Accepted manuscript to FEMS Yeast Research, Volume 20, Issue 2, March 2020, foaa006, https://doi.org/10.1093/femsyr/foaa006 Published: 18 February 2020

redLips: a comprehensive mechanistic model of the lipid metabolic network of yeast

Author list: S. Tsouka[1], V. Hatzimanikatis[1*]

[1]Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), CH-1015 Lausanne, Switzerland.

*corresponding author: vassily.hatzimanikatis@epfl.ch

Abstract

Over the last decades, yeast has become a key model organism for the study of lipid biochemistry. Because the regulation of lipids has been closely linked to various physiopathologies, the study of these biomolecules could lead to new diagnostics and treatments. Before the field can reach this point, however, sufficient tools for integrating and analyzing the ever-growing availability of lipidomics data will need to be developed. To this end, genome-scale models (GEMs) of metabolic networks are useful tools, though their large size and complexity introduces too much uncertainty in the accuracy of predicted outcomes. Ideally, therefore, a model for studying lipids would contain only the pathways required for the proper analysis of these biomolecules, but would not be an *ad hoc* reduction. We hereby present a metabolic model that focuses on lipid metabolism constructed through the integration of detailed lipid pathways into an already existing GEM of *Saccharomyces cerevisiae*. Our model was then systematically reduced around the subsystems defined by these pathways to provide a more manageable model size for complex studies. We show that this model is as consistent and inclusive as other yeast GEMs regarding the focus and detail on the lipid metabolism, and can be used as a scaffold for integrating lipidomics data to improve predictions in studies of lipid-related biological functions.

Introduction

Even the slightest changes in cellular membrane composition, which all serve a specific biological purpose, can affect many cellular functions from signaling cascades to the modulation of membrane fluidity (Guan, et al. 2009). Because they are the main structural component of cellular membranes, lipid imbalances have been shown to be involved in various physiopathologies concerning membrane lipid homeostasis (Holthuis and Menon 2014). Yeast is a very prominent model organism for the study of numerous parts of cell metabolism including, but not limited to, lipid-related cellular processes (de Kroon 2017, Klose, et al. 2012) because it is easy and inexpensive to cultivate and modify its genome for experiments, and it has a well-documented

genome sequence (Santos and Riezman 2012). Consequently, an increasing spectrum of yeast mutants has been made available, providing great opportunities for studies on the effects of lipid metabolism perturbations at molecular and cellular levels. One additional feature of yeast is its high homology to the human genome. Most importantly, the majority of regulatory mechanisms are preserved between the species (Petranovic, et al. 2010). This means that yeast could potentially be used as a platform to study lipid dysregulation in humans, making the study of potential causalities and treatments critically easier. The similarities and differences of the two organisms, along with the potentials for comparative analysis have been reviewed in detail by (Nielsen 2009). Unfortunately, the intricacies of how lipids tie to many biological functions, including those leading to disease, remain unknown. This means that comprehensive lipid identification and characterization and detailed studies of lipidomics are needed for a fundamental understanding of cellular metabolism (da Silveira Dos Santos, et al. 2014, Han 2016, Ivanova, et al. 2009, Kontush and Chapman 2010, Wenk 2005), and as such, recent interdisciplinary approaches are beginning to reveal novel lipid functions and interactions (Harayama and Riezman 2018). Eventually, lipidome profiling could be used as a predictive tool to further enhance our knowledge of the underlying molecular mechanisms typifying lipid dysregulation.

While traditionally established work on cellular lipid metabolism has been limited to the analysis of individual classes of lipids or specific lipid species, progress in mass spectrometry (MS)-based methodologies has allowed the analysis of the entirety of the lipids in a cell. Computational metabolic models of various pathways have emerged in an effort to evaluate the vast omics data available, and many different approaches for their construction and curation with incorporated omics data have been developed (Joyce and Palsson 2006). This mostly involves genome-scale models (GEMs) of metabolism, which are reconstructions of an organism's metabolism from genomic, biochemical, and physiological data, and in principle, contain the majority of known information for the modeled organism. With the increasing availability of omics data, however, comes increasing mathematical complexity, and it can be very complicated to handle the incorporation of experimental data in such large-scale models. The potential of dynamic modeling through the generation of appropriate sets of ordinary differential equations that describe the network topology is also hindered by the model's size. Mathematically, a larger model also leads to an increased solution space, which ultimately contributes to increased uncertainty in the model's predictions. Therefore, it is essential that a network is manageable with respect to size without a loss of information, so the redGEM framework was proposed as a way to systematically reduce GEMs around a biological context of interest with minimal loss of information and connectivity (Ataman, et al. 2017). On the other hand, due to the rapid discovery of novel species through innovative technologies, a gap is emerging between the existing pathway representations of lipids and lipid structure databases. An approach aiming to bridge this gap has been proposed, termed Network Integrated Computational Explorer for Lipidomics (NICELips) (Hadadi, et al. 2014), and this

framework can postulate novel lipid biosynthesis pathways using generalized enzymatic reaction rules. Specific to yeast metabolism, the first GEM of *S. cerevisiae* was published in 2003 (Forster, et al. 2003), and over the years, multiple yeast GEMs have been updated and published by several research groups (Lopes and Rocha 2017). Due to inconsistencies in annotation, a community consensus reconstruction has been developed, with its latest versions being Yeast 7 and Yeast 8 (Aung, et al. 2013, Lu, et al. 2019). Very recently, a novel method for the representation of lipid requirements in GEMs was proposed (Sanchez, et al. 2019).

