,model.rxns(445),...
    '(YOR388C or (YPL275W AND YPL276W))');

%rxn 470 is annotated as mitochondrial and is reversible in SGD. I made it
%reversible, but didn't move it to the mitochondria b/c I'm not sure how to
%reconsile that rxn with rxn 472 (below)
model=changeRxnBounds(model,model.rxns(470),-1000,'l');
model=changeRxnBounds(model,model.rxns(470),1000,'u');
%model.S([180 419 794 803 991 1198 1203],470)=0;
```

```
%model.S([182 421 799 1205],470)=1; %2-oxoglutarate, ammonium, H+, NADH
%model.S([807 993 1200],470)=-1; %H20, L-glutamate, NAD

%rxn 472 is annotated as mitochondrial in SGD - but y5 doesn't have
%mitochodrial L-glutamine, so I have to add that. Per SGD, it's synthesized
%by LGN1/YPR035W from glutamate. There is a glutamate transporter
%(AGC1/YPR021C, rxn 1193) in yeast (but SGD has LGN1 is in cytosol...)
%model.S([180 794 991 999 1198 1203],472)=0;
%model.S([182 799],472)=-1; %2-oxoglutarate, H+, L-glutamine (MISSING), NADH
%model.S(993,472)=2; %L-glutamate
%model.S(1200,472)=1; %NAD

%rxn 474 - SGD uses NADP as cofactor, not NAD (which is rxn 473).
model=changeRxnBounds(model,model.rxns(474),0,'l');
model=changeRxnBounds(model,model.rxns(474),0,'u');

%rxn 677 - SGD uses NADP as a cofactor, not NAD (which is rxn 678).
%Additionally, it doesn't involve ATP/ADP in this rxn (corrected below in
%NADP section)
model=changeRxnBounds(model,model.rxns(677),0,'l');
model=changeRxnBounds(model,model.rxns(677),0,'u');

%rxn 686 - SGD uses NADP as cofactor, not NAD (which is rxn 687).
model=changeRxnBounds(model,model.rxns(686),0,'l');
model=changeRxnBounds(model,model.rxns(686),0,'u');

%rxn 696 - no gene annotation; looks like maybe GRE2/YOL151W, which SGD has
%as NADP-dependent, and uses D- (R-)lactaldehyde as a substrate, not
%(S)-lactaldehyde. (R-) lactaldehyde isn't in the model. (S)-lactaldehyde
%is only in 2 rxns... I think there's something wrong here, but am not
%sure how to correct it...

%rxn 714 should be reversible
%*NOTE: this change increases max growth rate from .09 to .1, but now don't
%have fluxes through all rxns of TCA cycle - hopefully this fixes itself
%with further corrections*
%SGD says its regulated by glucose repression, and notes "unfavorable
%equilibrium for formation of oxaloacetate from malate" - but I'd interpret
%that to mean the reaction prefers to go in the to malate direction
%(negative flux)
%see also: http://dx.doi.org/10.1074/jbc.M213231200
%model.lb(714)=-1000; model.ub(714)=1000;
model=changeRxnBounds(model,model.rxns(714),0,'l');%a modeling constraint.
model=changeRxnBounds(model,model.rxns(714),1000,'u');

%rxn 782 - SGD has YGR010 as nuclear but YLR328W as cytoplasm - annotation
%change.
model=changeGeneAssociation(model,model.rxns(782),'YLR328W');

%rxn 1569 - no gene, no ref; KEGG rxn R01000, which KEGG doesn't include in
%Saccharomyces.
model=changeRxnBounds(model,model.rxns(1569),0,'l');
model=changeRxnBounds(model,model.rxns(1569),0,'u');

%rxn 1952 - no gene, no ref; KEGG rxn R05050, which KEGG annotates as
%YER073W. SGD says YER073W is mitochondrial, uses NADP. searching for
%acetamidobutanoate doesn't find any hits in SGD.
```

```
model=changeRxnBounds(model,model.rxns(1952),0,'l');
model=changeRxnBounds(model,model.rxns(1952),0,'u');

%rxn 1953 - no gene, no ref; KEGG rxn R00102, which KEGG doesn't include in
%Saccharomyces.
model=changeRxnBounds(model,model.rxns(1953),0,'l');
model=changeRxnBounds(model,model.rxns(1953),0,'u');

%rxn 1955 - no gene, no ref; a modeling rxn (when unconstrained, has
%negative flux in WT, and fewer TCA fluxes. When constrained, more TCA
%fluxes.
model=changeRxnBounds(model,model.rxns(1955),0,'l');
model=changeRxnBounds(model,model.rxns(1955),0,'u');

%rxn 1957 - no gene, no ref; I can't find evidence that yeast has a NADP
%phosphatase. - has activity in Kennedy KO1 KO2 KO3 KO4.
model=changeRxnBounds(model,model.rxns(1957),0,'l');
model=changeRxnBounds(model,model.rxns(1957),0,'u');

%rxn 2019 - no gene, no ref; KEGG rxn R01728, which KEGG doesn't have in
%yeast (though it does have a NADP prephenate dehydrogenase annotated as
%YBR166C, which is rxn 938)
model=changeRxnBounds(model,model.rxns(2019),0,'l');
model=changeRxnBounds(model,model.rxns(2019),0,'u');

%%
%cytoplasmic reactions involving NADP

%rxn 15 - SGD doesn't have NADP involved. KEGG rxn R09376 does. Lit appears
%to support the use of NADP (and maybe NAD), but I don't see lots of
%yeast-specific refs.

%rxn 80 - SGD has MET 13 in mitochondria (MET 12 seems less well
%characterized), but downstream MET6 is in cytoplasm, so don't relocate for
%now.

%rxn 164 - I'm not sure that the YHR104W annotation is correct.

%rxn 233 - annotation incorrect; per SGD, YHR042W is associated with Erg11,
%not erg5. However, Erg11 doesn't have this activity. So should just be
%YMR015C.
model=changeGeneAssociation(model,model.rxns(233),'YMR015C');

%rxn 234 - SGD localizes to ER. I didn't change it. Also - what kind of met
%are " zymosterol intermediate 1c" or "zymosterol intermediate 2"?

%rxn 324 - SGD doesn't include glucose as a substrate, and there's no ref
%in Y5. Remove rxn.
model=changeRxnBounds(model,model.rxns(324),0,'l');
model=changeRxnBounds(model,model.rxns(324),0,'u');

%I didn't look at fatty acid synthase rxns closely (rxns 385:395 397 398
%432:435)

%I didn't check annotation of rxn 481 too closely, but rxn agrees with SGD.

%rxn 487 is putative in SGD (but has no flux in WT, so leave as is)
```

```
%rxn 678 - SGD doesn't involve ATP/ADP in this rxn, and the lit ref isn't
%great for the stoichiometry. SGD lists water as a product, not
%diphosphate.
model.S(434,678)=0; model.S(423,678)=0; model.S(633,678)=0;
model.S(803,678)=1;

%rxn 687 - SGD has product of proline, not trans-4-hydroxy-L-proline. So,
%change this, and then don't need mito transport or met 1501. note: it
%doesn't appear that Y5 can synthesize proline prior to this change?
%my note is wrong - see rxn 956. Still not sure what rxn 687 is.
model.S(1501,687)=0; model.S(1035,687)=1;

%rxn 688 - not high confidence, but unchanged

%rxn 690 "The high Km for NADP+ may suggest that the equilibrium position
%of the YdfG and YMR226c reactions markedly favors the reduction of the
%l-2-amino-3-ketobutyrate." - PMID: 12535615; so, make irreversible.
model=changeRxnBounds(model,model.rxns(690),0,'l');
model=changeRxnBounds(model,model.rxns(690),1000,'u');

%rxn 763 - misannotated, perhaps should be DIT1/DIT2 instead of DIT2/NCP1?
%(I have low confidince in this, and suggest it due to SGD phrasing DIT2
%"involved in the production of N,N-bisformyl dityrosine" and DIT1
%"involved in the production of a soluble LL-dityrosine-containing
%precursor of the spore wall"
%this means adding DIT1 to the model.
%DIT1 is YDR403W ; DIT2 is YDR402C, model.genes(221); and NCP1 is YHR042W,
%model.genes(409)
model=changeGeneAssociation(model,model.rxns(763),'(YDR403W and YDR402C)');
model=changeGeneAssociation(model,model.rxns(919),'YDR402C');

%rxn 836 - should be in ER? SGD does not have NADP as a cofactor. NOTE:
%"OLE1 encodes the sole S. cerevisiae Delta-9 fatty acid desaturase, an ER
%membrane protein required for the production of monounsaturated fatty
%acids. Because these fatty acids are critical components of cell
%membranes, the OLE1 gene is essential unless the media are supplemented
%with unsaturated fatty acids." - so this (and rxn 1012) are critically
%important rxns. (I'm not sure if media has unsaturated FAs at the moment)
% At the moment, the only specifically named palmitoleoyl-CoA in model is
% in ER - met 1299, palmitoleoyl-CoA [endoplasmic reticulum]
%something is odd with rxn 836's products. For now, I'm just removing the
%NADP/H from the rxn.
model.S(794,836)=-2; %(2 H+, not 1)
model.S([1207 1212],836)=0;

%rxn 956 - how different from 687, above?

%rxn 988 - not reversible that I've seen. This prompts the idea: rather
%than assume rxns are reversible, what if we assume they're irreversible,
%except where documented otherwise?
model=changeRxnBounds(model,model.rxns(988),0,'l');
model=changeRxnBounds(model,model.rxns(988),1000,'u');

%rxn 997 - SGD has NAD as the cofactor.
model.S([1207 1212],997)=0; model.S(1203,997)=1; model.S(1198,997)=-1;
```

```
%rxn 1012 - SGD doesn't have NADP as a cofactor. (like rxn 836, above)
model.S([1207 1212],1212)=0;

%rxn 1026 - per SGD, should be backwards only, annotation should be AND,
%not OR
model=changeRxnBounds(model,model.rxns(1026),-1000,'l');
model=changeRxnBounds(model,model.rxns(1026),0,'u');
model=changeGeneAssociation(model,model.rxns(1026),...
    '(YFR030W and YJR137C)');

%rxn 1037 - strange rxn, seems pretty general in SGD. unchanged for now.

%rxn 1948 - it's not clear that this activity is in Yeast - no gene in
%model.
model=changeRxnBounds(model,model.rxns(1948),0,'l');
model=changeRxnBounds(model,model.rxns(1948),0,'u');

%rxn 1959 - I'm pretty sure NADP can't cross the cell membrane.
model=changeRxnBounds(model,model.rxns(1959),0,'l');
model=changeRxnBounds(model,model.rxns(1959),0,'u');

%rxn 1961 - I think NADP can't cross into the ER, either... This might
%be needed for some compartmentalization (it has a flux in WT sln),
%blocking blocks growth.
%model=changeRxnBounds(model,model.rxns(1961),0,'l');
%model=changeRxnBounds(model,model.rxns(1961),0,'u');

%rxn 1984 - the mechanism of the second unsaturation is unknown, but
%following the example of rxns 836 and 1012, I'll remove NADP from this
%rxn.
model.S([1207 1212],1984)=0;

%%
%other changes - random things I come across while validating this:

%rxn 1109 should be reversible - per SGD, "However, under certain
%conditions, such as during exponential growth on glucose under aerobic
%conditions, they act in the opposite direction, importing ATP into
%mitochondria"
model=changeRxnBounds(model,model.rxns(1109),-1000,'l');
model=changeRxnBounds(model,model.rxns(1109),1000,'u');

%rxn 444 citation is for a paper describing FDH1/FDH2, which doesn't use
%ubiquinol-6, but uses NAD - rxn 445. rxn 445 annotation changed above.
%block rxn 444.
model=changeRxnBounds(model,model.rxns(444),0,'l');
model=changeRxnBounds(model,model.rxns(444),0,'u');

%rxn 1227 annotation should include MPH2 and MPH3
model=changeGeneAssociation(model,model.rxns(1227),...
    '(YBR298C or YDL247W or YGR289C or YDL247W or YJR160C)');

%rxn 194 - per SGD, YBR001C "does not appear to encode significant
%trehalase activity, or be involved in trehalose catabolism"
model=changeGeneAssociation(model,model.rxns(194),...
    'YDR001C');
```

```
%rxn 80 - name needs a space
model.rxnNames(80)={'5,10-methylenetetrahydrofolate reductase (NADPH)'};

%rxn 906 - per SGD, this should be "a phosphopantothenoylcysteine
%decarboxylase (PPCDC; Cab3p, Sis2p, Vhs3p) complex" (cab3 is essential,
%sis2 is not essential, vhs3 is not essential). I suspect that sis2 and
%vhs3 were left out b/c they're inessential - I think they both encode
%negative regulatory subunits, so if they're missing, the enzyme function
%couldn't be regulated.
model=changeGeneAssociation(model,model.rxns(906),...
    '(YKL088W or (YKL088W and YKR072C) or (YKL088W and YOR054C) or (YKL088W
and YKR072C and YOR054C))');

%I'm not confident about rxn 1031 - esp the RMA1 annotation

%% Pathway coverage changes
% modifications based on comparison to yeastCyc pathway classification of
% genes

%rxn 762 currently misannotated as BNA3. per SGD, the activity described is
%catalyzed by BNA7: http://pathway.yeastgenome.org/YEAST/new-
image?type=PATHWAY&object=NADSYN-PWY&detail-level=2
%(per SGD, BNA3 catalyzes formation of kynurenic acid from kynurenine)
model=changeGeneAssociation(model,model.rxns(762),...
    'YDR428C');

%per SGD, rxn 2027 catalyzed by BUD16. currently not annotated. Lit ref is
%not a smoking gun, so perhaps lower confidence annotation.
model=changeGeneAssociation(model,model.rxns(2027),...
    'YEL029C');

%per SGD, rxn 905 is catalyzed by the phosphopantothenoylcysteine
%decarboxylase (PPCDC; Cab3p, Sis2p, Vhs3p) complex. Currently annotated
%with only CAB3 (note: check which might be essential - sis2 is not)
model=changeGeneAssociation(model,model.rxns(905),...
    '(YKL088W  or (YKR072C and YKL088W) or (YOR054C and YKL088W) or (YOR054C
and YKR072C and YKL088W))');

%sis2 and vhs3 also form a complex with ppz1 (which isn't in Y5).

%%
%check rxns of glycolosys, gluconeogenesis, TCA cycle, gloxylate cycle, and
%Pentose Phosphate pathway. At this point, there are changes between the
%pre-modified model and the post-modified model fluxes.

%An essential question is: what SHOULD the fluxes be?

%%
%glycolysis - the following rxns all have positive fluxes
%post-modifications (not all do prior to changes)

%pathway
%rxn 1165 -'glucose transport' - exc to cyto ->
%rxn 534 - 'hexokinase (D-glucose:ATP)' - cyto ->
%rxn 467 'glucose-6-phosphate isomerase' - cyto <->
%rxn 885 'phosphofructokinase' - cyto ->
%rxn 450 'fructose-bisphosphate aldolase' - cyto <->
```

```
%rxn 1053 'triose-phosphate isomerase' - cyto <->
%rxn 486 'glyceraldehyde-3-phosphate dehydrogenase' - cyto <->
%rxn 891 'phosphoglycerate kinase' - cyto <->
%rxn 892 'phosphoglycerate mutase' - cyto <->
%rxn 366 'enolase' - cyto <->
%rxn 961 'pyruvate kinase' - cyto ->
%rxn 958 'pyruvate decarboxylase' - cyto ->

%notes:
%g-6-p (from r534) participates in 7 rxns. 5 of them have fluxes
%f-6-p (from r467) participates in 11 rxns. 5 of them have fluxes
%f-1,6-bp (from r885) participates in 4 rxns. 2 have fluxes (glycolysis)
%DHAP (from r450) participates in 8. 3 have fluxes (glycolysis + r491)
%GADP (from r1053 and 450) in 7. all have fluxes.
%1,3BPG (from r486) in 3. 2 have fluxes (glycolysis)
%3PG (from r891) in 3. all have fluxes (glycolysis + TCA)
%2PG (from r892) in 2. both have fluxes

%PEP (from r366) in 6. 4 have fluxes. rxn 65 is an inferred gene function,
%but no FBA sln if I block it. but, YDR127W is not an essential gene.
%merits some more investiagation. (ah - not essential in complex media!)

%pyruvate (from r961) in 18. 5 have fluxes (203 957 958 961 2032). I think
%if r958 has a flux, ethanol should be produced...(maybe regulated, though,
%and I'm not very sure of this). As it is, acetaldehyde is transported to
%the mitochondria via rxn 1631, where it is converted to acetate via rxn
%175. The acetate goes to the cyto (rxn 1635), where there are 3 other rxns
%using acetate. FBA works if I block it. This also helps with TCA fluxes (8
%of 11 have fluxes instead of 6 of 11). Or, rxn 958 may be part of TCA or
%glyoxylate cycles?

%model.ub(958)=0; %possible regulatory constraint - I'm not sure.

%so, glycolysis fluxes seems mostly good at the moment. (pending issue in
%last paragraph

%%
%gluconeogenesis - post-modifications, rxns 883 and 449 have 0 flux (883
%has a flux prior to changes)

%pathway
%rxn 957 'pyruvate carboxylase'
%rxn 883 'phosphoenolpyruvate carboxykinase'
%rxn 366 'enolase' - cyto <->
%rxn 892 'phosphoglycerate mutase' - cyto <->
%rxn 891 'phosphoglycerate kinase' - cyto <->
%rxn 486 'glyceraldehyde-3-phosphate dehydrogenase' - cyto <->
%rxn 450 'fructose-bisphosphate aldolase' - cyto <->
%rxn 449 'fructose-bisphosphatase' - cyto ->
%rxn 467 'glucose-6-phosphate isomerase' - cyto <->

%notes
%so, gluconeogenesis fluxes seem good at the moment.

%%
%Pentose Phosphate pathway - post modifications, these the first 3 rxns on
%the following list don't have fluxes, and the last 4 are negative
```

```
%pathway
%rxn 466 'glucose 6-phosphate dehydrogenase'
%rxn 91 '6-phosphogluconolactonase'
%rxn 888 'phosphogluconate dehydrogenase'
%rxn 981 'ribose-5-phosphate isomerase'
%rxn 983 'ribulose 5-phosphate 3-epimerase'
%rxn 1048 'transketolase 1'
%rxn 1047 'transaldolase'
%rxn 1049 'transketolase 2'

%notes
%D-6-P-G-D-L (met 335 from rxn 466) is in 3 rxns. none have fluxes.
%6-P-D-G (met 340 from rxn 91) is in 2 rxns. niether has a flux.
%rxn 888 produced NADPH - should have a flux! (question: where is NADPH
%currently produced?)
%R-5-P (met 577 from rxn 888) is in 4 rxns. produced in rxn 983, consumed
%in 981 and 38. So where is the xylulose 5-p coming from?
%X-5-P (met 581 in rxn 983) in 5 rxns. 3 have - fluxes (all in PPP)

%I think that PPP can have these fluxes, in the absence of NADPH demand.
%This might be fixed with changed biomass def?
%What if add NADP sink/demand (I don't see the differenceat the moment)
%reaction?
%model=addSinkReactions(model,{'s_1212'},1,1000); %breaks FBA.

%rxn 1048 has negative flux. What if restrict direction?
model=changeRxnBounds(model,model.rxns(1048),0,'l'); %MODELING CONSTRAINT
%same for rxn 1049
model=changeRxnBounds(model,model.rxns(1049),0,'l'); %MODELING CONSTRAINT

%now first 6 rxns have fluxes; glycolysis looks good. TCA still needs
%attention.

% per Yeasts 2ed v3 p 223 "it should be emphasized that, contrary to an
% extended belief, this pathway cannot function in yeast as a cycle
% oxidizing glucose to 6 CO2 and 12 NADPH during hexose metabolism. This
% function would imply the operation of fructose-1,6-bipshosphatase, an
% enzyme repressed by hexoses in Sacch. cerevisiae.

%F-1,6-BP is YLR377C, catalyzing rxn 449, which is currently constrained to
%the forward direction.

%see http://dx.doi.org/10.1006/mben.1998.0110 - expect 20-50% of glucose is
%oxidized via PPP in aerobic conditions. So I think there should be fluxes
%through each rxn of the PPP.

%see also http://dx.doi.org/10.1099/00221287-129-4-953 - a very interesting
%ref!

%so, at this point, I've got some directional constraints on transketolase
%rxns (lit support?). With those, I get glycolysis and PPP fluxes I like,
%but not TCA yet.

%%
%TCA cycle (in mito) - post modifications, these the first 5 and final rxns
%on the following list have fluxes, and the others do not. If rxn 714 is
```

```
%irreversible, all TCA rxns have positive fluxes. I'll examine with
%irreversible rxn 714 first.

%pathway
% rxn 2032 'pyruvate transport'
% rxn 960 'pyruvate dehydrogenase'
% rxn 300 'citrate synthase'
% rxn 302 'citrate to cis-aconitate(3-)'
% rxn 280 'cis-aconitate(3-) to isocitrate'
% rxn 658 'isocitrate dehydrogenase (NAD+)'
% rxn 831 'oxoglutarate dehydrogenase (lipoamide)'
% rxn 830 'oxoglutarate dehydrogenase (dihydrolipoamide S-
succinyltransferase)'
% rxn 1021 'succinate-CoA ligase (ADP-forming)'
% rxn 1020 'succinate dehydrogenase (ubiquinone-6)'
% rxn 451 'fumarase'
% rxn 713 'malate dehydrogenase'

%notes
% pyruvate (met 1401) from r2032 in 11 rxns. 5 have fluxes consuming.
% acc-CoA (met 376) from r960 in 13 rxns. 3 have fluxes
% citrate (met 524) from r300 in 6 rxns. 4 have fluxes(vs 2 w/rev 714)

% citrate is going cyto in r1125 (importing malate), and returning in
% r1127(exporting isocitrate)

% cis-aconitate (met 517) from r302 in 3 rxns. 2 have fluxes
% isocitrate (met 941) from r280 in 4 rxns. 3 have fluxes, seem ok

% 2-oxoglutarate (met 182) from r658 and 660 in 14 rxns. 8 have fluxes. In
% reversible r714 case, r660 has flux, r831 does not.
% produced in r 118 (k), 217 (k), 674 (k), 1087 (k), 1098 (k, maybe glu
% repressed), consumed in rxn 1836

%rxn 1836 isn't annotated with a gene - looks like LYS20.
model=changeGeneAssociation(model,model.rxns(183),'YDL182W');

% succinyl-CoA (met 1464) from r830 in 3 rxns. 2 have fluxes
% succinate (met 1460) from r1021 in 7 rxns. 3 have fluxes.
% fumarate (met 727) from r1020 in 5 rxns. 2 have fluxes.

%rxn 1020 should use FAD? see http://dx.doi.org/10.1074/jbc.271.8.4055
%only required in 1 direction. see 10.1111/j.1432-1033.1994.tb18949.x

%"The covalent attachment of FAD is therefore necessary for succinate
%oxidation but is dispensable for both fumarate reduction and for the
%import and assembly of the flavoprotein subunit."

%so, maybe split into 3 rxns? (I haven't done this at the moment)
%succinate + ubiquinone-6 + FAD -> fumarate + ubiquinol-6 + FADH
%succinate + ubiquinone-6 + FAD <- fumarate + ubiquinol-6 + FADH
%succinate + ubiquinone-6 <- fumarate + ubiquinol-6

% malate (met 68) from r451 in 6 rxns. 4 have fluxes (including 2 exchange
% rxns that seem ok)
% r713 forms oxaloacetate, which participates in r300, completing cycle.
```

```
%%
%glyoxylate cycle (mostly in cyto). Post-changes (irreversible rxn 714), x
%of the following N rxns have fluxes.

%"In the presence of 1% glucose, the synthesis of all enzymes, except
%fumarase, was repressed. However, no indication for a specific regulation
%mechanism for the entire cycle could be found. Studies on the localization
%of the glyoxylate cycle enzymes in the yeast cell revealed that the key
%enzymes, isocitrate lyase and malate synthase, are located in the
%cytoplasm, whereas succinate dehydrogense was found only in the
%mitochondrial fraction. Activites of all other enzymes are found in the
%cytoplasm as well as in the mitochondira. In anaerobically grown cells, no
%mitochondria could be detected. Succinate dehydrogenase, isocritrate
%lyase, and malate synthase were absent. However, appreciable activities of
%citrate synthase, aconitase, fumarase, and malate dehydrogenase were found
%in the Cytoplasm." - http://dx.doi.org/10.1111/j.1432-1033.1969.tb00658.x

% pathway (starting with Malate)

% rxn 713 'malate dehydrogenase' - malate to oxaloacetate(mito)
% rxn 714 'malate dehydrogenase, cytoplasmic ' - 0 flux
% rxn 300 'citrate synthase' - mitochondrial generation of citrate
% MISSING RXN - 'citrate synthase' - cyto? - YCR005C?
% rxn 302 'citrate to cis-aconitate(3-)' - mitochondrial
% rxn 303 'citrate to cis-aconitate(3-), cytoplasmic'
% rxn 280 'cis-aconitate(3-) to isocitrate' - mitochondrial - YJL200C?
% MISSING RXN? 'cis-aconitate(3-) to isocitrate' - cyto? - YLR304C?
% rxn 662 'isocitrate lyase' - isocitrate to glyoxylate and succinate
% rxn 716 'malate synthase' - glyoxylate to malate (in cyto)
% rxn 1020 'succinate dehydrogenase (ubiquinone-6)' - succ to fum (mito)
% rxn 452 'fumarase' - fumarate to malate (cyto)
% rxn 451 'fumarase' - fumarate to malate (mito)

%tansport rxns
% rxn 1225 'malate transport' - to mitochondria
% rxn 1238 'oxaloacetate transport' - to mitochondria
% rxn 1127 'citrate transport' - to cyto
% rxn 1264 'succinate-fumarate transport' - succinate to mito

%notes
%again, 1020 should perhaps have FAD cofactor?