We thus sought to develop a metabolic model that could act as detailed repository of lipid metabolism for *S. cerevisiae*. Starting from the network provided by (Savoglidis, et al. 2016), we gathered all relevant reaction and pathway information available in the literature and databases. To ensure consistency with the well annotated GEMs, we incorporated these data into a GEM of the yeast *S. cerevisiae*, expanding its preexisting lipid description. We then performed a systematic reduction of the integrated model around the lipid subsystems to preserve the focus of the model on the lipid metabolic pathways and to simultaneously retain the connections to the rest of the cell metabolism. To ensure consistency of the cell biomass composition, we computed lumped reactions to establish the production of all biomass building blocks (BBBs). These steps made sure that our final model, termed “reduced lipids-centric model” (redLips), is inclusive yet concise and as consistent as the other available yeast GEMs. We have created a detailed thermodynamic database for all the metabolites of the network and performed a complete thermodynamic curation of redLips, a procedure that decreases the mathematical uncertainty and imposes physiological constraints. We also demonstrate how it can be used as a scaffold for lipidomic measurement implementation. In the future, redLips can be modified to accommodate simulations and predictions for human (or other) metabolism, thus creating a platform to study lipid regulation for applications across organisms.

Accepted manuscript to FEMS Yeast Research, Volume 20, Issue 2, March 2020, foaa006, https://doi.org/10.1093/femsyr/foaa006 Published: 18 February 2020

Materials & Methods

Starting reaction network

We used the model of (Savoglidis, et al. 2016) as a base to gather and reassemble the available knowledge on lipid metabolism to date. The LIPID MAPS classification system distinguishes eight major lipid categories: fatty acyls, glycerolipids, glycerophospholipids, sphingolipids, sterol lipids, prenol lipids, saccharolipids, and polyketides (Fahy, et al. 2005, Fahy, et al. 2009), and the above cited model focused on the sphingolipid biosynthesis pathway and included some of the glycerophospholipid biosynthetic route. The model was curated using thermodynamic and lipidomics data, and an extensive study on the control asserted by the highly multifunctional enzymes of the system was conducted.

The resulting gathered lipid reactions network (GLRN) was constructed by combining information found in the literature, GEMs, and databases. Primary sources of data include the online repositories *Saccharomyces Genome Database* (SGD, https://www.yeastgenome.org, (Cherry, et al. 1998)), KEGG (https://www.genome.jp/kegg/), and Lipid Maps (https://www.lipidmaps.org) as well as relevant journal publications and books (Dickinson and Schweizer 2004).

Consistent Reduction of Models

An issue that arises when modeling only a part of cell metabolism is the connection to the rest of the network. For example, if the lipid network was to be studied without including the TCA cycle, ATP would need to be obtained through an artificial transport reaction from the extra-model domain to the intra-model domain, though there is no such compartmental transport in reality. This can lead to uncertainty on the concentration levels as well as to a major question of the relevant flux constraints. To create a consistent and reliable model, we would need to constrain the flux values of all these transport reactions (which would include mostly cofactors) to realistic values.

To overcome this issue, we decided to effectively couple our model with a GEM that will account for any non-realistic assumptions that would have to be made. We did this by first incorporating our detailed lipid network into a GEM of choice, thus expanding the lipid metabolism pathways already present, then we utilized the redGEM framework to obtain a reduced model using our original subsystems as the starting network.

redGEM is a framework developed by (Ataman, et al. 2017) to systematically and consistently reduce genome-scale models. It focuses on chosen parts (subsystems) of the metabolic networks that are then connected to each other up to a user-defined degree of connection. This measure describes the distance in terms of reaction steps between a subsystem pair and can be either imposed by the user for all subsystem

pairs or can be equal to the intrinsic minimum distance between each pair. Subsequently, the resulting core network is connected to the biomass building blocks (BBBs) using lumpGEM (Ataman and Hatzimanikatis 2017).

In redGEM, a graph search algorithm is employed to identify all possible connections between metabolites belonging either to the same subsystem or different ones (excluding cofactors). The first step is the intra-expansion of the starting network, connecting metabolites within each subsystem of interest. Then these subsystems are step-wise connected to each other, first adding the one-step connections, then the two-step connections (which will involve an intermediate), etc., thus creating the core network. The degree of connection is symbolized as D#, where # is the number corresponding to the desired connection length.

After the network expansion, lumpGEM is used to identify sets of biosynthetic subnetworks that will synthesize each BBB that cannot already be produced by the core network. In other words, lumpGEM's objective is the minimization of the number of reactions that need to be added to the core to allow the production of each BBB. These sets are then collapsed into elementally balanced lumped reactions. lumpGEM first identifies the minimal subnetwork of reactions needed to connect the expanded network to each BBB. Subsequently, all alternative subnetworks of this minimal size can also be computed and are then translated into a single lumped reaction that is tested for feasibility in terms of stoichiometry and thermodynamics. Various consistency checks are performed to ensure the minimal loss of information during the reduction process. These checks include flux variability and essentiality studies in both stoichiometric and thermodynamic levels of curation between the GEM and the reduced model.

Genome Scale Model (before integration)

We integrated the GLRN into the well-known and well-studied iMM904 GEM (Mo, et al. 2009), which is annotated, ensuring that it was straightforward to match each reaction and its metabolites between the two networks. iMM904 also includes a large number of cellular compartments compared to most yeast GEMs. The interested reader may refer to (Lopes and Rocha 2017), (Sanchez and Nielsen 2015) and (Osterlund, et al. 2012) for a more detailed review of the development and evolution of various *S. cerevisiae* GEMs.

Gathered Lipid Reactions Network (before integration)

The GLRN encompasses more than 500 enzymatic reactions and 300 metabolites. We considered 7 cellular compartments where all the reactions take place, which are the cytosol, mitochondria, endoplasmic reticulum (ER), peroxisomes, Golgi apparatus, vacuole, and nucleus (as well as extracellular space). The model can be organized into 15 subsystems: glycolysis, pyruvate metabolism, fatty acid biosynthesis, fatty acid

mitochondrial biosynthesis, fatty acid elongation, fatty acid degradation, phospholipid biosynthesis, sphingolipid biosynthesis, sterol biosynthesis and esterification, sterol metabolism, mevalonate pathway, dolichol biosynthesis, cardiolipin biosynthesis, carnitine shuttle, and triacylglyceride decomposition. We did not consider any membrane compartments or lipid bodies in our study since our thermodynamic calculations do not hold for non-aqueous solutions, as will be explained shortly, and we instead opted for consistency over extensive detail.

The localization assignment for each reaction was made according to the Yeast7 and Yeast8 consensus GEMs (Aung, et al. 2013, Lu, et al. 2019). For the reactions that are not included in this model, the N-terminal amino acid sequence of the associated gene was used to predict localization (Emanuelsson, et al. 2007).

Genome Scale Model (after integration)

After we integrated the GLRN into the GEM, the integrated model had 2181 reactions and 1551 metabolites. The lipid-related reaction subsystems of iMM904 that were mostly expanded were the fatty acid biosynthesis and degradation, as well as the sterol biosynthesis and esterification, all of which existed mostly as lumped reactions or were missing parts of the pathways. The phospholipid and sphingolipid biosynthetic pathways originally included mostly mass-imbalanced and pooled reactions and were also greatly enhanced, with parts like phospholipid remodeling being added. Similarly, the lipid species that were added to the model mostly included fatty acids of different carbon chain lengths, complex sphingolipids, monolyso-glycerophospholipids, and fatty acid biosynthesis and degradation as well as sterol intermediates, over all of the cellular compartments.

Subsequently, we curated this model on both stoichiometric and thermodynamic levels. First, we removed all the reactions that were lumped reactions that we explicitly included in the GLRN, that were mis-assigned to other compartments, or that were in any other way rendered redundant by the integration. The curated integrated model includes 1531 reactions and 1078 metabolites.

Lipidomics – biosynthetic fluxes

Lipids are an essential component of the cell's various membranes and are critical for cell survival. Thus, being essential to biomass formation, they should be present in the modeled assumption of the biomass composition. However, few GEMs even consider these lipids as part of the growth requirements, let alone encompass the lipid network in detail. We have identified 37 metabolites that should be considered, which are 4 phospholipids (phosphatidylethanolamine, phosphatidylcholine, phosphatidylserine and phosphatidylinositol), 4 lyso-phospholipids (lyso-phosphatidylethanolamine, lyso-phosphatidylcholine, lyso-phosphatidylserine and lyso-phosphatidylinositol), 20

complex sphingolipids, ergosterol, 4 sterol esters (ergosterol, episterol, lanosterol and zymosterol esters), dolichol, as well as long and very long chain fatty acids.

We did not wish to alter the biomass reaction already defined in iMM904, so we defined 35 additional biosynthetic reactions. These reactions are all single (or double) species exchange reactions, all of which are essential to cell growth. This artificial representation corresponds to elementary fluxes of the aforementioned lipid species towards biomass formation. These fluxes can be constrained based on experimental concentration measurements (when available) as:

$$\mu(\overline{Y_i} - sd) \leq v_i \leq \mu(\overline{Y_i} + sd),$$

where v_i are the biosynthetic reaction fluxes, $\overline{Y_i}$ the mean of the lipidomic content measurements, sd the experimental measurements' standard deviation of $\overline{Y_i}$, and μ the specific growth rate of the cell as calculated from the flux through the biomass objective function.

It is important to note that when a species is already considered in the biomass composition of the GEM, the experimental constraint is altered accordingly to consider the corresponding amount required for each contribution.

Thermodynamics

Next, we performed a complete thermodynamic curation of the integrated model using TFA (Salvy, et al. 2019) to further reduce the solution space of the problem and help identify reaction directionalities. For a reaction to be feasible in the assumed directionality, the net change in Gibbs free energy of a reaction ($\Delta_r G'^o$) must be negative in this direction.

Lipids are very complex molecules, and thermodynamic information about them, such as the Gibbs free energies of formation and dissolution constants, is scarce. Where available, experimental observations indicating a pathway direction were used, which in turn provided insight for the whole reaction network. Otherwise, group contribution methods were used, which predict properties of complex molecules by using group or atom properties (Mavrovouniotis 1990, Mavrovouniotis 1991). Thus, very complicated molecules can be decomposed into a number of simple groups, and their individual contributions to the total properties can be estimated.

Since thermodynamic properties depend on the pH of the environment, we needed to assign a pH value to each considered compartment (Orij, et al. 2009, Paroutis, et al. 2004, Preston, et al. 1989). For cross-membrane transport reactions, we also needed to take the membrane potential difference, if any (Cohen and Venkatachalam 2014), into account. Finally, all of the calculated changes in Gibbs free energy needed to be adjusted with the associated compartmental ionic strength (Ataman 2016). All of these values can be found in Table 1.

Table 1. Values for pH and ionic strength (in M) for each model compartment, and cross-membrane potentials (in mV) for each set of these compartments (where applicable) – opposite arrow direction will correspond to the same value with opposite sign.

#	Compartment	pH	Ionic Strength (M)	Cross-membrane potential (mV)
I	Cytosol	7	0.25	n/a
II	Endoplasmic Reticulum	7.2	0	n/a
III	Golgi Apparatus	6.35	0	n/a
IV	Mitochondria	7.5	0.25	IV → I: 180
V	Nucleus	7	0	V → I: 15
VI	Peroxisome	8.2	0	n/a
VII	Vacuole	6.17	0	n/a
VIII	Extracellular	5	0	VIII → I: -60

To estimate the properties of a lipid containing a fatty acyl carbon chain, we needed to assume a chain length for each of the attached R groups. We chose C16:0 (where the first number denotes the carbon chain length and second denotes the number of unsaturations on this chain) for all species, since this chain length represents the vast majority of lipids in eukaryotes. Regardless, this assumption does not carry much weight in our model, since the group contribution method used to estimate the Gibbs free energy of a reaction considers only the groups that undergo a molecular change. Consequently, if the R group is not the reactive part of the molecule participating in the reaction, its length will not affect the calculated $\Delta_r G'^o$ value. One more assumption that needed to be made was that no reactions occurred inside membranes. It is known that this is not the case for numerous lipid biotransformations, but since all thermodynamic properties have been measured with the assumption of an aqueous solution and are computed accordingly, it was a necessary assumption.