%rxn 280 may be misannotated as YLR304C, which may catalyze cytosol rxn.
%280 could be annotated YJL200C. cytosol cis-Aconitate to Isocitrate may be
%missing.

%missing a cytosol citrate synthase?

%so it seems that enough have fluxes for the cycle to function.

%% output
rxnNames = {
    % glycolysis
    'GLYCOLYSIS'
    'glucose transport'
    'hexokinase (D-glucose:ATP)'
    'glucose-6-phosphate isomerase'
```

```
'phosphofructokinase'
'fructose-bisphosphate aldolase'
'triose-phosphate isomerase'
'glyceraldehyde-3-phosphate dehydrogenase'
'phosphoglycerate kinase'
'phosphoglycerate mutase'
'enolase'
'pyruvate kinase'
'pyruvate decarboxylase'
''
%ethanol production
%'alcohol dehydrogenase, reverse rxn (acetaldehyde -> ethanol)'
%should be 'alcohol dehydrogenase (ethanol)'
'ETHANOL PRODUCTION'
'alcohol dehydrogenase (ethanol)'
'ethanol transport'
''
% gluconeogenesis
'GLUCONEOGENESIS'
'pyruvate carboxylase'
'phosphoenolpyruvate carboxykinase'
'enolase'
'phosphoglycerate mutase'
'phosphoglycerate kinase'
'glyceraldehyde-3-phosphate dehydrogenase'
'fructose-bisphosphate aldolase'
'fructose-bisphosphatase'
'glucose-6-phosphate isomerase'
''
% PPP
'PENTOSE PHOSPHATE PATHWAY'
'glucose 6-phosphate dehydrogenase'
'6-phosphogluconolactonase'
'phosphogluconate dehydrogenase'
'ribose-5-phosphate isomerase'
'ribulose 5-phosphate 3-epimerase'
'transketolase 1'
'transaldolase'
'transketolase 2'
''
% TCA
'TCA CYCLE'
%'pyruvate transport' %commented out b/c multiple rxns have this name
'pyruvate dehydrogenase'
'citrate synthase'
'citrate to cis-aconitate(3-)'
'cis-aconitate(3-) to isocitrate'
'isocitrate dehydrogenase (NAD+)'
'oxoglutarate dehydrogenase (lipoamide)'
'oxoglutarate dehydrogenase (dihydrolipoamide S-succinyltransferase)'
'succinate-CoA ligase (ADP-forming)'
'succinate dehydrogenase (ubiquinone-6)'
'fumarase'
'malate dehydrogenase'
''
%glyoxylate cycle
'GLYOXYLATE CYCLE'
```

```
    'malate dehydrogenase'
    'malate dehydrogenase, cytoplasmic'
    'citrate synthase'
    'citrate synthase'
    'citrate to cis-aconitate(3-)'
    'citrate to cis-aconitate(3-), cytoplasmic'
    'cis-aconitate(3-) to isocitrate'
    'isocitrate lyase'
%    'malate synthase' %commented out b/c multiple rxns have this name
    'succinate dehydrogenase (ubiquinone-6)'
    'fumarase'
    'malate transport'
%    'oxaloacetate transport' %commented out b/c multiple rxns have this name
%    'citrate transport' %commented out b/c multiple rxns have this name
    'succinate-fumarate transport'
    ''
    % other
    };

fprintf('\nAerobic growth prior to
changes\nrate:\t%.2f\n\n',unchangedsolution.f);

for k = 1:length(rxnNames)
   ind = strcmp(rxnNames{k},model.rxnNames);
   if sum(ind)>1, disp(rxnNames{k}); end
   fprintf('%.2f\t%s\n',unchangedsolution.x(ind),rxnNames{k});
end

FBAsolution = optimizeCbModel(model,[],'one');
fprintf('Aerobic growth after changes\nrate:\t%.2f\n\n',FBAsolution.f);

for k = 1:length(rxnNames)
   ind = strcmp(rxnNames{k},model.rxnNames);
   if sum(ind)>1, disp(rxnNames{k}); end
   fprintf('%.2f\t%s\n',FBAsolution.x(ind),rxnNames{k});
end

% ana=model;
%
% ind = strcmp('oxygen exchange',ana.rxnNames); ana.lb(ind) = 0;
%
% ind = ismember(ana.rxnNames,{...
%     'lipid pseudoreaction [no 14-demethyllanosterol, no ergosta-
5,7,22,24(28)-tetraen-3beta-ol]'
%     'ergosterol exchange'
%     'lanosterol exchange'
%     'zymosterol exchange'
%     'phosphatidate exchange'
%     });
% ana.lb(ind) = -Inf;
% ana.ub(ind) = Inf;
%
% ind = strcmp('lipid pseudoreaction',ana.rxnNames); ana.ub(ind) = 0;
%
% ana_FBAsolution = optimizeCbModel(ana,[],'one');
% fprintf('anaerobic growth\nrate:\t%.2f\n\n',ana_FBAsolution.f);
%
```

```
% for k = 1:length(rxnNames)
%    ind = find(strcmp(rxnNames{k},ana.rxnNames));
%    fprintf('%.2f\t%s\n',ana_FBAsolution.x(ind),rxnNames{k});
% end

save('new_model','model');
```

Supplementary File 5.4.1: Y5_3_three_knockout1.m script

```
% FILE NAME: Y5_three_knockout1
%
% DATE CREATED: 9 Feb 2012
%
% PROGRAMMER: B. Heavner
% Department of Biological
% and Environmental Engineering
% Cornell University
% Ithaca, NY 14853
%
% LAST REVISED: 11 July 2012
%
% REVISED BY: B. Heavner
%
% REVISIONS:
%
% 11 July 2012 - BH - cleaned up for publication
% 22 May 2012 - BH - change to use Y5.30
% 13 May 2012 - BH - change media to have 18:1 and 16:1
% 11 May 2012 - BH - fix strain background misunderstanding
% 5 April 2012 - BH - get it working with modified Y5.
%
% PURPOSE:
% To attempt to reproduce the model results reported by Kennedy et al. with
% the Yeast 5.30 model instead of the iND750 model.
%
% In this script, I apply constraints to Y5 to duplicate the Kennedy et al.
% media and mutants to see if I still get a predicted formic acid secretion
% with the mutants reported in Kennedy et al. Table 1.
%
% REFERENCE:
% Kennedy CJ, Boyle PM, Waks Z, Silver PA. "Systems-Level Engineering of
% Nonfermentative Metabolism in Yeast" Genetics (2009) 183:385-397. DOI:
% 10.1534/genetics.109.105254
%
% VARIABLES:
% none - it's a script. But you need Y5_3.mat in the working directory (or
% load the Y5 GEM using readcbmodel, and comment out the "load" line of the
% script.
%
% EXPECTED OUTPUT:
% growth rates and fluxes for the mutants listed in Table 1 of Kennedy et
% al.
%
% REQUIRED SOFTWARE:
% this script calls functions from the COBRA Toolbox and Gurobi solver.
%
%-------------------------------------------------------------

%%
%first, load the Y5.30 model.

load Y5_3;
model=Y5_3;
```

```
fprintf(' \r');
fprintf('Yeast model v 5.30 loaded.\n');
fprintf(' \r');

%one more change found later - sent to Kieran to add to the model, too.
model.lb([490 491])=0;model.ub([490 491])=0;

%make sure we're using the gurobi solver (glpk gives different answers!)
changeCobraSolver('gurobi');

fprintf('LP solver changed to gurobi.\n');
fprintf(' \r');

%%
%next modify the changed model by relaxing modeling constraints I imposed
%to get PPP pathway working as expected, and to add some strain-specific
%changes.

%allow rxns 994 and 995 to be reversible (transketolase 1 and 2)
model=changeRxnBounds(model,model.rxns(994),-1000,'l');
model=changeRxnBounds(model,model.rxns(995),-1000,'l');

%add the alt2 gene, which is not in Y5.01 b/c it's a putative ORF in SGD.
model=addReaction(model,'NRXN11',{'s_0180', 's_0955', 's_0991', ...
    's_1399'}, [-1 -1 1 1], true, -1000, 1000, 0, '', 'YDR111C');

%add the YPL276W ORF, which is not in Y5.30 b/c it's not in s288c per SGD.
%rxn 429 (formate dehydrogenase) can be catalyzed by FDH2 in CEN.PK 113-7D
%strain (see pmid: 11921099)
model=changeGeneAssociation(model,model.rxns(429),...
    '(YOR388C or (YPL275W AND YPL276W))');

%paper wt background is fdh1 fdh2 deletion
[model,hasEffect,constrRxnNames,deletedGenes] = ...
        deleteModelGenes(model,{'YOR388C','YPL275W','YPL276W'});

%paper strain background is wt for ura3 his3 leu2 trp1, and a mal2 mutant
%MAL2 isn't in s288c background, so no locus in SGD.
%background strain codes for SUC2 invertase, others are not in s288c
%background, so no locus in SGD

%%
%next, make the media the same as that in Kennedy et al. (based on their
%supplemental table 1).
%
%NOTE: the COBRA exchange reaction convention means that I need to use the
%opposite sign than Kennedy et al.
%
%NOTE 2: Table S1 lists "pydxn", "4abz", "nac", and "dhf" as having an
%upper bound of 0.5. I think these are: Pyridoxine (met 875),
%4-Aminobenzoate (met 98), Nicotinate (met 714), and 7,8-Dihydrofolate (met
%383). They don't have exchange reactions in iND750. Kennedy et al. may
%have added them, but they don't have fluxes in Table S2, so they're not
%essential.
%
%NOTE 3: I'm guessing rxn names from their met names.
```

```
%start with a clean slate: set all exchange reactions to upper bound = 1000
%and lower bound = 0 (ie, unconstrained excretion, no uptake)

exchangeRxns = findExcRxns(model);
model.lb(exchangeRxns)=0;
model.ub(exchangeRxns)=1000;

%the following exchange rxns have an upper bound and lower bound of 20 in
%Table S1 (ie, a fixed rate of uptake). I'm switching signs, so do bounds
%of -20.

ub_lb_20={'D-glucose exchange'};

for i=1:length(ub_lb_20)
    index=find(strcmp(ub_lb_20(i),model.rxnNames));
    model.lb(index)=-20;
    model.ub(index)=-20;
end

%the following have upper and lower bounds of inf in Table S1.
%
%NOTE: oxygen not in table S1, but added to do aerobic, based on
%text of paper

ub_inf={'water exchange', 'ammonium exchange', 'phosphate exchange', ...
    'sulphate exchange', 'sodium exchange', 'potassium exchange', ...
    'carbon dioxide exchange', 'oxygen exchange'};

for i=1:length(ub_inf)
    index=find(strcmp(ub_inf(i),model.rxnNames));
    model.lb(index)=-1000;
end

%the following have an upper bound of 0.5 in Table S1 (but I'm switching
%signs, so do a lower bound of -0.5).

ub_05 = {'L-asparagine exchange', 'L-aspartate exchange', ...
    'L-valine exchange', 'L-tyrosine exchange', ...
    'L-tryptophan exchange', 'L-threonine exchange', ...
    'L-serine exchange', 'L-proline exchange', ...
    'L-phenylalanine exchange', 'L-methionine exchange', ...
    'L-lysine exchange', 'L-leucine exchange', ...
    'L-isoleucine exchange', 'L-histidine exchange', ...
    'glycine exchange', 'L-glutamate exchange', ...
    'L-glutamine exchange', 'L-cysteine exchange', ...
    'L-arginine exchange', 'L-alanine exchange' ...
    'riboflavin exchange', 'thiamine(1+) exchange', ...
    'zymosterol exchange', 'uracil exchange', ...
    '(R)-pantothenate exchange', 'linoleic acid exchange', ...
    'myo-inositol exchange', 'palmitoleate exchange', ...
    'ergosterol exchange', 'biotin exchange', 'adenine exchange'};

%note: Y5 doesn't have octadecynoate (n-C18:2) or octadecenoate (n-C18:1)
%exchange reactions. I used linoleic (18:1). It also doesn't have
%hexadecenoate (n-C16:1). I used palmitoleate (16:1) for that. I don't
%think they should really matter for aerobic simulations, though.
```

```
for i=1:length(ub_05)
    index=find(strcmp(ub_05(i),model.rxnNames));
    model.lb(index)=-0.5; %constrained uptake
end

%the following have a lower bound of -inf in Table S1 (but I'm switching
%signs, so do an upper bound of 1000)
%
% Note: 'Deoxycytidine exchange' not included in Table S1, but has a flux
% in Table S2, so I added it here.

lb_inf = {'(1->3)-beta-D-glucan exchange', ...
    'gamma-aminobutyrate exchange', '(S)-3-methyl-2-oxopentanoate exchange',
...
    ' 8-amino-7-oxononanoate exchange', 'L-arabinitol exchange', ...
    'acetaldehyde exchange', 'acetate exchange', 'adenosine exchange', ...
    ' 2-oxoglutarate exchange', 'allantoin exchange', ...
    'S-adenosyl-L-methionine exchange', 'D-arabinose exchange', ...
    'L-arabinose exchange', 'choline exchange', 'citrate(3-) exchange', ...
    '(R)-carnitine exchange', 'cytosine exchange', 'cytidine exchange', ...
    '2''-deoxyadenosine exchange', ' 7,8-diaminononanoate exchange', ...
    '2''-deoxyguanosine exchange', '2''-deoxyinosine exchange', ...
    'dTTP exchange', '2''-deoxyuridine exchange', 'ethanol exchange', ...
    'FMN exchange', 'palmitate exchange', 'H+ exchange', ...
    'hypoxanthine exchange', 'inosine exchange', '(S)-lactate exchange', ...
    '(S)-malate exchange', 'maltose exchange', 'D-mannose exchange', ...
    'melibiose exchange', 'S-methyl-L-methionine exchange', ...
    'NMN exchange', 'stearate exchange', ...
    'ornithine exchange', 'adenosine 3'',5''-bismonophosphate exchange', ...
    'peptide exchange', 'putrescine exchange', 'pyruvate exchange', ...
    'D-ribose exchange', 'D-glucitol exchange', 'L-glucitol exchange', ...
    'spermidine exchange', 'spermine exchange', 'L-sorbose exchange', ...
    'succinate exchange', 'sucrose exchange', ...
    'thiamine(1+) monophosphate exchange', 'thiamine(1+) diphosphate(1-)
exchange', ...
    'thymidine exchange', 'thymine exchange', 'alpha,alpha-trehalose
exchange', ...
    'myristate exchange', 'urea exchange', ...
    'uridine exchange', '9H-xanthine exchange', 'xanthosine exchange', ...
    'D-xylose exchange', 'xylitol exchange', 'deoxycytidine exchange'};

%note: Y5 doesn't have 4-Aminobutanoate exchangeI used gamma-aminobutyrate.
%Y5 doesn't have 5-Amino-4-oxopentanoate exchange. I used
%'(S)-3-methyl-2-oxopentanoate exchange'.

for i=1:length(lb_inf)
    index=find(strcmp(lb_inf(i),model.rxnNames));
    model.lb(index)=0;
    model.ub(index)=1000;
end

fprintf('Exhange reaction bounds set.\n');
fprintf(' \r');

%%
%next step: demonstrate the results of Table 1 (strains 1-5, b/c yat2 isn't
%in iND750 - I'm not sure how strains 6-8 were generated)
```

```
%

%Kennedy reported in silico knockouts:
%1 fdh1=YOR388C fdh2=YPL275W/YPL276W alt2=YDR111C fum1=YPL262W zwf1=YNL241C
%2 fdh1 fdh2 aat2=YLR027C fum1=YPL262W zwf1=YNL241C
%3 fdh1 fdh2 cat2=YML042W fum1=YPL262W zwf1=YNL241C
%4 fdh1 fdh2 cat2=YML042W fum1=YPL262W rpe1=YJL121C
%5 fdh1 fdh2 cat2=YML042W fbp1=YLR377C fum1=YPL262W
%6 fdh1 fdh2 cat2=YML042W yat2=YER024W slc1=YDL052C
%7 fdh1 fdh2 cat2=YML042W yat2=YER024W cho1=YER026C
%8 fdh1 fdh2 cat2=YML042W yat2=YER024W alt2=YDR111C

%%
%check that FBA works on the model
wt_sln=optimizeCbModel(model,[],'one');

formate_ex=find(strcmp('formate exchange',model.rxnNames));

fprintf('The WT growth flux is: %6.4f.\n', wt_sln.f); %2.5889
fprintf('Formate exchange is reaction %d.\n', formate_ex); %0
fprintf('The formate exchange flux is: %6.4f.\n', wt_sln.x(formate_ex));
fprintf(' \r');

%make the knockouts and test them
%fdh1/2 mutant applied earlier
%[model,hasEffect,constrRxnNames,deletedGenes] = ...
%        deleteModelGenes(model,{'YOR388C','YPL275W','YPL276W'});

[KO1,hasEffect,constrRxnNames,deletedGenes] = ...
        deleteModelGenes(model,{'YDR111C','YPL262W','YNL241C'});
KO1_sln=optimizeCbModel(KO1,[],'one');

fprintf('The KO1 growth flux is: %6.4f.\n', KO1_sln.f); %1.6065
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO1_sln.x(formate_ex));
%0.0001
fprintf(' \r');

[KO2,hasEffect,constrRxnNames,deletedGenes] = ...
        deleteModelGenes(model,{'YLR027C', 'YPL262W', 'YNL241C'});
KO2_sln=optimizeCbModel(KO2,[],'one');

fprintf('The KO2 growth flux is: %6.4f.\n', KO2_sln.f); %1.6065
fprintf('Formate exchange is reaction %d.\n', formate_ex); %
fprintf('The formate exchange flux is: %6.4f.\n', KO2_sln.x(formate_ex));
%0.0001
fprintf(' \r');

[KO3,hasEffect,constrRxnNames,deletedGenes] = ...
        deleteModelGenes(model,{'YML042W', 'YPL262W', 'YNL241C'});
KO3_sln=optimizeCbModel(KO3,[],'one');

fprintf('The KO3 growth flux is: %6.4f.\n', KO3_sln.f); %1.6065
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO3_sln.x(formate_ex));
%0.0001
fprintf(' \r');
```

```
[KO4,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YPL262W', 'YJL121C'});
KO4_sln=optimizeCbModel(KO4,[],'one');

fprintf('The KO4 growth flux is: %6.4f.\n', KO4_sln.f); %1.6065
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO4_sln.x(formate_ex));
%0.0001
fprintf(' \r');

KO5=model;
[KO5,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YLR377C', 'YPL262W'});
KO5_sln=optimizeCbModel(KO5,[],'one');

fprintf('The KO5 growth flux is: %6.4f.\n', KO5_sln.f);
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO5_sln.x(formate_ex));
fprintf(' \r');

KO6=model;
[KO6,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YER024W', 'YDL052C'});
KO6_sln=optimizeCbModel(KO6,[],'one');

fprintf('The KO6 growth flux is: %6.4f.\n', KO6_sln.f);
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO6_sln.x(formate_ex));
fprintf(' \r');

KO7=model;
[KO7,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YER024W', 'YER026C'});
KO7_sln=optimizeCbModel(KO7,[],'one');

fprintf('The KO7 growth flux is: %6.4f.\n', KO7_sln.f);
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO7_sln.x(formate_ex));
fprintf(' \r');

KO8=model;
[KO8,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YER024W', 'YDR111C'});
KO8_sln=optimizeCbModel(KO8,[],'one');

fprintf('The KO8 growth flux is: %6.4f.\n', KO8_sln.f);
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO8_sln.x(formate_ex));
fprintf(' \r');

%save('Y_5_3_KO_model','model');
```

Supplementary File 5.4.3: apply_y5_3_changes.m script

```
% FILE NAME: apply_y5_3_changes
%
% DATE CREATED: 23 May 2012
%
% PROGRAMMER: B. Heavner
% Department of Biological
% and Environmental Engineering
% Cornell University
% Ithaca, NY 14853
%
% LAST REVISED: 11 July 2012
%
% REVISIONS:
%
% 7/11/12 - BDH - cleaned up for publication
% 7/4/12 - BDH - added lines 396-421 (NAD/P constraints, and some rxn
% changes)
% 5/23/12 - BDH - started with Y5.30
%
% PURPOSE:
% To document and apply all changes I find to correct the Y5.30 model.
%
% REFERENCES:
% Various, included as comments to changes
%
% VARIABLES:
% none - it's a script. But you need Y5_3.mat in the working directory (or
% load the Y5 GEM using readcbmodel, and comment out the "load" line of the
% script.
%
% EXPECTED OUTPUT:
% updated model based on Y5.30
%
% REQUIRED SOFTWARE:
% this script calls functions from the COBRA Toolbox and the Gurobi solver.
%
%-------------------------------------------------------------

%%
%first, load the Y5.30 model. This matlab variable was made by running
% readcbmodel to load a .sbml file downloaded from yeast.sf.net on
% 5/21/12

load Y5_3.mat;
model=Y5_3;

fprintf(' \r');
fprintf('Yeast model v 5.30 loaded.\n');
fprintf(' \r');

%make sure we're using the gurobi solver (glpk gives different answers!)
changeCobraSolver('gurobi');

fprintf('LP solver changed to gurobi.\n');
fprintf(' \r');
```

```
%check the media:
exchangeRxns = findExcRxns(model);
temp=find(exchangeRxns);
temp2=find(model.lb(temp));
media_indexes=temp(temp2);

fprintf('Media rxns\n')
fprintf('rxnName \t lb\n')
for i=1:length(media_indexes)
    fprintf('%s\t%g\n',model.rxnNames{media_indexes(i)},
model.lb(media_indexes(i)));
end

fprintf(' \r');

%Check the solution prior to any changes
unchangedsolution = optimizeCbModel(model,[],'one');

%%
%model updates and fixes - note: add new to the end - numbers may change!
%I'm working off a list of reactions generated by examining genes that only
%catalyze one reaction in iND750, but catalyze more than one reaction in
%Y5.30

%%
%Start with a tricky one - succinate and fumarate stuff
%
%rxns 439 and 945 (r_0455 and r_1000) are annotated with YEL047C (FRD1),
%citing http://identifiers.org/kegg.reaction/R00408. However, that KEGG rxn
%isn't present in saccharomyces cerevsiae. Looking at SGD to determine
%YEL047C's function, it states "Soluble fumarate reductase, required with
%isoenzyme Osm1p for anaerobic growth". osm1 is YJR051W. of OSM1, SGD says
%"Fumarate reductase, catalyzes the reduction of fumarate to succinate,
%required for the reoxidation of intracellular NADH under anaerobic
%conditions"

%PMID: 8946166 says that FRD1 binds FAD. So, it may be that SGD is
%confusing or perhaps FRD1 uses FAD, and OSM1 uses NAD. for now, I'll stick
%with them both using FAD.

%per SGD, OSM1 localizes to the mitochondria, and FRD1 localizes to the
%cytosol (or mitochondria or ribosome)...

%OSM1 currently catalyzes mitochondrial rxns 438 and 440, involving FAD and
%FMN, not NAD/H..
% 438 cites http://identifiers.org/kegg.reaction/R00408 (which uses FAD,
% but is not in Yeast, per KEGG)
% 440 doesn't cite a source.
%
%looking at rxn 438, it's catalyzed by "(YJR051W or (YDR178W and YJL045W
%and YKL141W and YLL041C) or (YDR178W and YKL141W and YKL148C and
%YLL041C))"
%
%from SGD's description of these other ORFs, including YJR051W with them
%appears to be wrong. (the others form succinate dehydrogenase, which
%oxidizes succinate using ubiquinone as the electron acceptor, as shown in
```

```
%rxn 966).
%
%additionally, rxn 438 is constrained to work only fumarate -> succinate,
%which is opposite from what SGD describes and what YeastCyc shows for
%succinate dehydrogenase. And, Y5_3 has it using FAD, not ubiquinone, which
%differs from Yeastcyc's description of succinate dehydrogenase.

%proposed fixes: 1) reannoatate rxn 438 to use YEL047C or YEL047C (cite:
%PMID: 9587404 and PMID: 8946166). The rxn catalyzed by succinate
%dehydrogenase is more correctly covered by rxn 966; 2) remove rxn 440; 3)
%change rxn 439 to use FAD cofactor;

model=changeGeneAssociation(model,model.rxns(438), '(YEL047C or YEL047C)');
model=changeRxnBounds(model,model.rxns(440),0,'l');
model=changeRxnBounds(model,model.rxns(440),0,'u');

model.S([709 712],439)=0; model.S(685,439)=-1; model.S(683,439)=1;

%TCA cycle still ok

%%
%MetaCyc shows rxn 966 constrained to go in forward direction only
model=changeRxnBounds(model,model.rxns(966),0,'l');

%TCA cycle still ok

%%
%While looking at succinate dehydrogenase, notice rxn 965 - no ref, rxn
%shown is kinda like that annotated by YML120C in yeastCyc.. But YML120C
%annotates rxn 748 in Y5_3.