With these in mind, we curated a thermodynamic database containing all the thermodynamic properties of the model's metabolites, such as pKa, standard Gibbs free energy of formation, formula, charge, etc. These properties were calculated through Chemaxon (https://www.chemaxon.com). This database covers 90.4% of the integrated network's metabolites, which allowed us to calculate 87.4% of the $\Delta_r G'^o$ of the network reactions.

Assumptions made about the thermodynamic constraints, such as temperature and pH, or even uncertainty in the calculation and the standard deviation of measurements, can render networks computationally infeasible. Additionally, especially concerning lipid metabolism, channeling phenomena can lead to apparently infeasible reactions in a certain directionality. Regardless, since we were confident in most of the reaction directionalities in our network, we could adjust some thermodynamic constraints to attain feasible solutions in what we consider physiological conditions. More specifically, in order to retain consistency with yeast physiology, we relaxed 62 thermodynamic feasibility constraints in terms of the $\Delta_r G'^o$. These constraints

correspond in majority to lipid species transport reactions across intracellular compartments. This is actually a case for which our computations may not hold, though, since lipid species do not cross membranes in the same way as most others. The complete list of the $\Delta_r G'^o$ relaxations can be found in Supplementary Table S3.

FA chain lengths

Lipid species consist of R groups of acyl chains exchanged between themselves or provided by free fatty acids. These chains vary in size and usually contain an even number of carbon atoms from 4 to 28 as well as often one unsaturation. As mentioned previously, the most abundant fatty acids in yeast have a 16-carbon chain, and the second most abundant have an 18-carbon chain, which together comprise more than 70% of the total fatty acid population (Daum, et al. 1999, Schneiter, et al. 1999). In this model, we consider acyl chains only of even chain lengths varying from C8 to C26. Because any chain length or combination thereof could react to form a lipid species, we treated the fatty acids (in both inactive and coenzyme A [CoA]-activated form) as metabolite pools, which comprised all of the fatty acyl providers. We also defined a metabolite pool for polyprenol diphosphates, which include species possessing 14 to 22 prenyl units.

Lipidomics – concentrations

As mentioned above, experimental measurements can be used to constrain fluxes and effectively couple them to biomass formation. The metabolic concentrations of species can also be constrained through lipidomics as:

$$ln(\bar{X}_\iota - sd) \leq LC_i \leq ln(\bar{X}_\iota + sd),$$

where $\bar{X}_\iota$ is the mean of the concentration measurements, sd the standard deviation of the experimental measurements of $\bar{X}_\iota$, and LC_i are the natural logarithms of the concentrations for each compound.

Media

To ensure that the maximum growth rate predicted by the model reflects a typical growth rate for yeast in aerobic conditions (about 0.32-0.48 h^{-1}), we constrained the maximum uptake of glucose (which we considered to be the sole carbon source) to 4 $mmol \cdot gDW^{-1} \cdot h^{-1}$ (Orij, et al. 2012). The other uptakes allowed were the following inorganics: hydrogen, water, ammonium, oxygen (limited to 20 $mmol \cdot gDW^{-1} \cdot h^{-1}$), phosphate, and sulfate. We also had the option to allow a basal uptake of exogenous ethanolamine (we chose a value of up to 0.02 $mmol \cdot gDW^{-1} \cdot h^{-1}$) to activate the reaction catalyzed by ethanolamine kinase (EKI1), the first step of phosphatidylethanolamine (PE) synthesis via the Kennedy pathway.

Accepted manuscript to FEMS Yeast Research, Volume 20, Issue 2, March 2020, foaa006, https://doi.org/10.1093/femsyr/foaa006 Published: 18 February 2020

Results & Discussion

redGEM output model

To form our lipid-focused reduced metabolic model, we first applied the redGEM and lumpGEM algorithms to the previously defined subsystems of interest, GLRN, and the glycolysis pathway. We also included the electron transport chain (ETC) reactions to the starting subsystems to ensure consistent energy associations and that the growth rates were as equivalent to the GEM as possible. For this reduction, we set the degree of connection to 3, which means that pairwise subsystem connections of up to 3 steps each will be added during the subsequent network expansions. The resulting model encompassed 1130 reactions, of which 639 were enzymatic, 419 were transport or boundary, 35 were biosynthetic (as described in the Materials & Methods section), and 37 were lumped reactions. Additionally, the reduced model included 800 metabolites, 404 of which were unique across compartments.

After formation of the reduced model, to ensure and evaluate its function and the minimal loss of information from the integrated model, we conducted consistency checks in terms of enzyme essentiality and thermodynamic flux variability. The results from these tests can be found in Supplementary Tables S1 and S2. These tests showed that, as expected, redLips exhibits equal or less variability in terms of flux ranges compared to the integrated GEM, since some information will unavoidably be lost through the reduction process. In any case, all of the flux values in the solution space of redLips are a subset of the integrated GEM's solution space, as they would otherwise be inconsistent. Similarly, redLips has more essential enzymes than the integrated GEM, though the essential enzymes of the latter are all a subset of the former. This discrepancy can occur mainly because of two reasons: First, it is possible that some of the enzymes that are essential for redLips and non-essential for the GEM participate in lumped reactions, thus are indispensable for growth. Additionally, this means that these enzymes catalyze reactions that are present in all the computed alternatives for one (or more) BBB for the minimal subnetwork size. Second, some alternative pathways compensating for the loss of this enzymatic activity might be lost due to the reduction process, making it essential in the reduced model.

Overview of the reactions and metabolites in each expansion step

In order to get a clearer picture of redLips' structure and the overall network connectivity, we took a closer look at the reactions added to the model in each expansion step with respect to reactions that can carry flux. In the following discussion, the number of reactions comprising the graph search output will be given in parentheses next to the number of feasible (flux carrying) reaction additions. The starting subsystems include 540 reactions and 609 metabolites (307 unique ones

across compartments) in total. The complete list of subsystems included in redLips and the respective numbers of reactions in each of them, along with the number of reactions added in each expansion of the starting network are given in detail in Table 2.

In the D1 expansion of the model (one-step connections between core subsystems), most of the added reactions were transport reaction across compartments. The starting subsystems did not include any transport reactions, and all of the existing ones that connect metabolites belonging to the starting network were added at this stage since they are one step connections. Concerning the central carbon pathways, one reaction from the TCA cycle was included, namely the oxidation of succinate to fumarate and the reaction catalyzed by transaldolase from the pentose phosphate pathway. At this stage, a total of 199 (243) reactions were added to the core model along with 12 (22) new metabolites, of which 9 (17) were unique across compartments.

In the D2 expansion, a total of 44 (57) reactions were added to the D1 model. This seems like a significantly smaller number than in the previous step, though since the vast majority of the computed one-step reactions were transport reactions, this number is much larger than the enzymatic reactions that were added to the core during the D1 expansion. These reactions involve 35 (43) new metabolites, 30 (38) of which are unique, and include: the condensation of acetyl-CoA and oxaloacetate to form citrate in the cytosol and the peroxisomes (TCA cycle) and the reactions catalyzed by transketolase activity (pentose phosphate pathway).

The D3 expansion of the model encompassed 51 (71) additional reactions, including two more reactions from the pentose phosphate pathway. In terms of compounds, 57 (75) new metabolites were added, of which 56 (72) were unique across compartments. Lastly, a final graph search added the reactions in which only core metabolites participate and that had not already been added to the model in any expansion step, which were most commonly transport reactions for cofactors and boundary reactions. In our case, there were 250 (594) reactions that matched those criteria.

It is interesting to note that as we increased the user-defined degree of connection between the core subsystems, more amino acid biosynthetic routes were added. Also, some parts of the metabolism were located many steps away from our core network as we defined it, so ultimately were not added. One example of this was the TCA cycle that started to form for the D1 and D2 model expansions, though no new reactions were added in the D3 model. It therefore remained incomplete, missing three reactions to convert α-ketoglutarate to succinate through succinyl-CoA and one reaction to balance the intermediate byproducts. To ensure a more comprehensive and consistent network, we included these four reactions in our model *a posteriori*. For the complete list of reactions of the model, the interested reader may refer to Supplementary Table S3.

Table 2. List of subsystems included in redLips, and the corresponding number of reactions that were added in each step of the reduction process. Total number of reactions per subsystem and the percentage coverage of the corresponding integrated GEM subsystem is also reported. Boldface denotes the lipid pathway subsystems. FGS: Final Graph Search, PP: Post-Processing. (*)The biomass reaction representing cell growth is not part of either the starting network or any expansion step. (**)The lumped reactions are not part of the expansion steps, and they are computed and added to the model after D3 and before FGS.

Subsystem	Starting Network	D1	D2	D3	FGS	PP	Total # of reactions (% coverage of the integrated GEM)
Alanine and Aspartate Metabolism	-	-	4	-	-	-	4 (44.4%)
Alternate Carbon Metabolism	1	-	-	10	-	-	11 (40.7%)
Anapleurotic Reactions	-	5	-	-			7 (63.6%)
Arginine and Proline Metabolism	-	-	-	2	-	-	2 (6.1%)
Cardiolipin Biosynthesis	7	-	-	-	-	-	7 (100%)
Carnitine Shuttle	4	-	-	-	-	-	4 (100%)
Citric Acid Cycle	-	1	2	-	5	3	11 (84.6%)
Complex Alcohol Metabolism	-	-	-	2	-	-	2 (7.4%)
Cysteine Metabolism	-	-	-	1	2	-	3 (30%)
Dolichol Biosynthesis	30	-	-	-	-	-	30 (100%)
Fatty Acid Biosynthesis	67	-	-	-	-	-	67 (100%)
Fatty Acid Biosynthesis Mitochondrial	39	-	-	-	-	-	39 (100%)
Fatty Acid Degradation	99	-	-	-	-	-	99 (100%)
Fatty Acid Elongation	28	-	-	-	-	-	28 (100%)
Glutamate Metabolism	-	-	-	1	4	-	5 (29.4%)
Glutamine Metabolism	-	-	-	3	1	-	4 (100%)
Glycerolipid Metabolism	-	1	3	-	1	-	5 (55.6%)
Glycine and Serine Metabolism	-	-	1	5	1	1	8 (42.1%)
Glycolysis/Gluconeogenesis	12	-	4	3	1	-	20 (100%)
Glycoprotein Metabolism	-	-	2	1	-	-	3 (42.9%)
Histidine Metabolism	-	-	-	1	-	-	1 (7.1%)
Methane Metabolism	-	-	-	-	1	-	1 (50%)
Methionine Metabolism	-	-	-	1	3	-	4 (20%)
Mevalonate pathway	10	-	-	-	-	-	10 (100%)
NAD Biosynthesis	-	-	-	-	4	-	4 (16.7%)
Nucleotide Salvage Pathway	-	-	-	-	14	-	14 (16.9%)
Oxidative Phosphorylation	17	-	-	-	-	-	17 (89.5%)
Pentose Phosphate Pathway	-	1	2	2	-	-	5 (38.5%)
Phospholipid Biosynthesis	60	-	-	-	-	-	60 (100%)
Purine and Pyrimidine Biosynthesis	-	-	1	4	5	-	10 (19.2%)
Pyruvate Metabolism	8	1	2	1	-	-	12 (92.3%)