%proposed fixes: remove rxn 965, modify rxn 748, which per SGD shouldn't
%pump protons.

model=changeRxnBounds(model,model.rxns(965),0,'l');
model=changeRxnBounds(model,model.rxns(965),0,'u');

%TCA ok

%model.S([787 782],748)=0;
% breaks the TCA and glyoxylate cycle fluxes - perhaps the Yeast electron
% transport chain isn't correctly reconstructed?

%%
%pubmed: 16226833, cited for rxn 939, says YFL030W is highly specific for
%substrate - that described in rxn 153. It says there's very low activity
%for something like rxn 939. So, get rid of rxn 939 (note: this is related
%to smh1/2 and amino acid metabolism, as well as the glyoxylate pathway)

model=changeRxnBounds(model,model.rxns(939),0,'l');
model=changeRxnBounds(model,model.rxns(939),0,'u');

%%
%rxns catalyzed by YGL202W or YHR137W seem not quite right... the two genes
%catalyze different rxns, per SGD...

%Both YGL202W (ARO8) and YHR137W (ARO9) currently catalyze rxns 803 1001
```

```
%1008 (with OR)

% SGD gives the following refs for both ARO8 and ARO9: PMID: 9491082 PMID:
% 9491083

% rxn 803:  2-oxoglutarate + L-phenylalanine  <-> keto-phenylpyruvate +
% L-glutamate. SGD has this catalyzed by ARO8, reversible.

% rxn 1001:2-oxoglutarate + L-tryptophan <-> indole-3-pyruvate+
% L-glutamate. SGD has this catalyzed by ARO8, in -> direction.

% rxn 1008:3-(4-hydroxyphenyl)pyruvate + L-glutamate <-> 2-oxoglutarate+
% L-tyrosine. SGD has this catalyzed by ARO8, reversible.

%proposed fix: remove ARO9 from these rxns, add 3 more for it's work.
%Change rxn 1001 to be irreversible, change rxn 1008 to be reversible.

model=changeGeneAssociation(model,model.rxns(803), 'YGL202W');
model=changeGeneAssociation(model,model.rxns(1001), 'YGL202W');
model=changeGeneAssociation(model,model.rxns(1008), 'YGL202W');

model=changeRxnBounds(model,model.rxns(1001),0,'l');
model=changeRxnBounds(model,model.rxns(1008),-1000,'l');

% new ARO9 rxn 1:  pyruvate s_1399 + L-phenylalanine s_1032  <->
% keto-phenylpyruvate s_0951 + L-alanine s_0955. reversible.
model=addReaction(model,'NRXN1',{'s_1399', 's_1032', 's_0951', ...
     's_0955'}, [-1 -1 1 1], true, -1000, 1000,'', 'YHR137W');

% new ARO9 rxn 2: keto-phenylpyruvate s_0951 + L-tryptophan s_1048 ->
% indole-3-pyruvate s_0855 + L-phenylalanine s_1032. irreversible.
model=addReaction(model,'NRXN2',{'s_0951', 's_1048', 's_0855', ...
    's_1032'}, [-1 -1 1 1], false, 0, 1000,'', 'YHR137W');

% new ARO9 rxn 3: 3-(4-hydroxyphenyl)pyruvate s_0204 + L-alanine s_0955 <->
% pyruvate s_1399 + L-tyrosine s_1051. reversible.
model=addReaction(model,'NRXN3',{'s_0204', 's_0955', 's_1399', ...
    's_1051'}, [-1 -1 1 1], true, -1000, 1000,'', 'YHR137W');

%
%TCA fluxes are increased, at this point simulating Kennedy KOs gives same
%growth rates and formate fluxes for KO1,2,3,and 5.

%%
%rxns 754 and 755 are very similar, but use different species.
%Additionally, SGD localizes YGR010W to the nucleus, so it shouldn't
%annotate rxn 755. This merits more review - see the discussion of NAD
%biosynthesis at
%http://www.yeastgenome.org/cgi-bin/locus.fpl?locus=S000001116
%
%For now, it appears that both of these reactions are parts of NAD
%metabolism, though I think the reconstruction would benefit from closer
%review

%%
%should rxn 956 be annotated the same as rxn 955?
% from SGD, YHR042W seems incorrect for 956.
```

```
% but YNL111C seems wrong for 955, too. (and YIL043C and all others, which
% are all related to cytochrome B5, and locate in mitochondria)
%
% these reactions seems to incorporate ERG9's use of NAD(P), while SGD
% is not specific for electron donor/acceptor.

model=changeGeneAssociation(model,model.rxns(955), 'YGR175C');
model=changeGeneAssociation(model,model.rxns(956), 'YGR175C');

%%
%is rxn 1095 right? (1094 same as iND 549, FWIW)
%SGD says "it is still under debate as to whether Flx1p imports FAD into
%the mitochondria or exports it to the cytoplasm" but PMID: 14555654 says
%"Flx1p is proposed to be the mitochondrial FAD export carrier".
%rxn 1094 doesn't have a reference in Y5.30

%proposed fixes: remove rxn 1094. reverse constraints on rxn 1095:
model.lb(1094)=0;model.ub(1094)=0;
model.lb(1094)=-1000;model.ub(1094)=0;

%%
%rxns 271:273 are constrained; 274:285, catalyzed by same genes, are
%unconstrained. And 286:289, catalyzed by one of the genes, are also
%unconstrained. This portion of metabolism is related to protein
%glycosylation, which I know very little about. I suspect the reaction are
%catalyzed by both to support the observation that rer2 srt1 double
%deletion is lethal.
%
%the compartmentalization of these reactions differs from the
%compartmentalization of the genes in SGD.

%proposed mod: change annotation so all are catalyzed by both, all
%constrained
model=changeGeneAssociation(model,model.rxns(286), '(YBR002C or YMR101C');
model=changeGeneAssociation(model,model.rxns(287), '(YBR002C or YMR101C');
model=changeGeneAssociation(model,model.rxns(288), '(YBR002C or YMR101C');
model=changeGeneAssociation(model,model.rxns(289), '(YBR002C or YMR101C');
model.lb(274:289)=0;

%TCA fluxes are increased, at this point simulating Kennedy KOs gives same
%growth rates and formate fluxes for KO1,2,3,and 5.

%%
%rxns 266 and 267 seem similar, but one has more general species. Probably
%not quite right to have them both, but I don't know how they're connected
%to other portions of PC metabolism. 266 is unconstrained, 267 is
%constrained.
%suggested mod: constrain 266

model.lb(266)=0;

%%
%YOR128C catalyzes rxns 481 and 863. Though 863 is in iND too, neither
%seems to be the rxn specified by SGD/Yeastcyc:
%http://pathway.yeastgenome.org/YEAST/NEW-IMAGE?type=REACTION-IN-
PATHWAY&object=AIRCARBOXY-RXN&detail-level=2
```

```
%rxn 481 is annotated with http://identifiers.org/pubmed/21190580 - the
%iMM904 paper; and http://identifiers.org/kegg.reaction/R06974 - which
%doesn't link to a enzyme or pathway

%rxn 863 is annotated with http://identifiers.org/kegg.reaction/R04209,
%which is closer to what's in SGD/Yeastcyc

%Proposed fix: remove rxn 481, modify rxn 863 to match SGD/Yeastcyc
model.lb(481)=0;model.ub(481)=0;
model.S([433 791],863)=-1;model.S([393 1304],863)=1;

%%
% PPP checks in light of PMID: 21663798

% I don't know about rxn 839, but its reverse should exist. (there's a fair
% amount of lit supporting rxn 839 in other organisms, particularly
% rabbit).

% The reverse rxn is catalyzed by SHB17, YKR043C. And it doesn't take ATP,
% and it liberates a Phosphate.  (PMID: 21663798)

model=addReaction(model,'NRXN4',{'s_1426', 's_1427', 's_1322'}, [-1 1 1], ...
    false, 0, 1000,'', 'YKR043C');


%%
%changes from DOI: 10.1111/j.1742-4658.2012.08649.x
model.lb(108)=-1000; model.ub(108)=0;
model.lb([438 439 440])=-1000;
model.ub(440)=1000;
model.ub(470)=0; %not sure about ref on this one
model.lb(486)=0; %not definitive, but paper noted why
model.ub([639 640])=0; %a regulatory constraint
model.ub([997 998])=0; %to force TAG synthesis flux
model.lb(1054)=-1000; %but I thought ATP couldn't translate to mito...have to
review refs
model.lb(1575)=0;model.ub(1575)=0;
model.lb(1684)=-1000; %ref?
model.lb(1747)=-1000; %ref?
%NADP phosphatases have already been removed.

model=addReaction(model,'NRXN5',{'s_0061', 's_1198', 's_1269', 's_1203'}, ...
    [-1 -1 1 1], false, 0, 1000,'', 'YKL216W');
model=addReaction(model,'NRXN6',{'s_0973', 's_0441', 's_0734'}, [-1 1 1], ...
    false, 0, 1000,'', 'YMR250W ');
model=addReaction(model,'NRXN7',{'s_1262', 's_1263'}, [-1 1], ...
    false, 0, 1000,'', '');
model=addReaction(model,'NRXN8',{'s_1262', 's_1264'}, [-1 1], ...
    false, 0, 1000,'', '');
model=addReaction(model,'NRXN9',{'s_0529', 's_0531'}, [-1 1], ...
    true, -1000, 1000,'', '');
model=addReaction(model,'NRXN10',{'s_0799', 's_0794',}, [-1 1], ...
    false, 0, 0,'', ''); %to be tuned for P/O ratio..

%%
% NAD/P constraints
```

```
%change constraints as documented in 6/8/12 notes.
% or, see 6/13 notes:

%relax lit-supported NAD rxns
model.lb([163 166 176 179 183 428 454 469 690 933 1888])=-1000;
model.ub([163 166 176 179 183 428 454 469 690 933 1888])=1000;

%refs:
%163 - YDL168W catalyzes the following in Y5: [163 166 176 179 183 428].
%428 should be reversible, others also catalyzed by other genes:
%   YBR145W or YDL168W or YOL086C. For now, make these all reversible, but
%   a better fix could be to split rxns 163 166 176 179 183 to two, with
%   only the version catalyzed by YDL168W reversible - o ther ADH rxns may
%   not be reversible
%428  - sfa1 reversible per SGD
%454 - "overexpression of GDH2 can promote catalysis of the reverse reaction"
%469 - glyceraldehyde-3-phosphate dehydrogenase - glycolysis and
gluconeogenesis
%690 - malate dehydrogenase - catalyze interconversion of malate and
oxaloacetate (per SGD)
%933 - PMID: 418069 - but lysene biosynthsis in Yeast merits more attention
%1888 - NAD transport (nucleus/cyto)

%and a ref that BDH1 (rxn 3) isn't reversible: "Therefore, the enzyme would
%preferentially function as a reductase rather than as a dehydrogenase." -
%PMID 10938079
model.lb(3)=-1000; model.ub(3)=0;

%relax lit-supported NADP
model.lb([310])=-1000; %removed 455, 934
model.lb([656])=0; %PMID: 12535615

%refs:
%310 - http://dx.doi.org/10.1016/j.jmb.2007.05.097, but for different
%organism
%455 - ?
%934 - PMID 3098733 says reversible at pH 9.3

%other refs:
% 174 - aldehyde dehydrogenase, YPL061W
%   YPL061W is ALD6. DOI: 10.1111/j.1432-1033.1971.tb01574.x says
%   irreversible in pig brain. Make it irreversible.
%
% 464 - glutathione oxidoreductase, ((YCL035C and YPL091W) or (YDR098C and
%   YPL091W) or (YDR513W and YPL091W) or (YER174C and YPL091W))
%   - seems irreversible, but haven't found much lit to support yet.

% look at SGD lit links for YPL091W
% PMID: 22093810 - http://www.ncbi.nlm.nih.gov/pubmed/22093810
% PMID: 21549177 - http://www.ncbi.nlm.nih.gov/pubmed/21549177
%
% note: if I make 174 and 464 irreversible, big - fluxes through 310 and
375...
%
% 310 is D-arabinose 1-dehydrogenase (NADP), YBR149W
%   PMID: 17151466 suggests limited fluxes
%   http://dx.doi.org/10.1016/j.jmb.2007.05.097 says reversible
```

```
%
% 375 is  fatty acid synthase (n-C16:1), YJL196C
%   I think unlikely to be reversible.
%
%538    hydroxymethylglutaryl CoA reductase (YLR450W or YML075C)
%   YLR450W is HMG2, irreversible http://genomebiology.com/2004/5/11/248
%   YML075C is HMG1, irreversible http://genomebiology.com/2004/5/11/248
%
%   656 L-allo-threonine dehydrogenase YMR226C
%   YMR226C was TMA29, but doesn't have a name now..
%   should be irreversible, towards NADPH. (PMID: 12535615)

%%
% rxn 783 is annotated with http://identifiers.org/pubmed/8511969. Per that
% paper, the rxn should be annotated with cys3 = YAL012W. And it should be
% reversible

model.lb(783)=-1000;
model=changeGeneAssociation(model,model.rxns(783), 'YAL012W');

%%
%rxn 1010 (r_1065) is annotated with YLR027C, but should be YGL202W, and
%should be reversible (per KEGG). This is already rxn 1008, so 1010 should
%be removed.

model.lb(1010)=0;model.ub(1010)=0;

%%
%check rxns of glycolosys, gluconeogenesis, TCA cycle, gloxylate cycle, and
%Pentose Phosphate pathway. At this point, there are changes between the
%pre-modified model and the post-modified model fluxes.

%note that rxn numbers in comments below are wrong (they're based on 5.01)

%An essential question is: what SHOULD the fluxes be?

%%
%glycolysis -

%pathway
%rxn 1165 -'glucose transport' - exc to cyto ->
%rxn 534 - 'hexokinase (D-glucose:ATP)' - cyto ->
%rxn 467 'glucose-6-phosphate isomerase' - cyto <->
%rxn 885 'phosphofructokinase' - cyto ->
%rxn 450 'fructose-bisphosphate aldolase' - cyto <->
%rxn 1053 'triose-phosphate isomerase' - cyto <->
%rxn 486 'glyceraldehyde-3-phosphate dehydrogenase' - cyto <->
%rxn 891 'phosphoglycerate kinase' - cyto <->
%rxn 892 'phosphoglycerate mutase' - cyto <->
%rxn 366 'enolase' - cyto <->
%rxn 961 'pyruvate kinase' - cyto ->
%rxn 958 'pyruvate decarboxylase' - cyto ->

%%
%gluconeogenesis -

%pathway
```

```
%rxn 957 'pyruvate carboxylase'
%rxn 883 'phosphoenolpyruvate carboxykinase'
%rxn 366 'enolase' - cyto <->
%rxn 892 'phosphoglycerate mutase' - cyto <->
%rxn 891 'phosphoglycerate kinase' - cyto <->
%rxn 486 'glyceraldehyde-3-phosphate dehydrogenase' - cyto <->
%rxn 450 'fructose-bisphosphate aldolase' - cyto <->
%rxn 449 'fructose-bisphosphatase' - cyto ->
%rxn 467 'glucose-6-phosphate isomerase' - cyto <->

%%
%Pentose Phosphate pathway - post modifications, these the first 3 rxns on
%the following list don't have fluxes, and the last 4 are negative

%pathway
%rxn 466 'glucose 6-phosphate dehydrogenase'
%rxn 91 '6-phosphogluconolactonase'
%rxn 888 'phosphogluconate dehydrogenase'
%rxn 981 'ribose-5-phosphate isomerase'
%rxn 983 'ribulose 5-phosphate 3-epimerase'
%rxn 1048 'transketolase 1'
%rxn 1047 'transaldolase'
%rxn 1049 'transketolase 2'

%%
%TCA cycle (in mito) - post modifications, these the first 5 and final rxns
%on the following list have fluxes, and the others do not. If rxn 714 is
%irreversible, all TCA rxns have positive fluxes. I'll examine with
%irreversible rxn 714 first.

%pathway
% rxn 2032 'pyruvate transport'
% rxn 960 'pyruvate dehydrogenase'
% rxn 300 'citrate synthase'
% rxn 302 'citrate to cis-aconitate(3-)'
% rxn 280 'cis-aconitate(3-) to isocitrate'
% rxn 658 'isocitrate dehydrogenase (NAD+)'
% rxn 831 'oxoglutarate dehydrogenase (lipoamide)'
% rxn 830 'oxoglutarate dehydrogenase (dihydrolipoamide S-
succinyltransferase)'
% rxn 1021 'succinate-CoA ligase (ADP-forming)'
% rxn 1020 'succinate dehydrogenase (ubiquinone-6)'
% rxn 451 'fumarase'
% rxn 713 'malate dehydrogenase'


%%
%glyoxylate cycle (mostly in cyto). Post-changes (irreversible rxn 714), x
%of the following N rxns have fluxes.

%"In the presence of 1% glucose, the synthesis of all enzymes, except
%fumarase, was repressed. However, no indication for a specific regulation
%mechanism for the entire cycle could be found. Studies on the localization
%of the glyoxylate cycle enzymes in the yeast cell revealed that the key
%enzymes, isocitrate lyase and malate synthase, are located in the
%cytoplasm, whereas succinate dehydrogense was found only in the
%mitochondrial fraction. Activites of all other enzymes are found in the
```

```
%cytoplasm as well as in the mitochondira. In anaerobically grown cells, no
%mitochondria could be detected. Succinate dehydrogenase, isocritrate
%lyase, and malate synthase were absent. However, appreciable activities of
%citrate synthase, aconitase, fumarase, and malate dehydrogenase were found
%in the Cytoplasm." - http://dx.doi.org/10.1111/j.1432-1033.1969.tb00658.x

% pathway (starting with Malate)

% rxn 713 'malate dehydrogenase' - malate to oxaloacetate(mito)
% rxn 714 'malate dehydrogenase, cytoplasmic ' - 0 flux
% rxn 300 'citrate synthase' - mitochondrial generation of citrate
% MISSING RXN - 'citrate synthase' - cyto? - YCR005C?
% rxn 302 'citrate to cis-aconitate(3-)' - mitochondrial
% rxn 303 'citrate to cis-aconitate(3-), cytoplasmic'
% rxn 280 'cis-aconitate(3-) to isocitrate' - mitochondrial - YJL200C?
% MISSING RXN? 'cis-aconitate(3-) to isocitrate' - cyto? - YLR304C?
% rxn 662 'isocitrate lyase' - isocitrate to glyoxylate and succinate
% rxn 716 'malate synthase' - glyoxylate to malate (in cyto)
% rxn 1020 'succinate dehydrogenase (ubiquinone-6)' - succ to fum (mito)
% rxn 452 'fumarase' - fumarate to malate (cyto)
% rxn 451 'fumarase' - fumarate to malate (mito)

%tansport rxns
% rxn 1225 'malate transport' - to mitochondria
% rxn 1238 'oxaloacetate transport' - to mitochondria
% rxn 1127 'citrate transport' - to cyto
% rxn 1264 'succinate-fumarate transport' - succinate to mito

%% output
rxnNames = {
    % glycolysis
    'GLYCOLYSIS'
    'glucose transport'
    'hexokinase (D-glucose:ATP)'
    'glucose-6-phosphate isomerase'
    'phosphofructokinase'
    'fructose-bisphosphate aldolase'
    'triose-phosphate isomerase'
    'glyceraldehyde-3-phosphate dehydrogenase'
    'phosphoglycerate kinase'
    'phosphoglycerate mutase'
    'enolase'
    'pyruvate kinase'
    'pyruvate decarboxylase'
    ''
    %ethanol production
    %'alcohol dehydrogenase, reverse rxn (acetaldehyde -> ethanol)'
    %should be 'alcohol dehydrogenase (ethanol)'
    'ETHANOL PRODUCTION'
    'alcohol dehydrogenase (ethanol)'
    'ethanol transport'
    ''
    % gluconeogenesis
    'GLUCONEOGENESIS'
    'pyruvate carboxylase'
    'phosphoenolpyruvate carboxykinase'
    'enolase'
```

```
    'phosphoglycerate mutase'
    'phosphoglycerate kinase'
    'glyceraldehyde-3-phosphate dehydrogenase'
    'fructose-bisphosphate aldolase'
    'fructose-bisphosphatase'
    'glucose-6-phosphate isomerase'
    ''
    % PPP
    'PENTOSE PHOSPHATE PATHWAY'
    'glucose 6-phosphate dehydrogenase'
    '6-phosphogluconolactonase'
    'phosphogluconate dehydrogenase'
    'ribose-5-phosphate isomerase'
    'ribulose 5-phosphate 3-epimerase'
    'transketolase 1'
    'transaldolase'
    'transketolase 2'
    ''
    % TCA
    'TCA CYCLE'
    %'pyruvate transport' %commented out b/c multiple rxns have this name
    'pyruvate dehydrogenase'
    'citrate synthase'
    'citrate to cis-aconitate(3-)'
    'cis-aconitate(3-) to isocitrate'
    'isocitrate dehydrogenase (NAD+)'
    'oxoglutarate dehydrogenase (lipoamide)'
    'oxoglutarate dehydrogenase (dihydrolipoamide S-succinyltransferase)'
    'succinate-CoA ligase (ADP-forming)'
    'succinate dehydrogenase (ubiquinone-6)'
    'fumarase'
    'malate dehydrogenase'
    ''
    %glyoxylate cycle
    'GLYOXYLATE CYCLE'
    'malate dehydrogenase'
    'malate dehydrogenase, cytoplasmic'
    'citrate synthase'
    'citrate synthase'
    'citrate to cis-aconitate(3-)'
    'citrate to cis-aconitate(3-), cytoplasmic'
    'cis-aconitate(3-) to isocitrate'
    'isocitrate lyase'
%    'malate synthase' %commented out b/c multiple rxns have this name
    'succinate dehydrogenase (ubiquinone-6)'
    'fumarase'
    'malate transport'
%    'oxaloacetate transport' %commented out b/c multiple rxns have this name
%    'citrate transport' %commented out b/c multiple rxns have this name
    'succinate-fumarate transport'
    ''
    % other
    };

fprintf('\nAerobic growth prior to
changes\nrate:\t%.2f\n\n',unchangedsolution.f);
```

```
for k = 1:length(rxnNames)
   ind = strcmp(rxnNames{k},model.rxnNames);
   if sum(ind)>1, disp(rxnNames{k}); end
   fprintf('%.2f\t%s\n',unchangedsolution.x(ind),rxnNames{k});
end

FBAsolution = optimizeCbModel(model,[],'one');
fprintf('Aerobic growth after changes\nrate:\t%.2f\n\n',FBAsolution.f);

for k = 1:length(rxnNames)
   ind = strcmp(rxnNames{k},model.rxnNames);
   if sum(ind)>1, disp(rxnNames{k}); end
   fprintf('%.2f\t%s\n',FBAsolution.x(ind),rxnNames{k});
end

save('updated_Y5_3','model');
```

Supplementary File 5.4.3: Y5_3_three_knockout1_v2.m script

```
% FILE NAME: Y5_three_knockout1_v2
%
% DATE CREATED: 11 June 2012
%
% PROGRAMMER: B. Heavner
% Department of Biological
% and Environmental Engineering
% Cornell University
% Ithaca, NY 14853
%
% LAST REVISED: 7/11/12
%
% REVISED BY: BDH
%
% REVISIONS:
% 7/11/12 - cleaned up for publication
%
% PURPOSE:
% To attempt to reproduce the model results reported by Kennedy et al. with
% the updated Yeast 5.30 model instead of the iND750 model.
%
% This updated version incorporates various model updates that I've applied
% prior to the file creation date.
%
% REFERENCE:
% Kennedy CJ, Boyle PM, Waks Z, Silver PA. "Systems-Level Engineering of
% Nonfermentative Metabolism in Yeast" Genetics (2009) 183:385-397. DOI:
% 10.1534/genetics.109.105254
%
% VARIABLES:
% none - it's a script. But you need updated_Y5_3.mat in the working
% directory.
%
% EXPECTED OUTPUT:
% growth rates and fluxes for the mutants listed in Table 1 of
% Kennedy et al.
%
% REQUIRED SOFTWARE:
% this script calls functions from the COBRA Toolbox and Gurobi solver.
%
%-------------------------------------------------------------

%%
%first, load the Y5.30 model that I've updated with the apply_y5_3_changes
%script.

load updated_Y5_3;
%model=Y5_3;

fprintf(' \r');
fprintf('Updated Yeast model v 5.30 loaded.\n');
fprintf(' \r');

%make sure we're using the gurobi solver (glpk gives different answers!)
changeCobraSolver('gurobi');
```

```
fprintf('LP solver changed to gurobi.\n');
fprintf(' \r');

%%
% now begin modifying model.

%a change not yet incorporated.
model.lb([490 491])=0;model.ub([490 491])=0;

%next modify the changed model by relaxing modeling constraints I imposed
%to get PPP pathway working as expected, and to add some strain-specific
%changes.