Riboflavin Metabolism	-	-	-	1	-	-	1 (7.1%)
Sphingolipid Biosynthesis	58	-	-	-	-	-	58 (100%)
Sterol Biosynthesis	31	-	-	-	-	-	31 (100%)
Sterol Metabolism	10	3	1	-	-	-	14 (100%)
TAG Decomposition	3	-	-	-	-	-	3 (100%)
Threonine and Lysine Metabolism	-	-	-	1	-	-	1 (5.3%)
Tyrosine, Tryptophan, and Phenylalanine Metabolism	-	-	1	4	1	-	6 (13.6%)
Valine, Leucine, and Isoleucine Metabolism	-	-	-	3	-	-	3 (15.8%)
Other	-	-	-	-			2 (20%)
Biomass Synthesis	35	-	-	-	-	-	36(*) (100%)
Lumped Reactions	-	-	-	-	-	-	37(**) (n/a)
Pooling Reactions	21	-	-	-	-	-	21 (100%)
Exchange Reactions	-	-	-	-	60	-	60 (36.4%)
Transport Reactions	-	187	21	4	148	-	360 (59.6%)

Generated lumped reactions

After the network expansion, we generated lumped reactions connecting the required BBBs to ensure their adequate production for the desired amount of growth. The biomass composition, as defined in iMM904, is comprised of 42 BBBs, and 14 of those could be sufficiently produced by our generated core network. Therefore, we needed to generate associated lumped reactions for 28 BBBs. As mentioned previously, lumpGEM computes the minimal set of reactions (called a subnetwork) that need to be added to the core network to produce a target BBB, which are then lumped into one reaction and added to the core network. For each of these subnetworks, all alternative subnetworks of the same size were also computed, to allow for flexibility of the network in terms of biosynthetic routes. In total, 38 lumped reactions were computed that corresponded to lumped subnetworks of various numbers of reactions. A detailed report on the number of generated lumped reactions per BBB and the size of the computed minimal subnetworks can be found in Table 3.

At this point of the workflow, we made several interesting observations about the ability of the core network to produce several BBBs. Even though their biosynthetic routes were explicitly present in the model, two BBBs, namely PC and ergosterol, could not be produced by the core network. This production was hindered by the lack of adenosyl-methionine, which the core could synthesize from adenosine triphosphate (ATP) and methionine by methionine adenosyltransferase, though methionine was another BBB that could not be produced by the core network. Using lumpGEM, we estimated the minimal set of reactions that we would need to add to the model to enable the production of methionine. Two alternative subnetworks were computed, each consisting of 11 reactions. It is noteworthy that in both subnetworks, the algorithm

computes the most efficient methionine pathway to be the textbook biosynthetic route from aspartate. This additionally serves as an excellent validation point: the algorithm will always compute the most efficient biosynthetic pathways, which should be -and are- the physiologically observed ones. As seen in Figure 1, this pathway converts aspartate to homoserine, followed by homocysteine, which will finally be converted to methionine (Mountain, et al. 1991). The subnetworks also include methionine biosynthesis through sulfate assimilation (Thomas, et al. 1992), with a few extra reactions included for mass balancing. The only difference between the two alternative subnetworks lies in the dehydrogenation of L-aspartate semialdehyde to homoserine; this reaction can use NADH or NADPH as a cofactor (model reactions HSDxi and HSDyi, respectively).

Interestingly, the lumped reaction computed for the synthesis of methionine was still insufficient for the production of ergosterol and PC because, in both of these pathways, adenosyl homocysteine was produced though not consumed by any other reaction of the core. Therefore, additional lumped reactions needed to be generated to remove this product and mass balance the two pathways. It just so happens that the minimal subnetworks required for both of these cases produced methionine and were identical and unique. This subnetwork consisted of 4 reactions, which is considerably smaller size than the 11-reaction methionine subnetworks. Furthermore, since ergosterol and PC share the same subnetwork, the computed lumped reaction only needed to be added to the model once, resulting in the addition of 37 lumped reactions. Finally, these pathways can be observed graphically in Figure 2 and Figure 3. Figs 1,2 and 3 were created using the Escher web application (King, et al. 2015).