%allow rxns 994 and 995 to be reversible (transketolase 1 and 2)
model=changeRxnBounds(model,model.rxns(994),-1000,'l');
model=changeRxnBounds(model,model.rxns(995),-1000,'l');

%add the alt2 gene, which is not in Y5.30 b/c it's a putative ORF in SGD.
model=addReaction(model,'NRXN11',{'s_0180', 's_0955', 's_0991', ...
    's_1399'}, [-1 -1 1 1], true, -1000, 1000, 0, '', 'YDR111C');

%add the YPL276W ORF, which is not in Y5.30 b/c it's not in s288c per SGD.
%rxn 429 (formate dehydrogenase) can be catalyzed by FDH2 in CEN.PK 113-7D
%strain (see pmid: 11921099)
model=changeGeneAssociation(model,model.rxns(429),...
    '(YOR388C or (YPL275W AND YPL276W))');

%paper wt background is fdh1 fdh2 deletion
[model,hasEffect,constrRxnNames,deletedGenes] = ...
        deleteModelGenes(model,{'YOR388C','YPL275W','YPL276W'});

%paper strain background is wt for ura3 his3 leu2 trp1, and a mal2 mutant
%MAL2 isn't in s288c background, so no locus in SGD.
%background strain codes for SUC2 invertase, others are not in s288c
%background, so no locus in SGD

%the next step is optional, and I did it for testing.
%add the YML082W ORF, which is not in Y5.01 b/c it's a putative ORF in SGD,
%but is in iND750 and has a flux in the fum1/zwf1 mutant strain
%model=addReaction(model,'NRXN12',{'s_0803', 's_1241', 's_0178', ...
%    's_0794', 's_0419' ,'s_1458'}, [-1 -1 1 1 1 1], true, -1000, 1000, ...
%    0, '', 'YML082W');

%%
%next, make the media the same as that in Kennedy et al. (based on their
%supplemental table 1).
%
%NOTE: the COBRA exchange reaction convention means that I need to use the
%opposite sign than Kennedy et al.
%
%NOTE 2: Table S1 lists "pydxn", "4abz",
%"nac", and "dhf" as having an upper bound of 0.5. I think these are:
%Pyridoxine (met 875), 4-Aminobenzoate (met 98), Nicotinate (met 714), and
%7,8-Dihydrofolate (met 383). They don't have exchange reactions in iND750.
%Kennedy et al. may have added them, but they don't have fluxes in Table
%S2, so they're not essential.
```

```
%
%NOTE 3: I'm guessing rxn names from their met names.

%start with a clean slate: set all exchange reactions to upper bound = 1000
%and lower bound = 0 (ie, unconstrained excretion, no uptake)

exchangeRxns = findExcRxns(model);
model.lb(exchangeRxns)=0;
model.ub(exchangeRxns)=1000;

%the following exchange rxns have an upper bound and lower bound of 20 in
%Table S1 (ie, a fixed rate of uptake). I'm switching signs, so do bounds
%of -20.

ub_lb_20={'D-glucose exchange'};

for i=1:length(ub_lb_20)
    index=find(strcmp(ub_lb_20(i),model.rxnNames));
    model.lb(index)=-20;
    model.ub(index)=-20;
end

%the following have upper and lower bounds of inf in Table S1.
%
%NOTE: oxygen not in table S1, but added to do aerobic, based on
%text of paper

ub_inf={'water exchange', 'ammonium exchange', 'phosphate exchange', ...
    'sulphate exchange', 'sodium exchange', 'potassium exchange', ...
    'carbon dioxide exchange', 'oxygen exchange'};

for i=1:length(ub_inf)
    index=find(strcmp(ub_inf(i),model.rxnNames));
    model.lb(index)=-1000;
end

%the following have an upper bound of 0.5 in Table S1 (but I'm switching
%signs, so do a lower bound of -0.5).

ub_05 = {'L-asparagine exchange', 'L-aspartate exchange', ...
    'L-valine exchange', 'L-tyrosine exchange', ...
    'L-tryptophan exchange', 'L-threonine exchange', ...
    'L-serine exchange', 'L-proline exchange', ...
    'L-phenylalanine exchange', 'L-methionine exchange', ...
    'L-lysine exchange', 'L-leucine exchange', ...
    'L-isoleucine exchange', 'L-histidine exchange', ...
    'glycine exchange', 'L-glutamate exchange', ...
    'L-glutamine exchange', 'L-cysteine exchange', ...
    'L-arginine exchange', 'L-alanine exchange' ...
    'riboflavin exchange', 'thiamine(1+) exchange', ...
    'zymosterol exchange', 'uracil exchange', ...
    '(R)-pantothenate exchange', 'linoleic acid exchange', ...
    'myo-inositol exchange', 'palmitoleate exchange', ...
    'ergosterol exchange', 'biotin exchange', 'adenine exchange'};

%note: Y5 doesn't have octadecynoate (n-C18:2) or octadecenoate (n-C18:1)
%exchange reactions. I used linoleic (18:1). It also doesn't have
```

```
%hexadecenoate (n-C16:1). I used palmitoleate (16:1) for that. I don't
%think they should really matter for aerobic simulations, though.

for i=1:length(ub_05)
    index=find(strcmp(ub_05(i),model.rxnNames));
    model.lb(index)=-0.5; %constrained uptake
end

%the following have a lower bound of -inf in Table S1 (but I'm switching
%signs, so do an upper bound of 1000)
%
% Note: 'Deoxycytidine exchange' not included in Table S1, but has a flux
% in Table S2, so I added it here.

lb_inf = {'(1->3)-beta-D-glucan exchange', ...
    'gamma-aminobutyrate exchange', '(S)-3-methyl-2-oxopentanoate exchange',
...
    ' 8-amino-7-oxononanoate exchange', 'L-arabinitol exchange', ...
    'acetaldehyde exchange', 'acetate exchange', 'adenosine exchange', ...
    ' 2-oxoglutarate exchange', 'allantoin exchange', ...
    'S-adenosyl-L-methionine exchange', 'D-arabinose exchange', ...
    'L-arabinose exchange', 'choline exchange', 'citrate(3-) exchange', ...
    '(R)-carnitine exchange', 'cytosine exchange', 'cytidine exchange', ...
    '2''-deoxyadenosine exchange', ' 7,8-diaminononanoate exchange', ...
    '2''-deoxyguanosine exchange', '2''-deoxyinosine exchange', ...
    'dTTP exchange', '2''-deoxyuridine exchange', 'ethanol exchange', ...
    'FMN exchange', 'palmitate exchange', 'H+ exchange', ...
    'hypoxanthine exchange', 'inosine exchange', '(S)-lactate exchange', ...
    '(S)-malate exchange', 'maltose exchange', 'D-mannose exchange', ...
    'melibiose exchange', 'S-methyl-L-methionine exchange', ...
    'NMN exchange', 'stearate exchange', ...
    'ornithine exchange', 'adenosine 3'',5''-bismonophosphate exchange', ...
    'peptide exchange', 'putrescine exchange', 'pyruvate exchange', ...
    'D-ribose exchange', 'D-glucitol exchange', 'L-glucitol exchange', ...
    'spermidine exchange', 'spermine exchange', 'L-sorbose exchange', ...
    'succinate exchange', 'sucrose exchange', ...
    'thiamine(1+) monophosphate exchange', 'thiamine(1+) diphosphate(1-)
exchange', ...
    'thymidine exchange', 'thymine exchange', 'alpha,alpha-trehalose
exchange', ...
    'myristate exchange', 'urea exchange', ...
    'uridine exchange', '9H-xanthine exchange', 'xanthosine exchange', ...
    'D-xylose exchange', 'xylitol exchange', 'deoxycytidine exchange'};

%note: Y5 doesn't have 4-Aminobutanoate exchangeI used gamma-aminobutyrate.
%Y5 doesn't have 5-Amino-4-oxopentanoate exchange. I used
%'(S)-3-methyl-2-oxopentanoate exchange'.

for i=1:length(lb_inf)
    index=find(strcmp(lb_inf(i),model.rxnNames));
    model.lb(index)=0;
    model.ub(index)=1000;
end

fprintf('Exhange reaction bounds set.\n');
fprintf(' \r');
```

```
%%
%next step: demonstrate the results of Table 1 (strains 1-5, b/c yat2 isn't
%in iND750 - I'm not sure how strains 6-8 were generated)
%

%Kennedy reported in silico knockouts:
%1 fdh1=YOR388C fdh2=YPL275W/YPL276W alt2=YDR111C fum1=YPL262W zwf1=YNL241C
%2 fdh1 fdh2 aat2=YLR027C fum1=YPL262W zwf1=YNL241C
%3 fdh1 fdh2 cat2=YML042W fum1=YPL262W zwf1=YNL241C
%4 fdh1 fdh2 cat2=YML042W fum1=YPL262W rpe1=YJL121C
%5 fdh1 fdh2 cat2=YML042W fbp1=YLR377C fum1=YPL262W
%6 fdh1 fdh2 cat2=YML042W yat2=YER024W slc1=YDL052C
%7 fdh1 fdh2 cat2=YML042W yat2=YER024W cho1=YER026C
%8 fdh1 fdh2 cat2=YML042W yat2=YER024W alt2=YDR111C

%looking at Yeast 5.01:
%fdh1 = gene 848, rxns: 445
%fdh2 = gene added by my chages, rxns: 445
%alt2 - added to model, above
%fum1 = gene 883 rxns: 451 452
%zwf1 = gene 743 rxns: 466
%aat2 = gene 588 rxns: 216 218 681 683 1065
%cat2 = gene 659 rxns: 253 254
%rpe1 = gene 476 rxns: 984
%fbp1 = gene 645 rxns: 449
%yat2 = gene 263 rxns: 252
%slc1 = gene 130 rxns: 9
%cho1 = gene 264 rxns: 880 881 - essential for model.

%%
%check that FBA works on the model
wt_sln=optimizeCbModel(model,[],'one');

formate_ex=find(strcmp('formate exchange',model.rxnNames));

fprintf('The WT flux is: %6.4f.\n', wt_sln.f); %2.5889
fprintf('Formate exchange is reaction %d.\n', formate_ex); %0
fprintf('The formate exchange flux is: %6.4f.\n', wt_sln.x(formate_ex));
fprintf(' \r');

%make the knockouts and test them
%fdh1/2 mutant applied earlier
%[model,hasEffect,constrRxnNames,deletedGenes] = ...
%        deleteModelGenes(model,{'YOR388C','YPL275W','YPL276W'});

[KO1,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YDR111C','YPL262W','YNL241C'});
KO1_sln=optimizeCbModel(KO1,[],'one');

fprintf('The KO1 flux is: %6.4f.\n', KO1_sln.f); %1.6065
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO1_sln.x(formate_ex));
%0.0001
fprintf(' \r');

[KO2,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YLR027C', 'YPL262W', 'YNL241C'});
```

```
KO2_sln=optimizeCbModel(KO2,[],'one');

fprintf('The KO2 flux is: %6.4f.\n', KO2_sln.f); %1.6065
fprintf('Formate exchange is reaction %d.\n', formate_ex); %
fprintf('The formate exchange flux is: %6.4f.\n', KO2_sln.x(formate_ex));
%0.0001
fprintf(' \r');

[KO3,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YPL262W', 'YNL241C'});
KO3_sln=optimizeCbModel(KO3,[],'one');

fprintf('The KO3 flux is: %6.4f.\n', KO3_sln.f); %1.6065
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO3_sln.x(formate_ex));
%0.0001
fprintf(' \r');

[KO4,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YPL262W', 'YJL121C'});
KO4_sln=optimizeCbModel(KO4,[],'one');

fprintf('The KO4 flux is: %6.4f.\n', KO4_sln.f); %1.6065
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO4_sln.x(formate_ex));
%0.0001
fprintf(' \r');

KO5=model;
[KO5,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YLR377C', 'YPL262W'});
KO5_sln=optimizeCbModel(KO5,[],'one');

fprintf('The KO5 flux is: %6.4f.\n', KO5_sln.f);
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO5_sln.x(formate_ex));
fprintf(' \r');

KO6=model;
[KO6,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YER024W', 'YDL052C'});
KO6_sln=optimizeCbModel(KO6,[],'one');

fprintf('The KO6 flux is: %6.4f.\n', KO6_sln.f);
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO6_sln.x(formate_ex));
fprintf(' \r');

KO7=model;
[KO7,hasEffect,constrRxnNames,deletedGenes] = ...
       deleteModelGenes(model,{'YML042W', 'YER024W', 'YER026C'});
KO7_sln=optimizeCbModel(KO7,[],'one');

fprintf('The KO7 flux is: %6.4f.\n', KO7_sln.f);
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO7_sln.x(formate_ex));
fprintf(' \r');
```

```
KO8=model;
[KO8,hasEffect,constrRxnNames,deletedGenes] = ...
        deleteModelGenes(model,{'YML042W', 'YER024W', 'YDR111C'});
KO8_sln=optimizeCbModel(KO8,[],'one');

fprintf('The KO8 flux is: %6.4f.\n', KO8_sln.f);
fprintf('Formate exchange is reaction %d.\n', formate_ex);
fprintf('The formate exchange flux is: %6.4f.\n', KO8_sln.x(formate_ex));
fprintf(' \r');

save('updated_Y_5_3_KO_model','model');
```

Supplementary File 5.4.4: Y5_3_two_knockout2.m script

```
% FILE NAME: Y5_3_two_knockout2
%
% DATE CREATED: 2 February 2012
%
% PROGRAMMER: B. Heavner
% Department of Biological
% and Environmental Engineering
% Cornell University
% Ithaca, NY 14853
%
% LAST REVISED: 11 July 2012
%
% REVISED BY: B. Heavner
%
% REVISIONS:
% 11 July 2012 - BDH cleaned up for publishing
% 13 June 2012 - BDH edit to work with updated constraints of new Y5.30
% 22 May 2012 - BDH edit to work with Y5.30 instead of modified 5.01
%
% PURPOSE:
% A variation on the model results reported by Kennedy et al. which
% compares all viable 2-knockout mutants in iND750 to find those that will
% cause formic acid secretion.
%
% In this script, I apply a "global" search for double deletions that
% excrete formic acid to the the constrained Y5.30 I made in the
% Y5_3_three_knockout1_v2 script (the variable "model", which has the same
% media as Kennedy et al.)
%
%
% REFERENCE:
% Kennedy CJ, Boyle PM, Waks Z, Silver PA. "Systems-Level Engineering of
% Nonfermentative Metabolism in Yeast" Genetics (2009) 183:385-397. DOI:
% 10.1534/genetics.109.105254
%
% VARIABLES:
% this script requires the updated_Y_5_3_KO_model.mat variable built by
% Y5_3_three_kncokout1_v2.m in the working directory.
%
% EXPECTED OUTPUT:
% a list of double-gene deleted mutants that are predicted to excrete
% formate.
%
%------------------------------------------------------------

%%
%first, load the modified Y5.30

load('updated_Y_5_3_KO_model')

%double check that the model is a fdh1/2 background
[model,hasEffect,constrRxnNames,deletedGenes] = ...
        deleteModelGenes(model,{'YOR388C','YPL275W','YPL276W'});

fprintf(' \r');
```

```
fprintf('Y5.30 model, modified by Y5_3_three_knockout1 script output,
loaded.\n');
fprintf(' \r');

%%
%We need to restrict the list of genes which can be deleted to about 300.
%
%First, find reactions that are annotated with just 1 gene. Any of these
%reactions can be blocked with a single gene deletion, so combinations of
%two or three are feasible. Then, I'll remove duplicates and essential
%genes, to have a first list to screen for triple deletions.
%
%Anther thing to think about: "and" vs "or" genes for double and triples
%

%Start by finding all reactions annotated with only 1 gene in
%model.rxnGeneMat

number_of_genes=sum(model.rxnGeneMat');
%number_of_genes is a row vector whose entries correspond to the number of
%genes annotating each reaction (Ex: in iND750, there are 1266 reactions.
%number_of_genes(1265) is 1 b/c rxn 1265 is catalyzed by YOR011W.)

single_gene_reaction_genes=model.grRules(number_of_genes == 1);
%single_gene_reaction_genes is a list of genes which catalyze reactions,
%when the reaction is annotated as requiring only a single gene. There are
%834 such genes.

unique_single_gene_reaction_genes=unique(single_gene_reaction_genes);
%There are 478 unique genes which catalyze reactions annotated as only
%requiring a single gene (this is a nice reduction from the 926 genes in
%the model already!). How many of them are essential?

%next, let's remove known essential genes and dubious ORFs from our list.

%
% list of 1191 unique inviable ORFs taken from the Yeast Deletion Project (14
aug
% 11)
% http://www-
sequence.stanford.edu/group/yeast_deletion_project/downloads.html

sgd_essential_genes =
{'YAL001C';'YAL003W';'YAL025C';'YAL032C';'YAL033W';'YAL034W-a';'YAL035C-
A';'YAL038W';'YAL041W';'YAL043C';'YAR007C';'YAR008W';'YAR019C';'YBL004W';'YBL
014C';'YBL018C';'YBL020W';'YBL023C';'YBL026W';'YBL030C';'YBL034C';'YBL035C';'
YBL040C';'YBL041W';'YBL050W';'YBL073W';'YBL074C';'YBL076C';'YBL077W';'YBL084C
';'YBL092W';'YBL097W';'YBL105C';'YBR002C';'YBR004C';'YBR011C';'YBR029C';'YBR0
38W';'YBR049C';'YBR055C';'YBR060C';'YBR070C';'YBR079C';'YBR080C';'YBR087W';'Y
BR087W';'YBR088C';'YBR089W';'YBR091C';'YBR102C';'YBR109C';'YBR110W';'YBR123C'
;'YBR124W';'YBR135W';'YBR136W';'YBR140C';'YBR142W';'YBR143C';'YBR152W';'YBR15
3W';'YBR154C';'YBR155W';'YBR160W';'YBR167C';'YBR190W';'YBR192W';'YBR193C';'YB
R196C';'YBR198C';'YBR202W';'YBR211C';'YBR233W-A';'YBR233W-
A';'YBR234C';'YBR236C';'YBR237W';'YBR243C';'YBR247C';'YBR252W';'YBR253W';'YBR
254C';'YBR256C';'YBR257W';'YBR265W';'YCL003W';'YCL004W';'YCL017C';'YCL031C';'
YCL031C';'YCL031C';'YCL043C';'YCL052C';'YCL053C';'YCL054W';'YCL059C';'YCR012W
';'YCR012W';'YCR013C';'YCR013C';'YCR035C';'YCR052W';'YCR054C';'YCR057C';'YCR0
```

72C';'YCR093W';'YDL003W';'YDL004W';'YDL007W';'YDL008W';'YDL014W';'YDL015C';'Y
DL016C';'YDL017W';'YDL028C';'YDL029W';'YDL030W';'YDL031W';'YDL043C';'YDL045C'
;'YDL055C';'YDL058W';'YDL060W';'YDL064W';'YDL084W';'YDL087C';'YDL092W';'YDL09
7C';'YDL098C';'YDL102W';'YDL103C';'YDL105W';'YDL108W';'YDL111C';'YDL120W';'YD
L126C';'YDL132W';'YDL139C';'YDL140C';'YDL141W';'YDL143W';'YDL145C';'YDL147W';
'YDL148C';'YDL150W';'YDL152W';'YDL153C';'YDL163W';'YDL164C';'YDL165W';'YDL166
C';'YDL193W';'YDL195W';'YDL196W';'YDL205C';'YDL207W';'YDL208W';'YDL209C';'YDL
212W';'YDL217C';'YDL220C';'YDL221W';'YDL235C';'YDR002W';'YDR013W';'YDR016C';'
YDR021W';'YDR023W';'YDR037W';'YDR041W';'YDR044W';'YDR045C';'YDR047W';'YDR050C
';'YDR050C';'YDR052C';'YDR053W';'YDR054C';'YDR060W';'YDR062W';'YDR064W';'YDR0
81C';'YDR082W';'YDR086C';'YDR087C';'YDR088C';'YDR091C';'YDR113C';'YDR118W';'Y
DR141C';'YDR145W';'YDR160W';'YDR164C';'YDR166C';'YDR167W';'YDR168W';'YDR170C'
;'YDR172W';'YDR177W';'YDR180W';'YDR182W';'YDR187C';'YDR188W';'YDR189W';'YDR19
0C';'YDR196C';'YDR201W';'YDR208W';'YDR211W';'YDR212W';'YDR224C';'YDR224C';'YD
R224C';'YDR228C';'YDR232W';'YDR235W';'YDR236C';'YDR238C';'YDR240C';'YDR243C';
'YDR246W';'YDR267C';'YDR280W';'YDR288W';'YDR292C';'YDR299W';'YDR301W';'YDR302
W';'YDR303C';'YDR308C';'YDR311W';'YDR320C-A';'YDR320C-
A';'YDR324C';'YDR325W';'YDR327W';'YDR328C';'YDR331W';'YDR339C';'YDR341C';'YDR
353W';'YDR355C';'YDR356W';'YDR361C';'YDR362C';'YDR365C';'YDR367W';'YDR373W';'
YDR376W';'YDR381W';'YDR390C';'YDR394W';'YDR396W';'YDR397C';'YDR398W';'YDR404C
';'YDR407C';'YDR412W';'YDR413C';'YDR416W';'YDR427W';'YDR427W';'YDR429C';'YDR4
34W';'YDR437W';'YDR449C';'YDR454C';'YDR460W';'YDR464W';'YDR468C';'YDR472W';'Y
DR473C';'YDR478W';'YDR487C';'YDR489W';'YDR498C';'YDR499W';'YDR510W';'YDR526C'
;'YDR527W';'YDR531W';'YEL002C';'YEL019C';'YEL026W';'YEL032W';'YEL034W';'YEL03
5C';'YEL055C';'YEL058W';'YER003C';'YER006W';'YER008C';'YER009W';'YER012W';'YE
R013W';'YER018C';'YER021W';'YER022W';'YER023W';'YER025W';'YER029C';'YER029C';
'YER036C';'YER038C';'YER043C';'YER048W-A';'YER074W-A';'YER074W-
A';'YER082C';'YER093C';'YER094C';'YER104W';'YER112W';'YER125W';'YER126C';'YER
127W';'YER133W';'YER136W';'YER146W';'YER147C';'YER148W';'YER157W';'YER159C';'
YER165W';'YER168C';'YER171W';'YER172C';'YFL002C';'YFL005W';'YFL008W';'YFL009W
';'YFL017C';'YFL018W-A';'YFL022C';'YFL024C';'YFL029C';'YFL035C';'YFL035C-
A';'YFL037W';'YFL038C';'YFL039C';'YFL045C';'YFR002W';'YFR003C';'YFR004W';'YFR
005C';'YFR027W';'YFR028C';'YFR029W';'YFR031C';'YFR037C';'YFR042W';'YFR050C';'
YFR051C';'YFR052W';'YGL001C';'YGL008C';'YGL011C';'YGL018C';'YGL022W';'YGL030W
';'YGL040C';'YGL044C';'YGL044C';'YGL044C';'YGL047W';'YGL048C';'YGL055W';'YGL0
61C';'YGL065C';'YGL068W';'YGL069C';'YGL073W';'YGL074C';'YGL075C';'YGL091C';'Y
GL092W';'YGL093W';'YGL097W';'YGL097W';'YGL098W';'YGL099W';'YGL102C';'YGL103W'
;'YGL111W';'YGL112C';'YGL113W';'YGL116W';'YGL120C';'YGL122C';'YGL123W';'YGL12
8C';'YGL130W';'YGL137W';'YGL142C';'YGL145W';'YGL145W';'YGL145W';'YGL150C';'YG
L155W';'YGL169W';'YGL171W';'YGL172W';'YGL201C';'YGL207W';'YGL225W';'YGL233W';
'YGL238W';'YGL239C';'YGL239C';'YGL245W';'YGL247W';'YGR002C';'YGR005C';'YGR009
C';'YGR013W';'YGR024C';'YGR029W';'YGR029W';'YGR029W';'YGR030C';'YGR046W';'YGR
047C';'YGR048W';'YGR060W';'YGR065C';'YGR073C';'YGR074W';'YGR075C';'YGR082W';'
YGR083C';'YGR090W';'YGR091W';'YGR094W';'YGR095C';'YGR098C';'YGR099W';'YGR103W
';'YGR113W';'YGR114C';'YGR115C';'YGR116W';'YGR119C';'YGR120C';'YGR128C';'YGR1
40W';'YGR145W';'YGR147C';'YGR156W';'YGR158C';'YGR172C';'YGR175C';'YGR179C';'Y
GR185C';'YGR186W';'YGR190C';'YGR191W';'YGR195W';'YGR198W';'YGR211W';'YGR216C'
;'YGR218W';'YGR245C';'YGR246C';'YGR251W';'YGR253C';'YGR264C';'YGR265W';'YGR26
7C';'YGR274C';'YGR277C';'YGR278W';'YGR280C';'YHL015W';'YHR005C-
A';'YHR007C';'YHR019C';'YHR020W';'YHR023W';'YHR024C';'YHR036W';'YHR040W';'YHR
042W';'YHR058C';'YHR062C';'YHR065C';'YHR068W';'YHR069C';'YHR070W';'YHR072W';'
YHR072W-A';'YHR072W-
A';'YHR074W';'YHR083W';'YHR085W';'YHR088W';'YHR088W';'YHR088W';'YHR089C';'YHR
089C';'YHR101C';'YHR102W';'YHR102W';'YHR107C';'YHR118C';'YHR122W';'YHR128W';'
YHR128W';'YHR143W-
A';'YHR148W';'YHR164C';'YHR165C';'YHR165C';'YHR166C';'YHR169W';'YHR169W';'YHR
170W';'YHR172W';'YHR186C';'YHR188C';'YHR188C';'YHR188C';'YHR190W';'YHR196W';'

YHR197W';'YHR197W';'YHR197W';'YHR199C-A';'YHR199C-
A';'YIL003W';'YIL004C';'YIL019W';'YIL021W';'YIL022W';'YIL026C';'YIL031W';'YIL
046W';'YIL048W';'YIL051C';'YIL061C';'YIL062C';'YIL063C';'YIL068C';'YIL075C';'
YIL078W';'YIL083C';'YIL091C';'YIL104C';'YIL106W';'YIL106W';'YIL106W';'YIL109C
';'YIL115C';'YIL118W';'YIL126W';'YIL129C';'YIL142W';'YIL143C';'YIL144W';'YIL1
47C';'YIL150C';'YIL171W';'YIR006C';'YIR008C';'YIR010W';'YIR011C';'YIR012W';'Y
IR015W';'YIR022W';'YJL001W';'YJL002C';'YJL005W';'YJL008C';'YJL008C';'YJL008C'
;'YJL009W';'YJL010C';'YJL011C';'YJL014W';'YJL015C';'YJL018W';'YJL019W';'YJL02
5W';'YJL026W';'YJL031C';'YJL032W';'YJL033W';'YJL034W';'YJL035C';'YJL039C';'YJ
L041W';'YJL050W';'YJL054W';'YJL061W';'YJL069C';'YJL072C';'YJL074C';'YJL076W';
'YJL081C';'YJL085W';'YJL086C';'YJL087C';'YJL090C';'YJL091C';'YJL097W';'YJL104
W';'YJL109C';'YJL111W';'YJL125C';'YJL143W';'YJL156C';'YJL167W';'YJL173C';'YJL
174W';'YJL194W';'YJL195C';'YJL195C';'YJL202C';'YJL202C';'YJL203W';'YJR002W';'
YJR006W';'YJR007W';'YJR012C';'YJR013W';'YJR016C';'YJR017C';'YJR022W';'YJR023C
';'YJR041C';'YJR042W';'YJR045C';'YJR046W';'YJR046W';'YJR046W';'YJR057W';'YJR0
64W';'YJR065C';'YJR067C';'YJR068W';'YJR072C';'YJR076C';'YJR089W';'YJR089W';'Y
JR093C';'YJR112W';'YJR123W';'YJR141W';'YKL004W';'YKL006C-
A';'YKL012W';'YKL013C';'YKL014C';'YKL018W';'YKL019W';'YKL021C';'YKL022C';'YKL
024C';'YKL028W';'YKL033W';'YKL035W';'YKL036C';'YKL042W';'YKL045W';'YKL049C';'
YKL049C';'YKL049C';'YKL052C';'YKL058W';'YKL059C';'YKL060C';'YKL078W';'YKL082C
';'YKL083W';'YKL088W';'YKL089W';'YKL095W';'YKL099C';'YKL104C';'YKL108W';'YKL1
11C';'YKL112W';'YKL122C';'YKL125W';'YKL138C-A';'YKL138C-
A';'YKL141W';'YKL144C';'YKL145W';'YKL152C';'YKL153W';'YKL154W';'YKL165C';'YKL
172W';'YKL172W';'YKL172W';'YKL173W';'YKL180W';'YKL182W';'YKL186C';'YKL189W';'
YKL192C';'YKL193C';'YKL195W';'YKL196C';'YKL203C';'YKL210W';'YKR002W';'YKR004C
';'YKR008W';'YKR022C';'YKR025W';'YKR037C';'YKR038C';'YKR062W';'YKR063C';'YKR0
68C';'YKR071C';'YKR079C';'YKR081C';'YKR083C';'YKR086W';'YLL003W';'YLL004W';'Y
LL008W';'YLL011W';'YLL018C';'YLL031C';'YLL034C';'YLL035W';'YLL036C';'YLL037W'
;'YLL050C';'YLR002C';'YLR005W';'YLR007W';'YLR008C';'YLR009W';'YLR010C';'YLR02
2C';'YLR026C';'YLR029C';'YLR033W';'YLR045C';'YLR051C';'YLR060W';'YLR066W';'YL
R071C';'YLR075W';'YLR076C';'YLR078C';'YLR086W';'YLR088W';'YLR099W-
A';'YLR099W-
A';'YLR100W';'YLR101C';'YLR103C';'YLR105C';'YLR106C';'YLR115W';'YLR116W';'YLR
117C';'YLR127C';'YLR129W';'YLR132C';'YLR140W';'YLR141W';'YLR145W';'YLR147C';'
YLR153C';'YLR163C';'YLR166C';'YLR167W';'YLR175W';'YLR186W';'YLR186W';'YLR195C
';'YLR196W';'YLR197W';'YLR198C';'YLR208W';'YLR212C';'YLR215C';'YLR222C';'YLR2
23C';'YLR229C';'YLR230W';'YLR243W';'YLR249W';'YLR259C';'YLR272C';'YLR274W';'Y
LR275W';'YLR276C';'YLR277C';'YLR291C';'YLR293C';'YLR298C';'YLR305C';'YLR310C'
;'YLR314C';'YLR316C';'YLR316C';'YLR316C';'YLR317W';'YLR321C';'YLR323C';'YLR33
6C';'YLR339C';'YLR340W';'YLR347C';'YLR355C';'YLR359W';'YLR378C';'YLR379W';'YL
R383W';'YLR397C';'YLR409C';'YLR424W';'YLR430W';'YLR438C-
A';'YLR440C';'YLR457C';'YLR458W';'YLR459W';'YML010W';'YML015C';'YML015C';'YML
023C';'YML023C';'YML025C';'YML031W';'YML043C';'YML046W';'YML049C';'YML064C';'
YML065W';'YML069W';'YML077W';'YML085C';'YML091C';'YML092C';'YML092C';'YML092C
';'YML093W';'YML098W';'YML105C';'YML114C';'YML125C';'YML126C';'YML127W';'YML1
30C';'YMR001C';'YMR005W';'YMR005W';'YMR013C';'YMR028W';'YMR033W';'YMR033W';'Y
MR043W';'YMR047C';'YMR047C';'YMR049C';'YMR059W';'YMR059W';'YMR059W';'YMR061W'
;'YMR076C';'YMR079W';'YMR093W';'YMR094W';'YMR108W';'YMR108W';'YMR112C';'YMR11
3W';'YMR117C';'YMR128W';'YMR131C';'YMR134W';'YMR146C';'YMR149W';'YMR168C';'YM
R197C';'YMR200W';'YMR203W';'YMR208W';'YMR211W';'YMR213W';'YMR218C';'YMR220W';
'YMR227C';'YMR229C';'YMR235C';'YMR236W';'YMR239C';'YMR240C';'YMR260C';'YMR268
C';'YMR270C';'YMR277W';'YMR281W';'YMR288W';'YMR290C';'YMR290W-
A';'YMR296C';'YMR298W';'YMR301C';'YMR308C';'YMR309C';'YMR314W';'YNL002C';'YNL
006W';'YNL007C';'YNL024C-A';'YNL024C-
A';'YNL026W';'YNL036W';'YNL036W';'YNL038W';'YNL039W';'YNL061W';'YNL062C';'YNL
075W';'YNL088W';'YNL102W';'YNL103W';'YNL110C';'YNL112W';'YNL112W';'YNL112W';'
YNL113W';'YNL114C';'YNL118C';'YNL124W';'YNL126W';'YNL126W';'YNL131W';'YNL132W

```
';'YNL137C';'YNL138W-A';'YNL138W-
A';'YNL149C';'YNL150W';'YNL151C';'YNL152W';'YNL158W';'YNL161W';'YNL163C';'YNL
172W';'YNL178W';'YNL181W';'YNL182C';'YNL188W';'YNL189W';'YNL207W';'YNL216W';'
YNL221C';'YNL222W';'YNL232W';'YNL240C';'YNL244C';'YNL245C';'YNL247W';'YNL251C
';'YNL256W';'YNL258C';'YNL260C';'YNL261W';'YNL262W';'YNL263C';'YNL267W';'YNL2
72C';'YNL282W';'YNL287W';'YNL290W';'YNL306W';'YNL308C';'YNL310C';'YNL312W';'Y
NL313C';'YNL317W';'YNR003C';'YNR011C';'YNR016C';'YNR017W';'YNR026C';'YNR035C'
;'YNR038W';'YNR043W';'YNR046W';'YNR053C';'YNR054C';'YOL005C';'YOL010W';'YOL02
1C';'YOL022C';'YOL026C';'YOL034W';'YOL038W';'YOL040C';'YOL066C';'YOL069W';'YO
L077C';'YOL078W';'YOL094C';'YOL097C';'YOL102C';'YOL120C';'YOL123W';'YOL127W';
'YOL130W';'YOL133W';'YOL134C';'YOL135C';'YOL139C';'YOL142W';'YOL142W';'YOL144
W';'YOL146W';'YOL149W';'YOR004W';'YOR020C';'YOR046C';'YOR048C';'YOR056C';'YOR
057W';'YOR060C';'YOR063W';'YOR074C';'YOR075W';'YOR077W';'YOR095C';'YOR098C';'
YOR102W';'YOR103C';'YOR110W';'YOR116C';'YOR117W';'YOR119C';'YOR122C';'YOR143C
';'YOR145C';'YOR146W';'YOR148C';'YOR149C';'YOR151C';'YOR157C';'YOR159C';'YOR1
60W';'YOR168W';'YOR169C';'YOR174W';'YOR176W';'YOR181W';'YOR194C';'YOR203W';'Y
OR204W';'YOR206W';'YOR207C';'YOR210W';'YOR217W';'YOR218C';'YOR224C';'YOR232W'
;'YOR236W';'YOR244W';'YOR249C';'YOR250C';'YOR254C';'YOR256C';'YOR257W';'YOR25
9C';'YOR260W';'YOR261C';'YOR262W';'YOR272W';'YOR278W';'YOR281C';'YOR282W';'YO
R287C';'YOR294W';'YOR310C';'YOR319W';'YOR326W';'YOR329C';'YOR335C';'YOR336W';
'YOR340C';'YOR341W';'YOR353C';'YOR361C';'YOR362C';'YOR370C';'YOR372C';'YOR373
W';'YPL007C';'YPL010W';'YPL011C';'YPL012W';'YPL016W';'YPL020C';'YPL028W';'YPL
043W';'YPL044C';'YPL063W';'YPL076W';'YPL082C';'YPL083C';'YPL085W';'YPL093W';'
YPL094C';'YPL117C';'YPL122C';'YPL124W';'YPL126W';'YPL128C';'YPL131W';'YPL142C
';'YPL143W';'YPL146C';'YPL151C';'YPL153C';'YPL160W';'YPL169C';'YPL175W';'YPL1
90C';'YPL204W';'YPL209C';'YPL210C';'YPL211W';'YPL217C';'YPL218W';'YPL228W';'Y
PL231W';'YPL233W';'YPL235W';'YPL237W';'YPL238C';'YPL242C';'YPL243W';'YPL251W'
;'YPL252C';'YPL255W';'YPL266W';'YPR010C';'YPR016C';'YPR019W';'YPR025C';'YPR03
3C';'YPR034W';'YPR035W';'YPR041W';'YPR048W';'YPR055W';'YPR056W';'YPR082C';'YP
R085C';'YPR086W';'YPR088C';'YPR094W';'YPR103W';'YPR104C';'YPR105C';'YPR107C';
'YPR108W';'YPR110C';'YPR112C';'YPR113W';'YPR133C';'YPR136C';'YPR137W';'YPR142
C';'YPR143W';'YPR144C';'YPR161C';'YPR162C';'YPR165W';'YPR168W';'YPR169W';'YPR
175W';'YPR176C';'YPR177C';'YPR178W';'YPR180W';'YPR181C';'YPR182W';'YPR183W';'
YPR186C';'YPR187W';'YPR190C';};

% list of 4941 verified ORFs taken from SGD (14 aug 11)
% http://www.yeastgenome.org/cgi-
bin/search/featureSearch?featuretype=ORF&qualifier=Verified