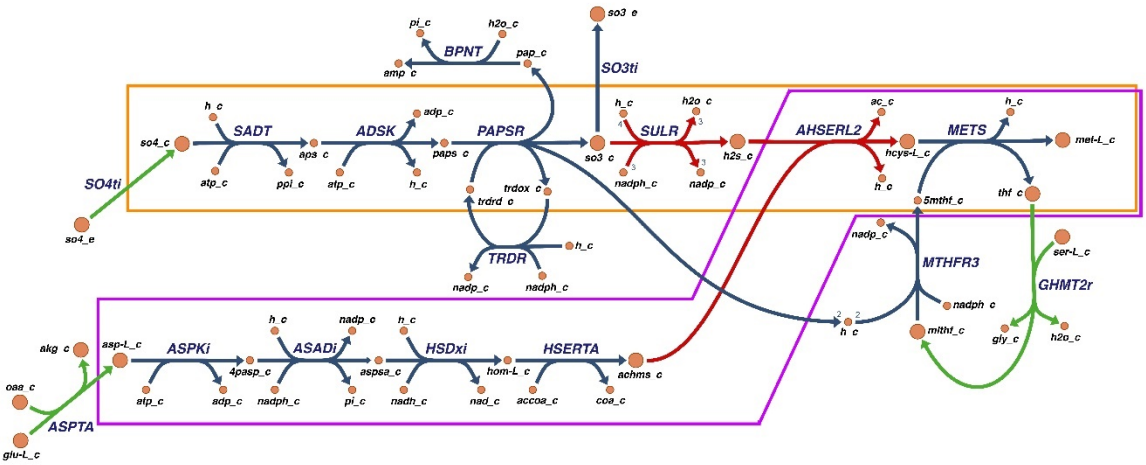
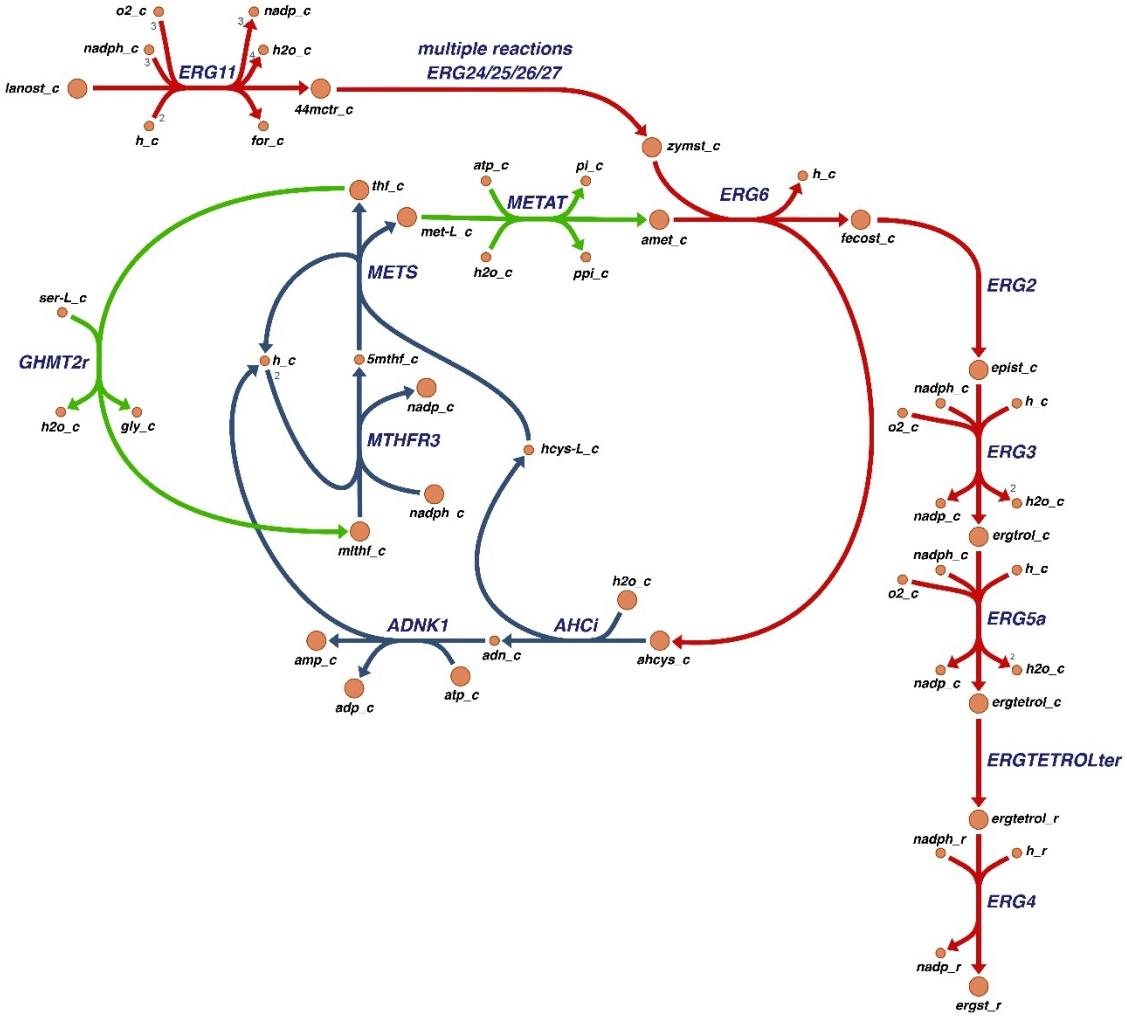


Figure 1. The L-methionine minimal subnetwork (in blue). The purple box highlights the textbook methionine biosynthetic route starting from aspartate. The orange box highlights the sulfate assimilation pathway for methionine biosynthesis. Reactions in red are part of the core network and part of the biosynthetic routes. Reactions in green are part of the core network but not part of the biosynthetic routes and serve the mass balancing of the subnetwork.



481

482 *Figure 2. The ergosterol minimal subnetwork (in blue). Reactions in red are part of the core network*
483 *and part of the biosynthetic route. Reactions in green are part of the core network but not part of the*
484 *biosynthetic route, serving instead as the mass balance for the subnetwork.*