verified_genes =
{'Q0045';'Q0050';'Q0055';'Q0060';'Q0065';'Q0070';'Q0080';'Q0085';'Q0105';'Q01
10';'Q0115';'Q0120';'Q0130';'Q0140';'Q0160';'Q0250';'Q0275';'R0010W';'R0020C'
;'R0030W';'R0040C';'YAL001C';'YAL002W';'YAL003W';'YAL005C';'YAL007C';'YAL008W
';'YAL009W';'YAL010C';'YAL011W';'YAL012W';'YAL013W';'YAL014C';'YAL015C';'YAL0
16W';'YAL017W';'YAL019W';'YAL020C';'YAL021C';'YAL022C';'YAL023C';'YAL024C';'Y
AL025C';'YAL026C';'YAL027W';'YAL028W';'YAL029C';'YAL030W';'YAL031C';'YAL032C'
;'YAL033W';'YAL034C';'YAL034W-
A';'YAL035W';'YAL036C';'YAL038W';'YAL039C';'YAL040C';'YAL041W';'YAL042W';'YAL
043C';'YAL044C';'YAL046C';'YAL047C';'YAL048C';'YAL049C';'YAL051W';'YAL053W';'
YAL054C';'YAL055W';'YAL056W';'YAL058W';'YAL059W';'YAL060W';'YAL062W';'YAL063C
';'YAL064W';'YAL067C';'YAL068C';'YAR002C-
A';'YAR002W';'YAR003W';'YAR007C';'YAR008W';'YAR014C';'YAR015W';'YAR018C';'YAR
019C';'YAR020C';'YAR027W';'YAR031W';'YAR033W';'YAR035W';'YAR042W';'YAR050W';'
YAR071W';'YBL001C';'YBL002W';'YBL003C';'YBL004W';'YBL005W';'YBL006C';'YBL007C
';'YBL008W';'YBL009W';'YBL011W';'YBL013W';'YBL014C';'YBL015W';'YBL016W';'YBL0
17C';'YBL018C';'YBL019W';'YBL020W';'YBL021C';'YBL022C';'YBL023C';'YBL024W';'Y
BL025W';'YBL026W';'YBL027W';'YBL028C';'YBL030C';'YBL031W';'YBL032W';'YBL033C'
```

;'YBL034C';'YBL035C';'YBL036C';'YBL037W';'YBL038W';'YBL039C';'YBL040C';'YBL04
1W';'YBL042C';'YBL043W';'YBL045C';'YBL046W';'YBL047C';'YBL049W';'YBL050W';'YB
L051C';'YBL052C';'YBL054W';'YBL055C';'YBL056W';'YBL057C';'YBL058W';'YBL059C-
A';'YBL060W';'YBL061C';'YBL063W';'YBL064C';'YBL066C';'YBL067C';'YBL068W';'YBL
069W';'YBL071W-
A';'YBL072C';'YBL074C';'YBL075C';'YBL076C';'YBL078C';'YBL079W';'YBL080C';'YBL
082C';'YBL084C';'YBL085W';'YBL087C';'YBL088C';'YBL089W';'YBL090W';'YBL091C';'
YBL091C-
A';'YBL092W';'YBL093C';'YBL097W';'YBL098W';'YBL099W';'YBL101C';'YBL102W';'YBL
103C';'YBL104C';'YBL105C';'YBL106C';'YBL108C-
A';'YBR001C';'YBR002C';'YBR003W';'YBR004C';'YBR005W';'YBR006W';'YBR008C';'YBR
009C';'YBR010W';'YBR011C';'YBR014C';'YBR015C';'YBR016W';'YBR017C';'YBR018C';'
YBR019C';'YBR020W';'YBR021W';'YBR022W';'YBR023C';'YBR024W';'YBR025C';'YBR026C
';'YBR028C';'YBR029C';'YBR030W';'YBR031W';'YBR034C';'YBR035C';'YBR036C';'YBR0
37C';'YBR038W';'YBR039W';'YBR040W';'YBR041W';'YBR042C';'YBR043C';'YBR044C';'Y
BR045C';'YBR046C';'YBR048W';'YBR049C';'YBR050C';'YBR052C';'YBR054W';'YBR055C'
;'YBR057C';'YBR058C';'YBR058C-
A';'YBR059C';'YBR060C';'YBR061C';'YBR065C';'YBR066C';'YBR067C';'YBR068C';'YBR
069C';'YBR070C';'YBR071W';'YBR072W';'YBR073W';'YBR076W';'YBR077C';'YBR078W';'
YBR079C';'YBR080C';'YBR081C';'YBR082C';'YBR083W';'YBR084C-
A';'YBR084W';'YBR085W';'YBR086C';'YBR087W';'YBR088C';'YBR089C-
A';'YBR091C';'YBR092C';'YBR093C';'YBR094W';'YBR095C';'YBR097W';'YBR098W';'YBR
101C';'YBR102C';'YBR103W';'YBR104W';'YBR105C';'YBR106W';'YBR107C';'YBR108W';'
YBR109C';'YBR110W';'YBR111C';'YBR111W-
A';'YBR112C';'YBR114W';'YBR115C';'YBR117C';'YBR118W';'YBR119W';'YBR120C';'YBR
121C';'YBR122C';'YBR123C';'YBR125C';'YBR126C';'YBR127C';'YBR128C';'YBR129C';'
YBR130C';'YBR131W';'YBR132C';'YBR133C';'YBR135W';'YBR136W';'YBR137W';'YBR139W
';'YBR140C';'YBR142W';'YBR143C';'YBR145W';'YBR146W';'YBR147W';'YBR148W';'YBR1
49W';'YBR150C';'YBR151W';'YBR152W';'YBR153W';'YBR154C';'YBR155W';'YBR156C';'Y
BR157C';'YBR158W';'YBR159W';'YBR160W';'YBR161W';'YBR162C';'YBR162W-
A';'YBR163W';'YBR164C';'YBR165W';'YBR166C';'YBR167C';'YBR168W';'YBR169C';'YBR
170C';'YBR171W';'YBR172C';'YBR173C';'YBR175W';'YBR176W';'YBR177C';'YBR179C';'
YBR180W';'YBR181C';'YBR182C';'YBR183W';'YBR185C';'YBR186W';'YBR188C';'YBR189W
';'YBR191W';'YBR192W';'YBR193C';'YBR194W';'YBR195C';'YBR196C';'YBR198C';'YBR1
99W';'YBR200W';'YBR201W';'YBR202W';'YBR203W';'YBR204C';'YBR205W';'YBR207W';'Y
BR208C';'YBR210W';'YBR211C';'YBR212W';'YBR213W';'YBR214W';'YBR215W';'YBR216C'
;'YBR217W';'YBR218C';'YBR221C';'YBR222C';'YBR223C';'YBR227C';'YBR228W';'YBR22
9C';'YBR230C';'YBR231C';'YBR233W';'YBR233W-
A';'YBR234C';'YBR236C';'YBR237W';'YBR238C';'YBR240C';'YBR243C';'YBR244W';'YBR
245C';'YBR247C';'YBR248C';'YBR249C';'YBR250W';'YBR251W';'YBR252W';'YBR253W';'
YBR254C';'YBR255W';'YBR256C';'YBR257W';'YBR258C';'YBR260C';'YBR261C';'YBR262C
';'YBR263W';'YBR264C';'YBR265W';'YBR267W';'YBR268W';'YBR272C';'YBR273C';'YBR2
74W';'YBR275C';'YBR276C';'YBR278W';'YBR279W';'YBR280C';'YBR281C';'YBR282W';'Y
BR283C';'YBR286W';'YBR288C';'YBR289W';'YBR290W';'YBR291C';'YBR293W';'YBR294W'
;'YBR295W';'YBR296C';'YBR297W';'YBR298C';'YBR299W';'YBR301W';'YBR302C';'YCL00
1W';'YCL004W';'YCL005W';'YCL005W-
A';'YCL008C';'YCL009C';'YCL010C';'YCL011C';'YCL012C';'YCL014W';'YCL016C';'YCL
017C';'YCL018W';'YCL024W';'YCL025C';'YCL026C-
A';'YCL027W';'YCL028W';'YCL029C';'YCL030C';'YCL031C';'YCL032W';'YCL033C';'YCL
034W';'YCL035C';'YCL036W';'YCL037C';'YCL038C';'YCL039W';'YCL040W';'YCL043C';'
YCL044C';'YCL045C';'YCL047C';'YCL048W';'YCL050C';'YCL051W';'YCL052C';'YCL054W
';'YCL055W';'YCL056C';'YCL057W';'YCL058C';'YCL058W-
A';'YCL059C';'YCL061C';'YCL063W';'YCL064C';'YCL066W';'YCL067C';'YCL069W';'YCL
073C';'YCR002C';'YCR003W';'YCR004C';'YCR005C';'YCR008W';'YCR009C';'YCR010C';'
YCR011C';'YCR012W';'YCR014C';'YCR017C';'YCR018C';'YCR019W';'YCR020C';'YCR020C
-A';'YCR020W-B';'YCR021C';'YCR023C';'YCR024C';'YCR024C-
A';'YCR026C';'YCR027C';'YCR028C';'YCR028C-

A';'YCR030C';'YCR031C';'YCR032W';'YCR033W';'YCR034W';'YCR035C';'YCR036W';'YCR
037C';'YCR038C';'YCR039C';'YCR040W';'YCR042C';'YCR044C';'YCR045C';'YCR046C';'
YCR047C';'YCR048W';'YCR052W';'YCR053W';'YCR054C';'YCR057C';'YCR059C';'YCR060W
';'YCR063W';'YCR065W';'YCR066W';'YCR067C';'YCR068W';'YCR069W';'YCR071C';'YCR0
72C';'YCR073C';'YCR073W-
A';'YCR075C';'YCR077C';'YCR079W';'YCR081W';'YCR082W';'YCR083W';'YCR084C';'YCR
086W';'YCR088W';'YCR089W';'YCR091W';'YCR092C';'YCR093W';'YCR094W';'YCR096C';'
YCR097W';'YCR098C';'YCR104W';'YCR105W';'YCR106W';'YCR107W';'YDL001W';'YDL002C
';'YDL003W';'YDL004W';'YDL005C';'YDL006W';'YDL007W';'YDL008W';'YDL010W';'YDL0
12C';'YDL013W';'YDL014W';'YDL015C';'YDL017W';'YDL018C';'YDL019C';'YDL020C';'Y
DL021W';'YDL022W';'YDL024C';'YDL025C';'YDL028C';'YDL029W';'YDL030W';'YDL031W'
;'YDL033C';'YDL035C';'YDL036C';'YDL037C';'YDL039C';'YDL040C';'YDL042C';'YDL04
3C';'YDL044C';'YDL045C';'YDL045W-
A';'YDL046W';'YDL047W';'YDL048C';'YDL049C';'YDL051W';'YDL052C';'YDL053C';'YDL
054C';'YDL055C';'YDL056W';'YDL058W';'YDL059C';'YDL060W';'YDL061C';'YDL063C';'
YDL064W';'YDL065C';'YDL066W';'YDL067C';'YDL069C';'YDL070W';'YDL072C';'YDL074C
';'YDL075W';'YDL076C';'YDL077C';'YDL078C';'YDL079C';'YDL080C';'YDL081C';'YDL0
82W';'YDL083C';'YDL084W';'YDL085W';'YDL087C';'YDL088C';'YDL089W';'YDL090C';'Y
DL091C';'YDL092W';'YDL093W';'YDL095W';'YDL097C';'YDL098C';'YDL099W';'YDL100C'
;'YDL101C';'YDL102W';'YDL103C';'YDL104C';'YDL105W';'YDL106C';'YDL107W';'YDL10
8W';'YDL110C';'YDL111C';'YDL112W';'YDL113C';'YDL115C';'YDL116W';'YDL117W';'YD
L120W';'YDL122W';'YDL123W';'YDL124W';'YDL125C';'YDL126C';'YDL127W';'YDL128W';
'YDL130W';'YDL130W-A';'YDL131W';'YDL132W';'YDL133C-
A';'YDL133W';'YDL134C';'YDL135C';'YDL136W';'YDL137W';'YDL138W';'YDL139C';'YDL
140C';'YDL141W';'YDL142C';'YDL143W';'YDL145C';'YDL146W';'YDL147W';'YDL148C';'
YDL149W';'YDL150W';'YDL153C';'YDL154W';'YDL155W';'YDL159W';'YDL160C';'YDL161W
';'YDL164C';'YDL165W';'YDL166C';'YDL167C';'YDL168W';'YDL169C';'YDL170W';'YDL1
71C';'YDL173W';'YDL174C';'YDL175C';'YDL176W';'YDL178W';'YDL179W';'YDL181W';'Y
DL182W';'YDL183C';'YDL184C';'YDL185W';'YDL188C';'YDL189W';'YDL190C';'YDL191W'
;'YDL192W';'YDL193W';'YDL194W';'YDL195W';'YDL197C';'YDL198C';'YDL200C';'YDL20
1W';'YDL202W';'YDL203C';'YDL204W';'YDL205C';'YDL207W';'YDL208W';'YDL209C';'YD
L210W';'YDL212W';'YDL213C';'YDL214C';'YDL215C';'YDL216C';'YDL217C';'YDL219W';
'YDL220C';'YDL222C';'YDL223C';'YDL224C';'YDL225W';'YDL226C';'YDL227C';'YDL229
W';'YDL230W';'YDL231C';'YDL232W';'YDL234C';'YDL235C';'YDL236W';'YDL237W';'YDL
238C';'YDL239C';'YDL240W';'YDL243C';'YDL244W';'YDL245C';'YDL247W';'YDL248W';'
YDR001C';'YDR002W';'YDR003W';'YDR004W';'YDR005C';'YDR006C';'YDR007W';'YDR009W
';'YDR011W';'YDR012W';'YDR013W';'YDR014W';'YDR014W-
A';'YDR016C';'YDR017C';'YDR019C';'YDR021W';'YDR022C';'YDR023W';'YDR025W';'YDR
026C';'YDR027C';'YDR028C';'YDR030C';'YDR031W';'YDR032C';'YDR033W';'YDR034C';'
YDR035W';'YDR036C';'YDR037W';'YDR038C';'YDR039C';'YDR040C';'YDR041W';'YDR043C
';'YDR044W';'YDR045C';'YDR046C';'YDR047W';'YDR049W';'YDR050C';'YDR051C';'YDR0
52C';'YDR054C';'YDR055W';'YDR057W';'YDR058C';'YDR059C';'YDR060W';'YDR062W';'Y
DR063W';'YDR064W';'YDR065W';'YDR068W';'YDR069C';'YDR071C';'YDR072C';'YDR073W'
;'YDR074W';'YDR075W';'YDR076W';'YDR077W';'YDR078C';'YDR079C-
A';'YDR079W';'YDR080W';'YDR081C';'YDR082W';'YDR083W';'YDR084C';'YDR085C';'YDR
086C';'YDR087C';'YDR088C';'YDR091C';'YDR092W';'YDR093W';'YDR096W';'YDR097C';'
YDR098C';'YDR099W';'YDR100W';'YDR101C';'YDR103W';'YDR104C';'YDR105C';'YDR106W
';'YDR107C';'YDR108W';'YDR110W';'YDR113C';'YDR116C';'YDR117C';'YDR118W';'YDR1
20C';'YDR121W';'YDR122W';'YDR123C';'YDR125C';'YDR126W';'YDR127W';'YDR128W';'Y
DR129C';'YDR130C';'YDR135C';'YDR137W';'YDR138W';'YDR139C';'YDR140W';'YDR141C'
;'YDR142C';'YDR143C';'YDR144C';'YDR145W';'YDR146C';'YDR147W';'YDR148C';'YDR15
0W';'YDR151C';'YDR152W';'YDR153C';'YDR155C';'YDR156W';'YDR158W';'YDR159W';'YD
R160W';'YDR162C';'YDR163W';'YDR164C';'YDR165W';'YDR166C';'YDR167W';'YDR168W';
'YDR169C';'YDR170C';'YDR171W';'YDR172W';'YDR173C';'YDR174W';'YDR175C';'YDR176
W';'YDR177W';'YDR178W';'YDR179C';'YDR180W';'YDR181C';'YDR182W';'YDR183W';'YDR
184C';'YDR185C';'YDR186C';'YDR188W';'YDR189W';'YDR190C';'YDR191W';'YDR192C';'
YDR194C';'YDR195W';'YDR196C';'YDR197W';'YDR198C';'YDR200C';'YDR201W';'YDR202C

';'YDR204W';'YDR205W';'YDR206W';'YDR207C';'YDR208W';'YDR211W';'YDR212W';'YDR2
13W';'YDR214W';'YDR216W';'YDR217C';'YDR218C';'YDR219C';'YDR221W';'YDR223W';'Y
DR224C';'YDR225W';'YDR226W';'YDR227W';'YDR228C';'YDR229W';'YDR231C';'YDR232W'
;'YDR233C';'YDR234W';'YDR235W';'YDR236C';'YDR237W';'YDR238C';'YDR239C';'YDR24
0C';'YDR242W';'YDR243C';'YDR244W';'YDR245W';'YDR246W';'YDR247W';'YDR251W';'YD
R252W';'YDR253C';'YDR254W';'YDR255C';'YDR256C';'YDR257C';'YDR258C';'YDR259C';
'YDR260C';'YDR261C';'YDR263C';'YDR264C';'YDR265W';'YDR266C';'YDR267C';'YDR268
W';'YDR270W';'YDR272W';'YDR273W';'YDR275W';'YDR276C';'YDR277C';'YDR279W';'YDR
280W';'YDR281C';'YDR283C';'YDR284C';'YDR285W';'YDR287W';'YDR288W';'YDR289C';'
YDR292C';'YDR293C';'YDR294C';'YDR295C';'YDR296W';'YDR297W';'YDR298C';'YDR299W
';'YDR300C';'YDR301W';'YDR302W';'YDR303C';'YDR304C';'YDR305C';'YDR308C';'YDR3
09C';'YDR310C';'YDR311W';'YDR312W';'YDR313C';'YDR314C';'YDR315C';'YDR316W';'Y
DR317W';'YDR318W';'YDR320C';'YDR320C-A';'YDR321W';'YDR322C-
A';'YDR322W';'YDR323C';'YDR324C';'YDR325W';'YDR326C';'YDR328C';'YDR329C';'YDR
330W';'YDR331W';'YDR332W';'YDR334W';'YDR335W';'YDR337W';'YDR339C';'YDR341C';'
YDR342C';'YDR343C';'YDR345C';'YDR346C';'YDR347W';'YDR349C';'YDR350C';'YDR351W
';'YDR353W';'YDR354W';'YDR356W';'YDR358W';'YDR359C';'YDR361C';'YDR362C';'YDR3
63W';'YDR363W-
A';'YDR364C';'YDR365C';'YDR367W';'YDR368W';'YDR369C';'YDR372C';'YDR373W';'YDR
375C';'YDR376W';'YDR377W';'YDR378C';'YDR379C-A';'YDR379W';'YDR380W';'YDR381C-
A';'YDR381W';'YDR382W';'YDR383C';'YDR384C';'YDR385W';'YDR386W';'YDR388W';'YDR
389W';'YDR390C';'YDR392W';'YDR393W';'YDR394W';'YDR395W';'YDR397C';'YDR398W';'
YDR399W';'YDR400W';'YDR402C';'YDR403W';'YDR404C';'YDR405W';'YDR406W';'YDR407C
';'YDR408C';'YDR409W';'YDR410C';'YDR411C';'YDR412W';'YDR414C';'YDR416W';'YDR4
18W';'YDR419W';'YDR420W';'YDR421W';'YDR422C';'YDR423C';'YDR424C';'YDR425W';'Y
DR427W';'YDR428C';'YDR429C';'YDR430C';'YDR432W';'YDR434W';'YDR435C';'YDR436W'
;'YDR437W';'YDR438W';'YDR439W';'YDR440W';'YDR441C';'YDR443C';'YDR446W';'YDR44
7C';'YDR448W';'YDR449C';'YDR450W';'YDR451C';'YDR452W';'YDR453C';'YDR454C';'YD
R456W';'YDR457W';'YDR458C';'YDR459C';'YDR460W';'YDR461W';'YDR462W';'YDR463W';
'YDR464W';'YDR465C';'YDR466W';'YDR468C';'YDR469W';'YDR470C';'YDR471W';'YDR472
W';'YDR473C';'YDR475C';'YDR477W';'YDR478W';'YDR479C';'YDR480W';'YDR481C';'YDR
482C';'YDR483W';'YDR484W';'YDR485C';'YDR486C';'YDR487C';'YDR488C';'YDR489W';'
YDR490C';'YDR492W';'YDR493W';'YDR494W';'YDR495C';'YDR496C';'YDR497C';'YDR498C
';'YDR499W';'YDR500C';'YDR501W';'YDR502C';'YDR503C';'YDR504C';'YDR505C';'YDR5
07C';'YDR508C';'YDR510W';'YDR511W';'YDR512C';'YDR513W';'YDR514C';'YDR515W';'Y
DR516C';'YDR517W';'YDR518W';'YDR519W';'YDR522C';'YDR523C';'YDR524C';'YDR525W-
A';'YDR527W';'YDR528W';'YDR529C';'YDR530C';'YDR531W';'YDR532C';'YDR533C';'YDR
534C';'YDR536W';'YDR538W';'YDR539W';'YDR540C';'YDR542W';'YDR545W';'YEL001C';'
YEL002C';'YEL003W';'YEL004W';'YEL005C';'YEL006W';'YEL009C';'YEL011W';'YEL012W
';'YEL013W';'YEL015W';'YEL016C';'YEL017C-
A';'YEL017W';'YEL018W';'YEL019C';'YEL020W-
A';'YEL021W';'YEL022W';'YEL024W';'YEL026W';'YEL027W';'YEL029C';'YEL030W';'YEL
031W';'YEL032W';'YEL034W';'YEL036C';'YEL037C';'YEL038W';'YEL039C';'YEL040W';'
YEL041W';'YEL042W';'YEL043W';'YEL044W';'YEL046C';'YEL047C';'YEL048C';'YEL049W
';'YEL050C';'YEL051W';'YEL052W';'YEL053C';'YEL054C';'YEL055C';'YEL056W';'YEL0
58W';'YEL059C-
A';'YEL060C';'YEL061C';'YEL062W';'YEL063C';'YEL064C';'YEL065W';'YEL066W';'YEL
069C';'YEL071W';'YEL072W';'YER001W';'YER002W';'YER003C';'YER004W';'YER005W';'
YER006W';'YER007C-
A';'YER007W';'YER008C';'YER009W';'YER010C';'YER011W';'YER012W';'YER013W';'YER
014C-A';'YER014W';'YER015W';'YER016W';'YER017C';'YER018C';'YER019C-
A';'YER019W';'YER020W';'YER021W';'YER022W';'YER023W';'YER024W';'YER025W';'YER
026C';'YER027C';'YER028C';'YER029C';'YER030W';'YER031C';'YER032W';'YER033C';'
YER035W';'YER036C';'YER037W';'YER038C';'YER039C';'YER040W';'YER041W';'YER042W
';'YER043C';'YER044C';'YER044C-
A';'YER045C';'YER046W';'YER047C';'YER048C';'YER048W-
A';'YER049W';'YER050C';'YER051W';'YER052C';'YER053C';'YER054C';'YER055C';'YER

056C';'YER056C-A';'YER057C';'YER058W';'YER059W';'YER060W';'YER060W-
A';'YER061C';'YER062C';'YER063W';'YER065C';'YER067W';'YER068W';'YER069W';'YER
070W';'YER072W';'YER073W';'YER074W';'YER074W-
A';'YER075C';'YER078C';'YER080W';'YER081W';'YER082C';'YER083C';'YER086W';'YER
087C-
B';'YER087W';'YER088C';'YER089C';'YER090W';'YER091C';'YER092W';'YER093C';'YER
093C-
A';'YER094C';'YER095W';'YER096W';'YER098W';'YER099C';'YER100W';'YER101C';'YER
102W';'YER103W';'YER104W';'YER105C';'YER106W';'YER107C';'YER109C';'YER110C';'
YER111C';'YER112W';'YER113C';'YER114C';'YER115C';'YER116C';'YER117W';'YER118C
';'YER119C';'YER120W';'YER122C';'YER123W';'YER124C';'YER125W';'YER126C';'YER1
27W';'YER128W';'YER129W';'YER131W';'YER132C';'YER133W';'YER134C';'YER136W';'Y
ER139C';'YER141W';'YER142C';'YER143W';'YER144C';'YER145C';'YER146W';'YER147C'
;'YER148W';'YER149C';'YER150W';'YER151C';'YER152C';'YER153C';'YER154W';'YER15
5C';'YER157W';'YER159C';'YER161C';'YER162C';'YER164W';'YER165W';'YER166W';'YE
R167W';'YER168C';'YER169W';'YER170W';'YER171W';'YER172C';'YER173W';'YER174C';
'YER175C';'YER176W';'YER177W';'YER178W';'YER179W';'YER180C';'YER180C-
A';'YER183C';'YER185W';'YER190W';'YFL001W';'YFL002C';'YFL003C';'YFL004W';'YFL
005W';'YFL007W';'YFL008W';'YFL009W';'YFL010C';'YFL010W-
A';'YFL011W';'YFL013C';'YFL014W';'YFL016C';'YFL017C';'YFL017W-
A';'YFL018C';'YFL020C';'YFL021W';'YFL022C';'YFL023W';'YFL024C';'YFL025C';'YFL
026W';'YFL027C';'YFL028C';'YFL029C';'YFL030W';'YFL031W';'YFL033C';'YFL034C-
A';'YFL034C-
B';'YFL036W';'YFL037W';'YFL038C';'YFL039C';'YFL041W';'YFL044C';'YFL045C';'YFL
047W';'YFL048C';'YFL049W';'YFL050C';'YFL053W';'YFL055W';'YFL056C';'YFL057C';'
YFL058W';'YFL059W';'YFL060C';'YFL062W';'YFR001W';'YFR002W';'YFR003C';'YFR004W
';'YFR005C';'YFR007W';'YFR008W';'YFR009W';'YFR010W';'YFR011C';'YFR013W';'YFR0
14C';'YFR015C';'YFR016C';'YFR017C';'YFR019W';'YFR021W';'YFR022W';'YFR023W';'Y
FR024C-
A';'YFR025C';'YFR026C';'YFR027W';'YFR028C';'YFR029W';'YFR030W';'YFR031C';'YFR
031C-A';'YFR032C-
A';'YFR033C';'YFR034C';'YFR036W';'YFR037C';'YFR038W';'YFR040W';'YFR041C';'YFR
042W';'YFR043C';'YFR044C';'YFR046C';'YFR047C';'YFR048W';'YFR049W';'YFR050C';'
YFR051C';'YFR052W';'YFR053C';'YGL001C';'YGL002W';'YGL003C';'YGL004C';'YGL005C
';'YGL006W';'YGL008C';'YGL009C';'YGL011C';'YGL012W';'YGL013C';'YGL014W';'YGL0
16W';'YGL017W';'YGL018C';'YGL019W';'YGL020C';'YGL021W';'YGL022W';'YGL023C';'Y
GL025C';'YGL026C';'YGL027C';'YGL028C';'YGL029W';'YGL030W';'YGL031C';'YGL032C'
;'YGL033W';'YGL035C';'YGL037C';'YGL038C';'YGL039W';'YGL040C';'YGL043W';'YGL04
4C';'YGL045W';'YGL047W';'YGL048C';'YGL049C';'YGL050W';'YGL051W';'YGL053W';'YG
L054C';'YGL055W';'YGL056C';'YGL057C';'YGL058W';'YGL059W';'YGL060W';'YGL061C';
'YGL062W';'YGL063W';'YGL064C';'YGL065C';'YGL066W';'YGL067W';'YGL068W';'YGL070
C';'YGL071W';'YGL073W';'YGL075C';'YGL076C';'YGL077C';'YGL078C';'YGL083W';'YGL
084C';'YGL086W';'YGL087C';'YGL089C';'YGL090W';'YGL091C';'YGL092W';'YGL093W';'
YGL094C';'YGL095C';'YGL096W';'YGL097W';'YGL098W';'YGL099W';'YGL100W';'YGL103W
';'YGL104C';'YGL105W';'YGL106W';'YGL107C';'YGL110C';'YGL111W';'YGL112C';'YGL1
13W';'YGL115W';'YGL116W';'YGL119W';'YGL120C';'YGL121C';'YGL122C';'YGL123W';'Y
GL124C';'YGL125W';'YGL126W';'YGL127C';'YGL128C';'YGL129C';'YGL130W';'YGL131C'
;'YGL133W';'YGL134W';'YGL135W';'YGL136C';'YGL137W';'YGL139W';'YGL141W';'YGL14
2C';'YGL143C';'YGL144C';'YGL145W';'YGL147C';'YGL148W';'YGL150C';'YGL151W';'YG
L153W';'YGL154C';'YGL155W';'YGL156W';'YGL157W';'YGL158W';'YGL160W';'YGL161C';
'YGL162W';'YGL163C';'YGL164C';'YGL166W';'YGL167C';'YGL168W';'YGL169W';'YGL170
C';'YGL171W';'YGL172W';'YGL173C';'YGL174W';'YGL175C';'YGL178W';'YGL179C';'YGL
180W';'YGL181W';'YGL183C';'YGL184C';'YGL186C';'YGL187C';'YGL189C';'YGL190C';'
YGL191W';'YGL192W';'YGL194C';'YGL195W';'YGL196W';'YGL197W';'YGL198W';'YGL200C
';'YGL201C';'YGL202W';'YGL203C';'YGL205W';'YGL206C';'YGL207W';'YGL208W';'YGL2
09W';'YGL210W';'YGL211W';'YGL212W';'YGL213C';'YGL215W';'YGL216W';'YGL219C';'Y
GL220W';'YGL221C';'YGL222C';'YGL223C';'YGL224C';'YGL225W';'YGL226C-

A';'YGL226W';'YGL227W';'YGL228W';'YGL229C';'YGL231C';'YGL232W';'YGL233W';'YGL
234W';'YGL236C';'YGL237C';'YGL238W';'YGL240W';'YGL241W';'YGL243W';'YGL244W';'
YGL245W';'YGL246C';'YGL247W';'YGL248W';'YGL249W';'YGL250W';'YGL251C';'YGL252C
';'YGL253W';'YGL254W';'YGL255W';'YGL256W';'YGL257C';'YGL258W';'YGL263W';'YGR0
02C';'YGR003W';'YGR004W';'YGR005C';'YGR006W';'YGR007W';'YGR008C';'YGR009C';'Y
GR010W';'YGR012W';'YGR013W';'YGR014W';'YGR019W';'YGR020C';'YGR023W';'YGR024C'
;'YGR027C';'YGR028W';'YGR029W';'YGR030C';'YGR031C-
A';'YGR031W';'YGR032W';'YGR033C';'YGR034W';'YGR036C';'YGR037C';'YGR038W';'YGR
040W';'YGR041W';'YGR043C';'YGR044C';'YGR046W';'YGR047C';'YGR048W';'YGR049W';'
YGR054W';'YGR055W';'YGR056W';'YGR057C';'YGR058W';'YGR059W';'YGR060W';'YGR061C
';'YGR062C';'YGR063C';'YGR065C';'YGR068C';'YGR070W';'YGR072W';'YGR074W';'YGR0
75C';'YGR076C';'YGR077C';'YGR078C';'YGR080W';'YGR081C';'YGR082W';'YGR083C';'Y
GR084C';'YGR085C';'YGR086C';'YGR087C';'YGR088W';'YGR089W';'YGR090W';'YGR091W'
;'YGR092W';'YGR094W';'YGR095C';'YGR096W';'YGR097W';'YGR098C';'YGR099W';'YGR10
0W';'YGR101W';'YGR102C';'YGR103W';'YGR104C';'YGR105W';'YGR106C';'YGR108W';'YG
R109C';'YGR110W';'YGR112W';'YGR113W';'YGR116W';'YGR118W';'YGR119C';'YGR120C';
'YGR121C';'YGR122W';'YGR123C';'YGR124W';'YGR128C';'YGR129W';'YGR130C';'YGR131
W';'YGR132C';'YGR133W';'YGR134W';'YGR135W';'YGR136W';'YGR138C';'YGR140W';'YGR
141W';'YGR142W';'YGR143W';'YGR144W';'YGR145W';'YGR146C';'YGR147C';'YGR148C';'
YGR150C';'YGR152C';'YGR154C';'YGR155W';'YGR156W';'YGR157W';'YGR158C';'YGR159C
';'YGR162W';'YGR163W';'YGR165W';'YGR166W';'YGR167W';'YGR169C';'YGR170W';'YGR1
71C';'YGR172C';'YGR173W';'YGR174C';'YGR175C';'YGR177C';'YGR178C';'YGR179C';'Y
GR180C';'YGR181W';'YGR183C';'YGR184C';'YGR185C';'YGR186W';'YGR187C';'YGR188C'
;'YGR189C';'YGR191W';'YGR192C';'YGR193C';'YGR194C';'YGR195W';'YGR196C';'YGR19
7C';'YGR198W';'YGR199W';'YGR200C';'YGR202C';'YGR203W';'YGR204W';'YGR205W';'YG
R206W';'YGR207C';'YGR208W';'YGR209C';'YGR211W';'YGR212W';'YGR213C';'YGR214W';
'YGR215W';'YGR216C';'YGR217W';'YGR218W';'YGR220C';'YGR221C';'YGR222W';'YGR223
C';'YGR224W';'YGR225W';'YGR227W';'YGR229C';'YGR230W';'YGR231C';'YGR232W';'YGR
233C';'YGR234W';'YGR236C';'YGR238C';'YGR239C';'YGR240C';'YGR241C';'YGR244C';'
YGR245C';'YGR246C';'YGR247W';'YGR248W';'YGR249W';'YGR250C';'YGR251W';'YGR252W
';'YGR253C';'YGR254W';'YGR255C';'YGR256W';'YGR257C';'YGR258C';'YGR260W';'YGR2
61C';'YGR262C';'YGR263C';'YGR264C';'YGR266W';'YGR267C';'YGR268C';'YGR270W';'Y
GR271C-
A';'YGR271W';'YGR274C';'YGR275W';'YGR276C';'YGR277C';'YGR278W';'YGR279C';'YGR
280C';'YGR281W';'YGR282C';'YGR283C';'YGR284C';'YGR285C';'YGR286C';'YGR287C';'
YGR288W';'YGR289C';'YGR292W';'YGR294W';'YGR295C';'YGR296W';'YHL001W';'YHL002W
';'YHL003C';'YHL004W';'YHL006C';'YHL007C';'YHL009C';'YHL010C';'YHL011C';'YHL0
13C';'YHL014C';'YHL015W';'YHL016C';'YHL019C';'YHL020C';'YHL021C';'YHL022C';'Y
HL023C';'YHL024W';'YHL025W';'YHL027W';'YHL028W';'YHL030W';'YHL031C';'YHL032C'
;'YHL033C';'YHL034C';'YHL035C';'YHL036W';'YHL038C';'YHL039W';'YHL040C';'YHL04
3W';'YHL046C';'YHL047C';'YHL048W';'YHR001W';'YHR001W-
A';'YHR002W';'YHR003C';'YHR004C';'YHR005C';'YHR005C-
A';'YHR006W';'YHR007C';'YHR008C';'YHR010W';'YHR011W';'YHR012W';'YHR013C';'YHR
014W';'YHR015W';'YHR016C';'YHR017W';'YHR018C';'YHR019C';'YHR020W';'YHR021C';'
YHR023W';'YHR024C';'YHR025W';'YHR026W';'YHR027C';'YHR028C';'YHR029C';'YHR030C
';'YHR031C';'YHR032W';'YHR034C';'YHR036W';'YHR037W';'YHR038W';'YHR039C';'YHR0
39C-
A';'YHR040W';'YHR041C';'YHR042W';'YHR043C';'YHR044C';'YHR046C';'YHR047C';'YHR
049W';'YHR050W';'YHR051W';'YHR052W';'YHR053C';'YHR055C';'YHR056C';'YHR057C';'
YHR058C';'YHR059W';'YHR060W';'YHR061C';'YHR062C';'YHR063C';'YHR064C';'YHR065C
';'YHR066W';'YHR067W';'YHR068W';'YHR069C';'YHR070W';'YHR071W';'YHR072W';'YHR0
72W-A';'YHR073W';'YHR074W';'YHR075C';'YHR076W';'YHR077C';'YHR079C';'YHR079C-
A';'YHR080C';'YHR081W';'YHR082C';'YHR083W';'YHR084W';'YHR085W';'YHR086W';'YHR
087W';'YHR088W';'YHR089C';'YHR090C';'YHR091C';'YHR092C';'YHR094C';'YHR096C';'
YHR098C';'YHR099W';'YHR100C';'YHR101C';'YHR102W';'YHR103W';'YHR104W';'YHR105W
';'YHR106W';'YHR107C';'YHR108W';'YHR109W';'YHR110W';'YHR111W';'YHR112C';'YHR1
13W';'YHR114W';'YHR115C';'YHR116W';'YHR117W';'YHR118C';'YHR119W';'YHR120W';'Y

HR121W';'YHR122W';'YHR123W';'YHR124W';'YHR127W';'YHR128W';'YHR129C';'YHR132C'
;'YHR132W-
A';'YHR133C';'YHR134W';'YHR135C';'YHR136C';'YHR137W';'YHR139C';'YHR141C';'YHR
142W';'YHR143W';'YHR143W-
A';'YHR144C';'YHR146W';'YHR147C';'YHR148W';'YHR149C';'YHR150W';'YHR151C';'YHR
152W';'YHR153C';'YHR154W';'YHR155W';'YHR156C';'YHR157W';'YHR158C';'YHR160C';'
YHR161C';'YHR163W';'YHR164C';'YHR165C';'YHR166C';'YHR167W';'YHR168W';'YHR169W
';'YHR170W';'YHR171W';'YHR172W';'YHR174W';'YHR175W';'YHR176W';'YHR178W';'YHR1
79W';'YHR181W';'YHR183W';'YHR184W';'YHR185C';'YHR186C';'YHR187W';'YHR188C';'Y
HR189W';'YHR190W';'YHR191C';'YHR193C';'YHR194W';'YHR195W';'YHR196W';'YHR197W'
;'YHR198C';'YHR199C';'YHR199C-
A';'YHR200W';'YHR201C';'YHR203C';'YHR204W';'YHR205W';'YHR206W';'YHR207C';'YHR
208W';'YHR209W';'YHR211W';'YHR215W';'YHR216W';'YIL002C';'YIL003W';'YIL004C';'
YIL005W';'YIL006W';'YIL007C';'YIL008W';'YIL009C-
A';'YIL009W';'YIL010W';'YIL011W';'YIL013C';'YIL014W';'YIL015W';'YIL016W';'YIL
017C';'YIL018W';'YIL019W';'YIL020C';'YIL021W';'YIL022W';'YIL023C';'YIL026C';'
YIL027C';'YIL030C';'YIL031W';'YIL033C';'YIL034C';'YIL035C';'YIL036W';'YIL037C
';'YIL038C';'YIL039W';'YIL040W';'YIL041W';'YIL042C';'YIL043C';'YIL044C';'YIL0
45W';'YIL046W';'YIL047C';'YIL048W';'YIL049W';'YIL050W';'YIL051C';'YIL052C';'Y
IL053W';'YIL056W';'YIL057C';'YIL061C';'YIL062C';'YIL063C';'YIL064W';'YIL065C'
;'YIL066C';'YIL068C';'YIL069C';'YIL070C';'YIL071C';'YIL072W';'YIL073C';'YIL07
4C';'YIL075C';'YIL076W';'YIL078W';'YIL079C';'YIL083C';'YIL084C';'YIL085C';'YI
L087C';'YIL088C';'YIL089W';'YIL090W';'YIL091C';'YIL093C';'YIL094C';'YIL095W';
'YIL097W';'YIL098C';'YIL099W';'YIL101C';'YIL103W';'YIL104C';'YIL105C';'YIL106
W';'YIL107C';'YIL109C';'YIL110W';'YIL111W';'YIL112W';'YIL113W';'YIL114C';'YIL
115C';'YIL116W';'YIL117C';'YIL118W';'YIL119C';'YIL120W';'YIL121W';'YIL122W';'
YIL123W';'YIL124W';'YIL125W';'YIL126W';'YIL128W';'YIL129C';'YIL130W';'YIL131C
';'YIL132C';'YIL133C';'YIL134W';'YIL135C';'YIL136W';'YIL137C';'YIL138C';'YIL1
39C';'YIL140W';'YIL142W';'YIL143C';'YIL144W';'YIL145C';'YIL146C';'YIL147C';'Y
IL148W';'YIL149C';'YIL150C';'YIL153W';'YIL154C';'YIL155C';'YIL156W';'YIL157C'
;'YIL158W';'YIL159W';'YIL160C';'YIL162W';'YIL164C';'YIL172C';'YIL173W';'YIR00
1C';'YIR002C';'YIR003W';'YIR004W';'YIR005W';'YIR006C';'YIR008C';'YIR009W';'YI
R010W';'YIR011C';'YIR012W';'YIR013C';'YIR015W';'YIR017C';'YIR018W';'YIR019C';
'YIR021W';'YIR022W';'YIR023W';'YIR024C';'YIR025W';'YIR026C';'YIR027C';'YIR028
W';'YIR029W';'YIR030C';'YIR031C';'YIR032C';'YIR033W';'YIR034C';'YIR037W';'YIR
038C';'YIR039C';'YIR041W';'YJL001W';'YJL002C';'YJL003W';'YJL004C';'YJL005W';'
YJL006C';'YJL008C';'YJL010C';'YJL011C';'YJL012C';'YJL013C';'YJL014W';'YJL019W
';'YJL020C';'YJL023C';'YJL024C';'YJL025W';'YJL026W';'YJL028W';'YJL029C';'YJL0
30W';'YJL031C';'YJL033W';'YJL034W';'YJL035C';'YJL036W';'YJL037W';'YJL038C';'Y
JL039C';'YJL041W';'YJL042W';'YJL044C';'YJL045W';'YJL046W';'YJL047C';'YJL048C'
;'YJL050W';'YJL051W';'YJL052W';'YJL053W';'YJL054W';'YJL056C';'YJL057C';'YJL05
8C';'YJL059W';'YJL060W';'YJL061W';'YJL062W';'YJL062W-
A';'YJL063C';'YJL065C';'YJL066C';'YJL068C';'YJL069C';'YJL071W';'YJL072C';'YJL
073W';'YJL074C';'YJL076W';'YJL077C';'YJL078C';'YJL079C';'YJL080C';'YJL081C';'
YJL082W';'YJL083W';'YJL084C';'YJL085W';'YJL087C';'YJL088W';'YJL089W';'YJL090C
';'YJL091C';'YJL092W';'YJL093C';'YJL094C';'YJL095W';'YJL096W';'YJL097W';'YJL0
98W';'YJL099W';'YJL100W';'YJL101C';'YJL102W';'YJL103C';'YJL104W';'YJL105W';'Y
JL106W';'YJL108C';'YJL109C';'YJL110C';'YJL111W';'YJL112W';'YJL115W';'YJL116C'
;'YJL117W';'YJL118W';'YJL121C';'YJL122W';'YJL123C';'YJL124C';'YJL125C';'YJL12
6W';'YJL127C';'YJL128C';'YJL129C';'YJL130C';'YJL131C';'YJL133W';'YJL134W';'YJ
L136C';'YJL137C';'YJL138C';'YJL139C';'YJL140W';'YJL141C';'YJL143W';'YJL144W';
'YJL145W';'YJL146W';'YJL148W';'YJL149W';'YJL151C';'YJL153C';'YJL154C';'YJL155
C';'YJL156C';'YJL157C';'YJL158C';'YJL159W';'YJL162C';'YJL164C';'YJL165C';'YJL
166W';'YJL167W';'YJL168C';'YJL170C';'YJL171C';'YJL172W';'YJL173C';'YJL174W';'
YJL176C';'YJL177W';'YJL178C';'YJL179W';'YJL180C';'YJL183W';'YJL184W';'YJL186W
';'YJL187C';'YJL189W';'YJL190C';'YJL191W';'YJL192C';'YJL194W';'YJL196C';'YJL1
97W';'YJL198W';'YJL200C';'YJL201W';'YJL203W';'YJL204C';'YJL205C';'YJL207C';'Y

```
JL208C';'YJL209W';'YJL210W';'YJL212C';'YJL213W';'YJL214W';'YJL216C';'YJL217W'
;'YJL219W';'YJL221C';'YJL222W';'YJL223C';'YJR001W';'YJR002W';'YJR004C';'YJR00
5W';'YJR006W';'YJR007W';'YJR009C';'YJR010C-
A';'YJR010W';'YJR013W';'YJR014W';'YJR016C';'YJR017C';'YJR019C';'YJR021C';'YJR
022W';'YJR024C';'YJR025C';'YJR031C';'YJR032W';'YJR033C';'YJR034W';'YJR035W';'
YJR036C';'YJR040W';'YJR041C';'YJR042W';'YJR043C';'YJR044C';'YJR045C';'YJR046W
';'YJR047C';'YJR048W';'YJR049C';'YJR050W';'YJR051W';'YJR052W';'YJR053W';'YJR0
55W';'YJR057W';'YJR058C';'YJR059W';'YJR060W';'YJR062C';'YJR063W';'YJR064W';'Y
JR065C';'YJR066W';'YJR067C';'YJR068W';'YJR069C';'YJR070C';'YJR072C';'YJR073C'
;'YJR074W';'YJR075W';'YJR076C';'YJR077C';'YJR078W';'YJR080C';'YJR082C';'YJR08
3C';'YJR084W';'YJR086W';'YJR088C';'YJR089W';'YJR090C';'YJR091C';'YJR092W';'YJ
R093C';'YJR094C';'YJR094W-
A';'YJR095W';'YJR096W';'YJR097W';'YJR099W';'YJR100C';'YJR101W';'YJR102C';'YJR
103W';'YJR104C';'YJR105W';'YJR106W';'YJR108W';'YJR109C';'YJR110W';'YJR112W';'
YJR113C';'YJR117W';'YJR118C';'YJR119C';'YJR120W';'YJR121W';'YJR122W';'YJR123W
';'YJR125C';'YJR126C';'YJR127C';'YJR130C';'YJR131W';'YJR132W';'YJR133W';'YJR1
34C';'YJR135C';'YJR135W-
A';'YJR136C';'YJR137C';'YJR138W';'YJR139C';'YJR140C';'YJR143C';'YJR144W';'YJR
145C';'YJR147W';'YJR148W';'YJR150C';'YJR151C';'YJR152W';'YJR153W';'YJR155W';'
YJR156C';'YJR158W';'YJR159W';'YJR160C';'YJR161C';'YKL001C';'YKL002W';'YKL003C
';'YKL004W';'YKL005C';'YKL006C-
A';'YKL006W';'YKL007W';'YKL008C';'YKL009W';'YKL010C';'YKL011C';'YKL012W';'YKL
013C';'YKL014C';'YKL015W';'YKL016C';'YKL017C';'YKL018W';'YKL019W';'YKL020C';'
YKL021C';'YKL022C';'YKL024C';'YKL025C';'YKL026C';'YKL027W';'YKL028W';'YKL029C
';'YKL032C';'YKL033W';'YKL034W';'YKL035W';'YKL037W';'YKL038W';'YKL039W';'YKL0
40C';'YKL041W';'YKL042W';'YKL043W';'YKL045W';'YKL046C';'YKL048C';'YKL049C';'Y
KL050C';'YKL051W';'YKL052C';'YKL053C-
A';'YKL054C';'YKL055C';'YKL056C';'YKL057C';'YKL058W';'YKL059C';'YKL060C';'YKL
062W';'YKL064W';'YKL065C';'YKL067W';'YKL068W';'YKL069W';'YKL072W';'YKL073W';'
YKL074C';'YKL078W';'YKL079W';'YKL080W';'YKL081W';'YKL082C';'YKL084W';'YKL085W
';'YKL086W';'YKL087C';'YKL088W';'YKL089W';'YKL090W';'YKL091C';'YKL092C';'YKL0
93W';'YKL094W';'YKL095W';'YKL096W';'YKL096W-
A';'YKL098W';'YKL099C';'YKL101W';'YKL103C';'YKL104C';'YKL106W';'YKL108W';'YKL
109W';'YKL110C';'YKL112W';'YKL113C';'YKL114C';'YKL116C';'YKL117W';'YKL119C';'
YKL120W';'YKL122C';'YKL124W';'YKL125W';'YKL126W';'YKL127W';'YKL128C';'YKL129C
';'YKL130C';'YKL132C';'YKL134C';'YKL135C';'YKL137W';'YKL138C';'YKL138C-
A';'YKL139W';'YKL140W';'YKL141W';'YKL142W';'YKL143W';'YKL144C';'YKL145W';'YKL
146W';'YKL148C';'YKL149C';'YKL150W';'YKL152C';'YKL154W';'YKL155C';'YKL156W';'
YKL157W';'YKL159C';'YKL160W';'YKL161C';'YKL163W';'YKL164C';'YKL165C';'YKL166C
';'YKL167C';'YKL168C';'YKL170W';'YKL171W';'YKL172W';'YKL173W';'YKL174C';'YKL1
75W';'YKL176C';'YKL178C';'YKL179C';'YKL180W';'YKL181W';'YKL182W';'YKL183W';'Y
KL184W';'YKL185W';'YKL186C';'YKL188C';'YKL189W';'YKL190W';'YKL191W';'YKL192C'
;'YKL193C';'YKL194C';'YKL195W';'YKL196C';'YKL197C';'YKL198C';'YKL201C';'YKL20
3C';'YKL204W';'YKL205W';'YKL206C';'YKL207W';'YKL208W';'YKL209C';'YKL210W';'YK
L211C';'YKL212W';'YKL213C';'YKL214C';'YKL215C';'YKL216W';'YKL217W';'YKL218C';
'YKL219W';'YKL220C';'YKL221W';'YKL222C';'YKL224C';'YKR001C';'YKR002W';'YKR003
W';'YKR004C';'YKR006C';'YKR007W';'YKR008W';'YKR009C';'YKR010C';'YKR013W';'YKR
014C';'YKR016W';'YKR019C';'YKR020W';'YKR021W';'YKR022C';'YKR024C';'YKR025W';'
YKR026C';'YKR027W';'YKR028W';'YKR029C';'YKR030W';'YKR031C';'YKR034W';'YKR035W
-
A';'YKR036C';'YKR037C';'YKR038C';'YKR039W';'YKR041W';'YKR042W';'YKR043C';'YKR
044W';'YKR046C';'YKR048C';'YKR049C';'YKR050W';'YKR052C';'YKR053C';'YKR054C';'
YKR055W';'YKR056W';'YKR057W';'YKR058W';'YKR059W';'YKR060W';'YKR061W';'YKR062W
';'YKR063C';'YKR064W';'YKR065C';'YKR066C';'YKR067W';'YKR068C';'YKR069W';'YKR0
71C';'YKR072C';'YKR074W';'YKR076W';'YKR077W';'YKR078W';'YKR079C';'YKR080W';'Y
KR081C';'YKR082W';'YKR083C';'YKR084C';'YKR085C';'YKR086W';'YKR087C';'YKR088C'
;'YKR089C';'YKR090W';'YKR091W';'YKR092C';'YKR093W';'YKR094C';'YKR095W';'YKR09
```

5W-
A';'YKR096W';'YKR097W';'YKR098C';'YKR099W';'YKR100C';'YKR101W';'YKR102W';'YKR
103W';'YKR104W';'YKR106W';'YLL001W';'YLL002W';'YLL003W';'YLL004W';'YLL005C';'
YLL006W';'YLL008W';'YLL009C';'YLL010C';'YLL011W';'YLL012W';'YLL013C';'YLL014W
';'YLL015W';'YLL018C';'YLL018C-
A';'YLL019C';'YLL021W';'YLL022C';'YLL023C';'YLL024C';'YLL025W';'YLL026W';'YLL
027W';'YLL028W';'YLL029W';'YLL031C';'YLL032C';'YLL033W';'YLL034C';'YLL035W';'
YLL036C';'YLL038C';'YLL039C';'YLL040C';'YLL041C';'YLL042C';'YLL043W';'YLL045C
';'YLL046C';'YLL048C';'YLL049W';'YLL050C';'YLL051C';'YLL052C';'YLL055W';'YLL0
57C';'YLL060C';'YLL061W';'YLL062C';'YLL063C';'YLL064C';'YLR002C';'YLR003C';'Y
LR004C';'YLR005W';'YLR006C';'YLR007W';'YLR008C';'YLR009W';'YLR010C';'YLR011W'
;'YLR013W';'YLR014C';'YLR015W';'YLR016C';'YLR017W';'YLR018C';'YLR019W';'YLR02
0C';'YLR021W';'YLR022C';'YLR023C';'YLR024C';'YLR025W';'YLR026C';'YLR027C';'YL
R028C';'YLR029C';'YLR032W';'YLR033W';'YLR034C';'YLR035C';'YLR037C';'YLR038C';
'YLR039C';'YLR043C';'YLR044C';'YLR045C';'YLR047C';'YLR048W';'YLR051C';'YLR052
W';'YLR054C';'YLR055C';'YLR056W';'YLR058C';'YLR059C';'YLR060W';'YLR061W';'YLR
064W';'YLR066W';'YLR067C';'YLR068W';'YLR069C';'YLR070C';'YLR071C';'YLR073C';'
YLR074C';'YLR075W';'YLR077W';'YLR078C';'YLR079W';'YLR080W';'YLR081W';'YLR082C
';'YLR083C';'YLR084C';'YLR085C';'YLR086W';'YLR087C';'YLR088W';'YLR089C';'YLR0
90W';'YLR091W';'YLR092W';'YLR093C';'YLR094C';'YLR095C';'YLR096W';'YLR097C';'Y
LR098C';'YLR099C';'YLR100W';'YLR102C';'YLR103C';'YLR105C';'YLR106C';'YLR107W'
;'YLR109W';'YLR110C';'YLR113W';'YLR114C';'YLR115W';'YLR116W';'YLR117C';'YLR11
8C';'YLR119W';'YLR120C';'YLR121C';'YLR127C';'YLR128W';'YLR129W';'YLR130C';'YL
R131C';'YLR133W';'YLR134W';'YLR135W';'YLR136C';'YLR137W';'YLR138W';'YLR139C';
'YLR141W';'YLR142W';'YLR144C';'YLR145W';'YLR146C';'YLR147C';'YLR148W';'YLR150
W';'YLR151C';'YLR153C';'YLR154C';'YLR154W-
C';'YLR155C';'YLR157C';'YLR158C';'YLR160C';'YLR162W';'YLR163C';'YLR164W';'YLR
165C';'YLR166C';'YLR167W';'YLR168C';'YLR170C';'YLR172C';'YLR174W';'YLR175W';'
YLR176C';'YLR178C';'YLR179C';'YLR180W';'YLR181C';'YLR182W';'YLR183C';'YLR185W
';'YLR186W';'YLR188W';'YLR189C';'YLR190W';'YLR191W';'YLR192C';'YLR193C';'YLR1
94C';'YLR195C';'YLR196W';'YLR197W';'YLR199C';'YLR200W';'YLR201C';'YLR203C';'Y
LR204W';'YLR205C';'YLR206W';'YLR207W';'YLR208W';'YLR209C';'YLR210W';'YLR212C'
;'YLR213C';'YLR214W';'YLR215C';'YLR216C';'YLR218C';'YLR219W';'YLR220W';'YLR22
1C';'YLR222C';'YLR223C';'YLR226W';'YLR227C';'YLR228C';'YLR229C';'YLR231C';'YL
R233C';'YLR234W';'YLR237W';'YLR238W';'YLR239C';'YLR240W';'YLR242C';'YLR244C';
'YLR245C';'YLR246W';'YLR247C';'YLR248W';'YLR249W';'YLR250W';'YLR251W';'YLR254
C';'YLR256W';'YLR258W';'YLR259C';'YLR260W';'YLR262C';'YLR262C-
A';'YLR263W';'YLR264W';'YLR265C';'YLR266C';'YLR268W';'YLR270W';'YLR272C';'YLR
273C';'YLR274W';'YLR275W';'YLR276C';'YLR277C';'YLR284C';'YLR285W';'YLR286C';'
YLR287C-
A';'YLR288C';'YLR289W';'YLR291C';'YLR292C';'YLR293C';'YLR295C';'YLR298C';'YLR
299W';'YLR300W';'YLR301W';'YLR303W';'YLR304C';'YLR305C';'YLR306W';'YLR307W';'
YLR308W';'YLR309C';'YLR310C';'YLR312W-
A';'YLR313C';'YLR314C';'YLR315W';'YLR316C';'YLR318W';'YLR319C';'YLR320W';'YLR
321C';'YLR323C';'YLR324W';'YLR325C';'YLR327C';'YLR328W';'YLR329W';'YLR330W';'
YLR332W';'YLR333C';'YLR335W';'YLR336C';'YLR337C';'YLR340W';'YLR341W';'YLR342W
';'YLR343W';'YLR344W';'YLR347C';'YLR348C';'YLR350W';'YLR351C';'YLR353W';'YLR3
54C';'YLR355C';'YLR356W';'YLR357W';'YLR359W';'YLR360W';'YLR361C';'YLR362W';'Y
LR363C';'YLR364W';'YLR367W';'YLR368W';'YLR369W';'YLR370C';'YLR371W';'YLR372W'
;'YLR373C';'YLR375W';'YLR376C';'YLR377C';'YLR378C';'YLR380W';'YLR381W';'YLR38
2C';'YLR383W';'YLR384C';'YLR385C';'YLR386W';'YLR387C';'YLR388W';'YLR389C';'YL
R390W';'YLR390W-
A';'YLR392C';'YLR393W';'YLR394W';'YLR395C';'YLR396C';'YLR397C';'YLR398C';'YLR
399C';'YLR401C';'YLR403W';'YLR404W';'YLR405W';'YLR406C';'YLR409C';'YLR410W';'
YLR411W';'YLR412W';'YLR414C';'YLR417W';'YLR418C';'YLR420W';'YLR421C';'YLR423C
';'YLR424W';'YLR425W';'YLR427W';'YLR429W';'YLR430W';'YLR431C';'YLR432W';'YLR4
33C';'YLR435W';'YLR436C';'YLR437C';'YLR438C-

A';'YLR438W';'YLR439W';'YLR440C';'YLR441C';'YLR442C';'YLR443W';'YLR447C';'YLR
448W';'YLR449W';'YLR450W';'YLR451W';'YLR452C';'YLR453C';'YLR457C';'YLR459W';'
YLR461W';'YLR466W';'YLR467W';'YML001W';'YML004C';'YML005W';'YML006C';'YML007W
';'YML008C';'YML009C';'YML010W';'YML011C';'YML012W';'YML013W';'YML014W';'YML0
15C';'YML016C';'YML017W';'YML019W';'YML021C';'YML022W';'YML023C';'YML024W';'Y
ML025C';'YML026C';'YML027W';'YML028W';'YML029W';'YML030W';'YML031W';'YML032C'
;'YML034W';'YML035C';'YML036W';'YML038C';'YML041C';'YML042W';'YML043C';'YML04
6W';'YML047C';'YML048W';'YML049C';'YML050W';'YML051W';'YML052W';'YML054C';'YM
L055W';'YML056C';'YML057W';'YML058W';'YML058W-
A';'YML059C';'YML060W';'YML061C';'YML062C';'YML063W';'YML064C';'YML065W';'YML
066C';'YML067C';'YML068W';'YML069W';'YML070W';'YML071C';'YML072C';'YML073C';'
YML074C';'YML075C';'YML076C';'YML077W';'YML078W';'YML080W';'YML081C-
A';'YML081W';'YML085C';'YML086C';'YML087C';'YML088W';'YML091C';'YML092C';'YML
093W';'YML094W';'YML095C';'YML097C';'YML098W';'YML099C';'YML100W';'YML101C';'
YML102W';'YML103C';'YML104C';'YML105C';'YML106W';'YML107C';'YML109W';'YML110C
';'YML111W';'YML112W';'YML113W';'YML114C';'YML115C';'YML116W';'YML117W';'YML1
18W';'YML120C';'YML121W';'YML123C';'YML124C';'YML125C';'YML126C';'YML127W';'Y
ML128C';'YML129C';'YML130C';'YML132W';'YMR001C';'YMR002W';'YMR003W';'YMR004W'
;'YMR005W';'YMR006C';'YMR008C';'YMR009W';'YMR011W';'YMR012W';'YMR013C';'YMR01
4W';'YMR015C';'YMR016C';'YMR017W';'YMR019W';'YMR020W';'YMR021C';'YMR022W';'YM
R023C';'YMR024W';'YMR025W';'YMR026C';'YMR028W';'YMR029C';'YMR030W';'YMR031C';
'YMR032W';'YMR033W';'YMR035W';'YMR036C';'YMR037C';'YMR038C';'YMR039C';'YMR040
W';'YMR041C';'YMR042W';'YMR043W';'YMR044W';'YMR047C';'YMR048W';'YMR049C';'YMR
052W';'YMR053C';'YMR054W';'YMR055C';'YMR056C';'YMR058W';'YMR059W';'YMR060C';'
YMR061W';'YMR062C';'YMR063W';'YMR064W';'YMR065W';'YMR066W';'YMR067C';'YMR068W
';'YMR069W';'YMR070W';'YMR071C';'YMR072W';'YMR073C';'YMR074C';'YMR075W';'YMR0
76C';'YMR077C';'YMR078C';'YMR079W';'YMR080C';'YMR081C';'YMR083W';'YMR086W';'Y
MR087W';'YMR088C';'YMR089C';'YMR091C';'YMR092C';'YMR093W';'YMR094W';'YMR095C'
;'YMR096W';'YMR097C';'YMR098C';'YMR099C';'YMR100W';'YMR101C';'YMR104C';'YMR10
5C';'YMR106C';'YMR107W';'YMR108W';'YMR109W';'YMR110C';'YMR112C';'YMR113W';'YM
R114C';'YMR115W';'YMR116C';'YMR117C';'YMR119W';'YMR120C';'YMR121C';'YMR123W';
'YMR125W';'YMR127C';'YMR128W';'YMR129W';'YMR131C';'YMR133W';'YMR135C';'YMR136
W';'YMR137C';'YMR138W';'YMR139W';'YMR140W';'YMR142C';'YMR143W';'YMR145C';'YMR
146C';'YMR148W';'YMR149W';'YMR150C';'YMR152W';'YMR153W';'YMR154C';'YMR156C';'
YMR157C';'YMR158W';'YMR159C';'YMR161W';'YMR162C';'YMR163C';'YMR164C';'YMR165C
';'YMR167W';'YMR168C';'YMR169C';'YMR170C';'YMR171C';'YMR172W';'YMR173W';'YMR1
74C';'YMR175W';'YMR176W';'YMR177W';'YMR179W';'YMR180C';'YMR182C';'YMR183C';'Y
MR184W';'YMR186W';'YMR188C';'YMR189W';'YMR190C';'YMR191W';'YMR192W';'YMR193W'
;'YMR194C-
B';'YMR194W';'YMR195W';'YMR197C';'YMR198W';'YMR199W';'YMR200W';'YMR201C';'YMR
202W';'YMR203W';'YMR204C';'YMR205C';'YMR207C';'YMR208W';'YMR210W';'YMR211W';'
YMR212C';'YMR213W';'YMR214W';'YMR215W';'YMR216C';'YMR217W';'YMR218C';'YMR219W
';'YMR220W';'YMR222C';'YMR223W';'YMR224C';'YMR225C';'YMR226C';'YMR227C';'YMR2
28W';'YMR229C';'YMR230W';'YMR231W';'YMR232W';'YMR233W';'YMR234W';'YMR235C';'Y
MR236W';'YMR237W';'YMR238W';'YMR239C';'YMR240C';'YMR241W';'YMR242C';'YMR243C'
;'YMR246W';'YMR247C';'YMR250W';'YMR251W';'YMR251W-
A';'YMR255W';'YMR256C';'YMR257C';'YMR258C';'YMR260C';'YMR261C';'YMR263W';'YMR
264W';'YMR266W';'YMR267W';'YMR268C';'YMR269W';'YMR270C';'YMR271C';'YMR272C';'
YMR273C';'YMR274C';'YMR275C';'YMR276W';'YMR277W';'YMR278W';'YMR279C';'YMR280C
';'YMR281W';'YMR282C';'YMR283C';'YMR284W';'YMR285C';'YMR286W';'YMR287C';'YMR2
88W';'YMR289W';'YMR290C';'YMR291W';'YMR292W';'YMR293C';'YMR294W';'YMR295C';'Y
MR296C';'YMR297W';'YMR298W';'YMR299C';'YMR300C';'YMR301C';'YMR302C';'YMR303C'
;'YMR304W';'YMR305C';'YMR306W';'YMR307W';'YMR308C';'YMR309C';'YMR311C';'YMR31
2W';'YMR313C';'YMR314W';'YMR315W';'YMR316W';'YMR318C';'YMR319C';'YMR323W';'YM
R325W';'YNL001W';'YNL002C';'YNL003C';'YNL004W';'YNL005C';'YNL006W';'YNL007C';
'YNL008C';'YNL009W';'YNL012W';'YNL014W';'YNL015W';'YNL016W';'YNL020C';'YNL021
W';'YNL023C';'YNL024C-

A';'YNL025C';'YNL026W';'YNL027W';'YNL029C';'YNL030W';'YNL031C';'YNL032W';'YNL
036W';'YNL037C';'YNL038W';'YNL039W';'YNL041C';'YNL042W';'YNL044W';'YNL045W';'
YNL047C';'YNL048W';'YNL049C';'YNL051W';'YNL052W';'YNL053W';'YNL054W';'YNL055C
';'YNL056W';'YNL059C';'YNL061W';'YNL062C';'YNL063W';'YNL064C';'YNL065W';'YNL0
66W';'YNL067W';'YNL068C';'YNL069C';'YNL070W';'YNL071W';'YNL072W';'YNL073W';'Y
NL074C';'YNL075W';'YNL076W';'YNL077W';'YNL078W';'YNL079C';'YNL080C';'YNL081C'
;'YNL082W';'YNL083W';'YNL084C';'YNL085W';'YNL087W';'YNL088W';'YNL090W';'YNL09
1W';'YNL093W';'YNL094W';'YNL096C';'YNL097C';'YNL098C';'YNL099C';'YNL100W';'YN
L101W';'YNL102W';'YNL103W';'YNL104C';'YNL106C';'YNL107W';'YNL110C';'YNL111C';
'YNL112W';'YNL113W';'YNL116W';'YNL117W';'YNL118C';'YNL119W';'YNL121C';'YNL123
W';'YNL124W';'YNL125C';'YNL126W';'YNL127W';'YNL128W';'YNL129W';'YNL130C';'YNL
131W';'YNL132W';'YNL133C';'YNL135C';'YNL136W';'YNL137C';'YNL138W';'YNL138W-
A';'YNL139C';'YNL141W';'YNL142W';'YNL145W';'YNL147W';'YNL148C';'YNL149C';'YNL
151C';'YNL152W';'YNL153C';'YNL154C';'YNL156C';'YNL157W';'YNL158W';'YNL159C';'
YNL160W';'YNL161W';'YNL162W';'YNL163C';'YNL164C';'YNL166C';'YNL167C';'YNL169C
';'YNL172W';'YNL173C';'YNL175C';'YNL177C';'YNL178W';'YNL180C';'YNL182C';'YNL1
83C';'YNL185C';'YNL186W';'YNL187W';'YNL188W';'YNL189W';'YNL191W';'YNL192W';'Y
NL194C';'YNL197C';'YNL199C';'YNL201C';'YNL202W';'YNL204C';'YNL206C';'YNL207W'
;'YNL208W';'YNL209W';'YNL210W';'YNL212W';'YNL213C';'YNL214W';'YNL215W';'YNL21
6W';'YNL218W';'YNL219C';'YNL220W';'YNL221C';'YNL222W';'YNL223W';'YNL224C';'YN
L225C';'YNL227C';'YNL229C';'YNL230C';'YNL231C';'YNL232W';'YNL233W';'YNL234W';
'YNL236W';'YNL237W';'YNL238W';'YNL239W';'YNL240C';'YNL241C';'YNL242W';'YNL243
W';'YNL244C';'YNL245C';'YNL246W';'YNL247W';'YNL248C';'YNL249C';'YNL250W';'YNL
251C';'YNL252C';'YNL253W';'YNL254C';'YNL255C';'YNL256W';'YNL257C';'YNL258C';'
YNL259C';'YNL261W';'YNL262W';'YNL263C';'YNL264C';'YNL265C';'YNL267W';'YNL268W
';'YNL269W';'YNL270C';'YNL271C';'YNL272C';'YNL273W';'YNL274C';'YNL275W';'YNL2
77W';'YNL278W';'YNL279W';'YNL280C';'YNL281W';'YNL282W';'YNL283C';'YNL284C';'Y
NL286W';'YNL287W';'YNL288W';'YNL289W';'YNL290W';'YNL291C';'YNL292W';'YNL293W'
;'YNL294C';'YNL297C';'YNL298W';'YNL299W';'YNL301C';'YNL302C';'YNL304W';'YNL30
5C';'YNL306W';'YNL307C';'YNL308C';'YNL309W';'YNL310C';'YNL311C';'YNL312W';'YN
L313C';'YNL314W';'YNL315C';'YNL316C';'YNL317W';'YNL318C';'YNL321W';'YNL322C';
'YNL323W';'YNL325C';'YNL326C';'YNL327W';'YNL328C';'YNL329C';'YNL330C';'YNL331
C';'YNL332W';'YNL333W';'YNL334C';'YNL336W';'YNL339C';'YNR001C';'YNR002C';'YNR
003C';'YNR006W';'YNR007C';'YNR008W';'YNR009W';'YNR010W';'YNR011C';'YNR012W';'
YNR013C';'YNR015W';'YNR016C';'YNR017W';'YNR018W';'YNR019W';'YNR020C';'YNR022C
';'YNR023W';'YNR024W';'YNR026C';'YNR027W';'YNR028W';'YNR030W';'YNR031C';'YNR0
32C-
A';'YNR032W';'YNR033W';'YNR034W';'YNR035C';'YNR036C';'YNR037C';'YNR038W';'YNR
039C';'YNR041C';'YNR043W';'YNR044W';'YNR045W';'YNR046W';'YNR047W';'YNR048W';'
YNR049C';'YNR050C';'YNR051C';'YNR052C';'YNR053C';'YNR054C';'YNR055C';'YNR056C
';'YNR057C';'YNR058W';'YNR059W';'YNR060W';'YNR064C';'YNR067C';'YNR069C';'YNR0
72W';'YNR074C';'YNR075W';'YNR076W';'YOL001W';'YOL002C';'YOL003C';'YOL004W';'Y
OL005C';'YOL006C';'YOL007C';'YOL008W';'YOL009C';'YOL010W';'YOL011W';'YOL012C'
;'YOL013C';'YOL015W';'YOL016C';'YOL017W';'YOL018C';'YOL020W';'YOL021C';'YOL02
2C';'YOL023W';'YOL025W';'YOL026C';'YOL027C';'YOL028C';'YOL030W';'YOL031C';'YO
L032W';'YOL033W';'YOL034W';'YOL038W';'YOL039W';'YOL040C';'YOL041C';'YOL042W';
'YOL043C';'YOL044W';'YOL045W';'YOL049W';'YOL051W';'YOL052C';'YOL052C-
A';'YOL053W';'YOL054W';'YOL055C';'YOL056W';'YOL057W';'YOL058W';'YOL059W';'YOL
060C';'YOL061W';'YOL062C';'YOL063C';'YOL064C';'YOL065C';'YOL066C';'YOL067C';'
YOL068C';'YOL069W';'YOL070C';'YOL071W';'YOL072W';'YOL076W';'YOL077C';'YOL077W
-
A';'YOL078W';'YOL080C';'YOL081W';'YOL082W';'YOL083W';'YOL084W';'YOL086C';'YOL
088C';'YOL089C';'YOL090W';'YOL091W';'YOL093W';'YOL094C';'YOL095C';'YOL096C';'
YOL097C';'YOL100W';'YOL101C';'YOL102C';'YOL103W';'YOL104C';'YOL105C';'YOL108C
';'YOL109W';'YOL110W';'YOL111C';'YOL112W';'YOL113W';'YOL115W';'YOL116W';'YOL1
17W';'YOL119C';'YOL120C';'YOL121C';'YOL122C';'YOL123W';'YOL124C';'YOL125W';'Y
OL126C';'YOL127W';'YOL128C';'YOL129W';'YOL130W';'YOL132W';'YOL133W';'YOL135C'

;'YOL136C';'YOL137W';'YOL138C';'YOL139C';'YOL140W';'YOL141W';'YOL142W';'YOL14
3C';'YOL144W';'YOL145C';'YOL146W';'YOL147C';'YOL148C';'YOL149W';'YOL151W';'YO
L152W';'YOL154W';'YOL155C';'YOL156W';'YOL157C';'YOL158C';'YOL159C';'YOL159C-
A';'YOL161C';'YOL164W';'YOL165C';'YOR001W';'YOR002W';'YOR003W';'YOR004W';'YOR
005C';'YOR006C';'YOR007C';'YOR008C';'YOR009W';'YOR010C';'YOR011W';'YOR014W';'
YOR016C';'YOR017W';'YOR018W';'YOR019W';'YOR020C';'YOR023C';'YOR025W';'YOR026W
';'YOR027W';'YOR028C';'YOR030W';'YOR031W';'YOR032C';'YOR033C';'YOR034C';'YOR0
35C';'YOR036W';'YOR037W';'YOR038C';'YOR039W';'YOR040W';'YOR042W';'YOR043W';'Y
OR044W';'YOR045W';'YOR046C';'YOR047C';'YOR048C';'YOR049C';'YOR051C';'YOR052C'
;'YOR054C';'YOR056C';'YOR057W';'YOR058C';'YOR060C';'YOR061W';'YOR063W';'YOR06
4C';'YOR065W';'YOR066W';'YOR067C';'YOR068C';'YOR069W';'YOR070C';'YOR071C';'YO
R073W';'YOR074C';'YOR075W';'YOR076C';'YOR077W';'YOR078W';'YOR079C';'YOR080W';
'YOR081C';'YOR083W';'YOR084W';'YOR085W';'YOR086C';'YOR087W';'YOR089C';'YOR090
C';'YOR091W';'YOR092W';'YOR094W';'YOR095C';'YOR096W';'YOR098C';'YOR099W';'YOR
100C';'YOR101W';'YOR103C';'YOR104W';'YOR106W';'YOR107W';'YOR108W';'YOR109W';'
YOR110W';'YOR112W';'YOR113W';'YOR115C';'YOR116C';'YOR117W';'YOR118W';'YOR119C
';'YOR120W';'YOR122C';'YOR123C';'YOR124C';'YOR125C';'YOR126C';'YOR127W';'YOR1
28C';'YOR129C';'YOR130C';'YOR132W';'YOR133W';'YOR134W';'YOR136W';'YOR137C';'Y
OR138C';'YOR140W';'YOR141C';'YOR142W';'YOR143C';'YOR144C';'YOR145C';'YOR147W'
;'YOR148C';'YOR149C';'YOR150W';'YOR151C';'YOR153W';'YOR155C';'YOR156C';'YOR15
7C';'YOR158W';'YOR159C';'YOR160W';'YOR161C';'YOR162C';'YOR163W';'YOR164C';'YO
R165W';'YOR166C';'YOR167C';'YOR168W';'YOR171C';'YOR172W';'YOR173W';'YOR174W';
'YOR175C';'YOR176W';'YOR177C';'YOR178C';'YOR179C';'YOR180C';'YOR181W';'YOR182
C';'YOR184W';'YOR185C';'YOR187W';'YOR188W';'YOR189W';'YOR190W';'YOR191W';'YOR
192C';'YOR193W';'YOR194C';'YOR195W';'YOR196C';'YOR197W';'YOR198C';'YOR201C';'
YOR202W';'YOR204W';'YOR205C';'YOR206W';'YOR207C';'YOR208W';'YOR209C';'YOR210W
';'YOR211C';'YOR212W';'YOR213C';'YOR215C';'YOR216C';'YOR217W';'YOR219C';'YOR2
21C';'YOR222W';'YOR223W';'YOR224C';'YOR226C';'YOR227W';'YOR228C';'YOR229W';'Y
OR230W';'YOR231W';'YOR232W';'YOR233W';'YOR234C';'YOR236W';'YOR237W';'YOR239W'
;'YOR241W';'YOR242C';'YOR243C';'YOR244W';'YOR245C';'YOR247W';'YOR249C';'YOR25
0C';'YOR251C';'YOR252W';'YOR253W';'YOR254C';'YOR255W';'YOR256C';'YOR257W';'YO
R258W';'YOR259C';'YOR260W';'YOR261C';'YOR262W';'YOR264W';'YOR265W';'YOR266W';
'YOR267C';'YOR269W';'YOR270C';'YOR272W';'YOR273C';'YOR274W';'YOR275C';'YOR276
W';'YOR278W';'YOR279C';'YOR280C';'YOR281C';'YOR283W';'YOR284W';'YOR285W';'YOR
286W';'YOR287C';'YOR288C';'YOR290C';'YOR291W';'YOR293W';'YOR294W';'YOR295W';'
YOR297C';'YOR298C-
A';'YOR298W';'YOR299W';'YOR301W';'YOR302W';'YOR303W';'YOR304W';'YOR305W';'YOR
306C';'YOR307C';'YOR308C';'YOR310C';'YOR311C';'YOR312C';'YOR313C';'YOR315W';'
YOR316C';'YOR317W';'YOR319W';'YOR320C';'YOR321W';'YOR322C';'YOR323C';'YOR324C
';'YOR326W';'YOR327C';'YOR328W';'YOR329C';'YOR330C';'YOR332W';'YOR334W';'YOR3
35C';'YOR336W';'YOR337W';'YOR339C';'YOR340C';'YOR341W';'YOR344C';'YOR346W';'Y
OR347C';'YOR348C';'YOR349W';'YOR350C';'YOR351C';'YOR353C';'YOR354C';'YOR355W'
;'YOR356W';'YOR357C';'YOR358W';'YOR359W';'YOR360C';'YOR361C';'YOR362C';'YOR36
3C';'YOR367W';'YOR368W';'YOR369C';'YOR370C';'YOR371C';'YOR372C';'YOR373W';'YO
R374W';'YOR375C';'YOR377W';'YOR380W';'YOR381W';'YOR382W';'YOR383C';'YOR384W';
'YOR386W';'YOR388C';'YOR391C';'YOR393W';'YOR394W';'YPL001W';'YPL002C';'YPL003
W';'YPL004C';'YPL005W';'YPL006W';'YPL007C';'YPL008W';'YPL009C';'YPL010W';'YPL
011C';'YPL012W';'YPL013C';'YPL015C';'YPL016W';'YPL017C';'YPL018W';'YPL019C';'
YPL020C';'YPL021W';'YPL022W';'YPL023C';'YPL024W';'YPL026C';'YPL027W';'YPL028W
';'YPL029W';'YPL030W';'YPL031C';'YPL032C';'YPL033C';'YPL036W';'YPL037C';'YPL0
38W';'YPL040C';'YPL042C';'YPL043W';'YPL045W';'YPL046C';'YPL047W';'YPL048W';'Y
PL049C';'YPL050C';'YPL051W';'YPL052W';'YPL053C';'YPL054W';'YPL055C';'YPL057C'
;'YPL058C';'YPL059W';'YPL060W';'YPL061W';'YPL063W';'YPL064C';'YPL065W';'YPL06
9C';'YPL070W';'YPL072W';'YPL074W';'YPL075W';'YPL076W';'YPL078C';'YPL079W';'YP
L081W';'YPL082C';'YPL083C';'YPL084W';'YPL085W';'YPL086C';'YPL087W';'YPL089C';
'YPL090C';'YPL091W';'YPL092W';'YPL093W';'YPL094C';'YPL095C';'YPL096C-
A';'YPL096W';'YPL097W';'YPL098C';'YPL099C';'YPL100W';'YPL101W';'YPL103C';'YPL

```
104W';'YPL105C';'YPL106C';'YPL110C';'YPL111W';'YPL112C';'YPL115C';'YPL116W';'
YPL117C';'YPL118W';'YPL119C';'YPL120W';'YPL121C';'YPL122C';'YPL123C';'YPL124W
';'YPL125W';'YPL126W';'YPL127C';'YPL128C';'YPL129W';'YPL130W';'YPL131W';'YPL1
32W';'YPL133C';'YPL134C';'YPL135W';'YPL137C';'YPL138C';'YPL139C';'YPL140C';'Y
PL141C';'YPL143W';'YPL144W';'YPL145C';'YPL146C';'YPL147W';'YPL148C';'YPL149W'
;'YPL151C';'YPL152W';'YPL153C';'YPL154C';'YPL155C';'YPL156C';'YPL157W';'YPL15
8C';'YPL159C';'YPL160W';'YPL161C';'YPL163C';'YPL164C';'YPL165C';'YPL166W';'YP
L167C';'YPL169C';'YPL170W';'YPL171C';'YPL172C';'YPL173W';'YPL174C';'YPL175W';
'YPL176C';'YPL177C';'YPL178W';'YPL179W';'YPL180W';'YPL181W';'YPL183C';'YPL183
W-A';'YPL184C';'YPL186C';'YPL187W';'YPL188W';'YPL189C-
A';'YPL189W';'YPL190C';'YPL192C';'YPL193W';'YPL194W';'YPL195W';'YPL196W';'YPL
198W';'YPL200W';'YPL201C';'YPL202C';'YPL203W';'YPL204W';'YPL206C';'YPL207W';'
YPL208W';'YPL209C';'YPL210C';'YPL211W';'YPL212C';'YPL213W';'YPL214C';'YPL215W
';'YPL217C';'YPL218W';'YPL219W';'YPL220W';'YPL221W';'YPL223C';'YPL224C';'YPL2
25W';'YPL226W';'YPL227C';'YPL228W';'YPL230W';'YPL231W';'YPL232W';'YPL233W';'Y
PL234C';'YPL235W';'YPL237W';'YPL239W';'YPL240C';'YPL241C';'YPL242C';'YPL243W'
;'YPL244C';'YPL246C';'YPL248C';'YPL249C';'YPL249C-
A';'YPL250C';'YPL252C';'YPL253C';'YPL254W';'YPL255W';'YPL256C';'YPL258C';'YPL
259C';'YPL262W';'YPL263C';'YPL265W';'YPL266W';'YPL267W';'YPL268W';'YPL269W';'
YPL270W';'YPL271W';'YPL273W';'YPL274W';'YPL281C';'YPL282C';'YPL283C';'YPR001W
';'YPR002W';'YPR004C';'YPR005C';'YPR006C';'YPR007C';'YPR008W';'YPR009W';'YPR0
10C';'YPR016C';'YPR017C';'YPR018W';'YPR019W';'YPR020W';'YPR021C';'YPR023C';'Y
PR024W';'YPR025C';'YPR026W';'YPR028W';'YPR029C';'YPR030W';'YPR031W';'YPR032W'
;'YPR033C';'YPR034W';'YPR035W';'YPR036W';'YPR036W-
A';'YPR037C';'YPR040W';'YPR041W';'YPR042C';'YPR043W';'YPR045C';'YPR046W';'YPR
047W';'YPR048W';'YPR049C';'YPR051W';'YPR052C';'YPR054W';'YPR055W';'YPR056W';'
YPR057W';'YPR058W';'YPR060C';'YPR061C';'YPR062W';'YPR065W';'YPR066W';'YPR067W
';'YPR068C';'YPR069C';'YPR070W';'YPR072W';'YPR073C';'YPR074C';'YPR075C';'YPR0
79W';'YPR080W';'YPR081C';'YPR082C';'YPR083W';'YPR085C';'YPR086W';'YPR088C';'Y
PR091C';'YPR093C';'YPR094W';'YPR095C';'YPR096C';'YPR097W';'YPR098C';'YPR100W'
;'YPR101W';'YPR102C';'YPR103W';'YPR104C';'YPR105C';'YPR106W';'YPR107C';'YPR10
8W';'YPR110C';'YPR111W';'YPR112C';'YPR113W';'YPR115W';'YPR116W';'YPR118W';'YP
R119W';'YPR120C';'YPR121W';'YPR122W';'YPR124W';'YPR125W';'YPR127W';'YPR128C';
'YPR129W';'YPR131C';'YPR132W';'YPR133C';'YPR133W-
A';'YPR134W';'YPR135W';'YPR137W';'YPR138C';'YPR139C';'YPR140W';'YPR141C';'YPR
143W';'YPR144C';'YPR145W';'YPR148C';'YPR149W';'YPR151C';'YPR152C';'YPR153W';'
YPR154W';'YPR155C';'YPR156C';'YPR158W';'YPR159W';'YPR160W';'YPR161C';'YPR162C
';'YPR163C';'YPR164W';'YPR165W';'YPR166C';'YPR167C';'YPR168W';'YPR169W';'YPR1
71W';'YPR173C';'YPR175W';'YPR176C';'YPR178W';'YPR179C';'YPR180W';'YPR181C';'Y
PR182W';'YPR183W';'YPR184W';'YPR185W';'YPR186C';'YPR187W';'YPR188C';'YPR189W'
;'YPR190C';'YPR191W';'YPR192W';'YPR193C';'YPR194C';'YPR198W';'YPR199C';'YPR20
0C';'YPR201W';'YPR204W';};

included_dubiousORFs = setdiff(...
    unique_single_gene_reaction_genes, verified_genes);
%there is 1 dubious ORF on the list of rxns catalyzed by 1 gene

included_essential=intersect(...
    unique_single_gene_reaction_genes,sgd_essential_genes);
%there are 104 essential genes on the list of rxns catalyzed by 1 gene

candidate_singles =
setxor(included_dubiousORFs,unique_single_gene_reaction_genes);
candidate_singles = setxor(included_essential,candidate_singles);

%So, now there are 373 genes on the list of possible candidates. Let's also
%remove any that are essential per model simulation (even if it's
```

```
%incorrect)

%the list of essential genes is as followes:
%ko_tol = 1e-6;
%grRatio = singleGeneDeletion(model);
%model_essential = model.genes(grRatio < ko_tol);

model_essential={'YBR021W';'YBR029C';'YBR041W';'YBR126C';'YBR192W';'YBR265W';
'YCR034W';'YDL015C';'YDL055C';'YDR062W';'YDR074W';'YDR353W';'YDR367W';'YDR454
C';'YER003C';'YER026C';'YER043C';'YFL045C';'YGR037C';'YGR175C';'YHR007C';'YHR
042W';'YHR072W';'YHR190W';'YJL167W';'YJR073C';'YKL004W';'YKL024C';'YKL067W';'
YKL182W';'YLR372W';'YML008C';'YML126C';'YMR202W';'YMR208W';'YMR217W';'YMR220W
';'YMR296C';'YMR298W';'YNL003C';'YNL280C';'YNR016C';'YNR043W';'YOR074C';'YOR0
95C';'YOR175C';'YOR236W';'YPL028W';'YPL117C';'YPL231W';'YPR113W';'YPR183W';};

%As I work, I've found other genes which shouldn't be included
others_to_remove = {'YNL280C',... %SGD annotates as inviable in aerobic
    'YER005W',... %ynd1 excluded b/c it shows up unexpectedly w/high flux
    'YML120C',... %ndi1 excluded b/c 353/355 top predicted involve it, and
I'm not confident about NAD metabolism
    model_essential{:}};

included_model_essential = intersect(others_to_remove,candidate_singles);

candidate_singles = setxor(included_model_essential,candidate_singles);
%Now there are 360 genes on the list of possible candidates.

candidate_singles(end+1) = {'YDR111C'};
candidate_singles=unique(candidate_singles);
save('Y5_3_candidate_singles','candidate_singles');

%%
%So now I have a list of genes that annotate reactions that are only
%annotated with a single genes. I (think) I can knock out any three of
%these genes and look for increased formic acid secretion.

doubles = combntns((1:length(candidate_singles)),2); %there are 64980
length(doubles)
save('Y5_doubles','doubles')

formate_ex=find(strcmp('formate exchange',model.rxnNames));

%make sure we're using the gurobi solver (glpk gives different answers!)
changeCobraSolver('gurobi');

fprintf(' \r');
fprintf('LP solver changed to gurobi.\n');
fprintf(' \r');

double_formate_fluxes=[];

h = waitbar(0,'Checking double knockouts ...');

for i=1:length(doubles)
    if mod(i,100) == 0
        waitbar(i/length(doubles),h);
    end
```

```
    clear mex;
    [newmodel,hasEffect,constrRxnNames,deletedGenes] = ...
        deleteModelGenes(model,candidate_singles(doubles(i,:)));
    if hasEffect
        flux=optimizeCbModel(newmodel,[],'one');
        if numel(flux.x)==0
            i
        elseif (flux.x(formate_ex)>0 && (flux.f>0.5518)) %formate flux + >20%
WT growth rate
            double_formate_fluxes = [double_formate_fluxes; i
flux.x(formate_ex) flux.f];
            sprintf('%12i %12.4f %12.4f\n',i, flux.x(formate_ex), flux.f)
        end
    end
end
close(h);

%the double_formate_fluxes variable should now be 3 columns, with each
%row consisting of the KO number in column 1 (the index of "doubles"), and
%the formate flux in column 2, and the growth rate in column 3.

save('Y5_3_double_formate_fluxes','double_formate_fluxes');

% To find which deletion gives a particular formate flux, do something like
% this:
max_formate_flux=max(double_formate_fluxes(:,2));
index=find(double_formate_fluxes(:,2)==max_formate_flux);
max_double=double_formate_fluxes(index,1);
singles_index=doubles(max_double,:);
max_formate_genes=candidate_singles(singles_index)
```