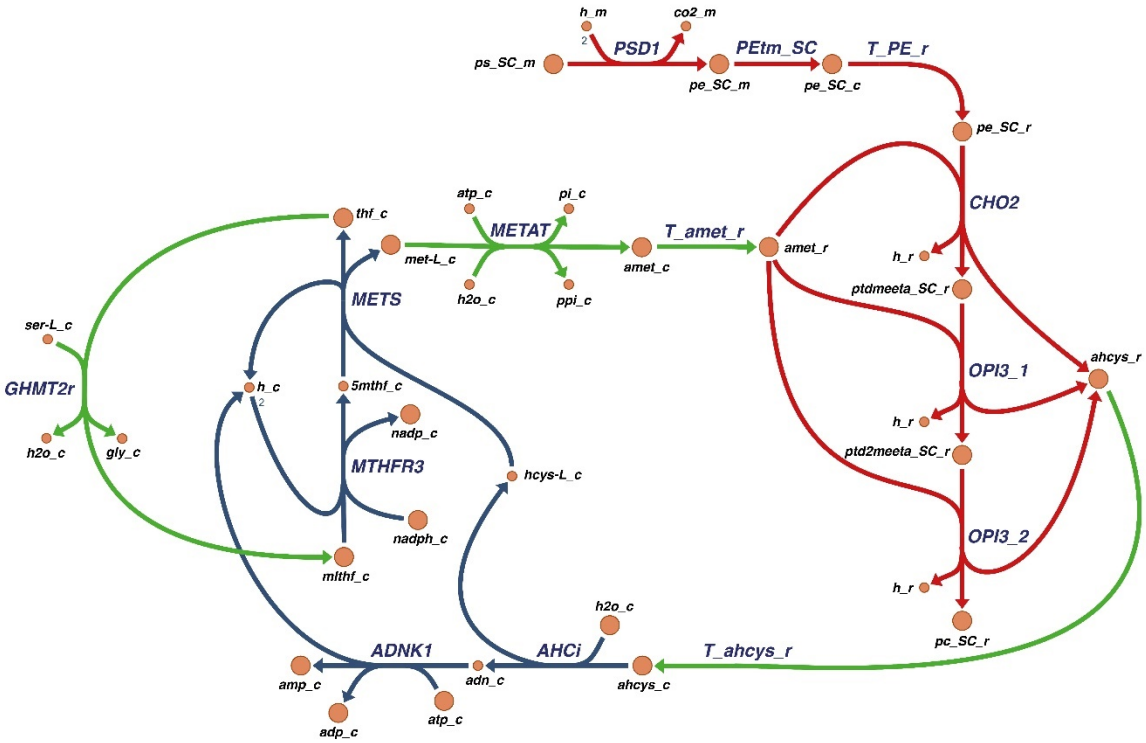


Figure 3. The phosphatidylcholine (PC) minimal subnetwork (in blue). Reactions in red are part of the core network and part of the biosynthetic route. Reactions in green are part of the core network but not part of the biosynthetic route, serving instead as the mass balance for the subnetwork.

Thermodynamics

The thermodynamic curation of redLips stemmed from the curation of the integrated GEM, as described in the Materials and Methods section. We used the same data to ensure that all reactions in the network are thermodynamically feasible, by imposing the relevant physiological constraints, The compounds whose properties could not be computed contain an acyl-carrier protein (ACP) molecule, which is a large and complicated molecule with a stereochemical structure that cannot be computed by GCM, as well as other related or bound species. The coverage of our database amounts to 89% of the metabolites of redLips, meaning 85.5% of the $\Delta_r G'^o$ values for the network reactions could be computed. These computations included the relaxation of 62 thermodynamic constraints, as described in the Materials and Methods section. The complete thermodynamic curation data can be found in Supplementary Table S3.

Table 3. Biomass building blocks for iMM904, the size of the subnetworks generated by lumpGEM, and the corresponding number of lumped reactions. (*)Produced by the core network.

Biomass Building Block	Size of Subnetwork	# of generated lumped reactions
1,3-beta-D-Glucan	3	1
AMP	10	1
L-Arginine	7	1
L-Asparagine	1	1
CMP	7	1
L-Cysteine	5	1
dAMP	13	1
dCMP	10	1
dGMP	15	2
dTMP	12	2
Ergosterol	4	1
Glycogen	3	1
GMP	12	1
L-Histidine	12	1
L-Isoleucine	9	4
L-Leucine	8	1
L-Lysine	8	2
L-Methionine	11	2
Phosphatidylcholine	4	1
L-Phenylalanine	7	1
L-Proline	4	2
Riboflavin	18	1
L-Threonine	5	2
Trehalose	2	1
L-Tryptophan	9	1
L-Tyrosine	7	2
UMP	6	1
L-Valine	3	1
Glycine	(*)	(*)
L-Alanine	(*)	(*)
L-Aspartate	(*)	(*)
L-Glutamate	(*)	(*)
L-Glutamine	(*)	(*)
L-Serine	(*)	(*)
Mannan	(*)	(*)
Phosphatidate	(*)	(*)
Sulfate	(*)	(*)
Phosphatidyl-1D-myo-inositol	(*)	(*)
Phosphatidylethanolamine	(*)	(*)
Phosphatidylserine	(*)	(*)
Triglyceride	(*)	(*)
Zymosterol	(*)	(*)

Gene Essentiality Analysis and Comparison

redLips was curated to include gene-reaction relationships in the form of logical rules. These rules were assigned through an exhaustive search in other yeast GEMs, majorly iIN800 (Nookaew, et al. 2008) and iMM904 (Mo, et al. 2009, Zomorrodi and Maranas 2010), and in literature through the Saccharomyces Genome Database (SGD, https://www.yeastgenome.org, (Cherry, et al. 1998)). Available experimental evidence for gene essentiality were gathered from literature through the Phenotype repository of SGD. All the genes whose deletions would result in inviability or auxotrophy beyond our defined media were classified as essential.

In order to benchmark and evaluate the performance of our model, we performed gene essentiality analysis for redLips and iMM904, for single-gene knockouts, and compared the results. The detailed predictions for each of the models are available in Supplementary Table S4. redLips encompasses 459 genes opposed to 905 for iMM904. Out of these, 439 are common between the two models (Figure 4a). The 20 genes that are part of redLips and not iMM904 are all encoding enzymes catalyzing lipid related reactions.

redLips predicted correctly 50 genes as essential (true positive) and 372 genes as non-essential (true negative). Nine genes were predicted falsely as essential (false positive) and 28 as non-essential (false negative) (Figure 4b). Out of the 28 false negative predictions, one gene, namely YJL097W (PHS1) is not part of the iMM904 gene annotation. PHS1 encodes the enzyme that catalyzes the elongation of very long chain fatty acids, which are then used as building blocks for complex sphingolipids. The false negative prediction occurred because of the definition of biomass composition requirements in the model. Sphingolipids were not considered as BBBs, thus their formation, or lack thereof, does not affect the predicted growth. Moreover, among the rest of the false negative predictions we identified four genes, namely YBR265W (TSC10), YDL015C (TSC13), YKL004W (AUR1), and YMR296C (LCB1), which all are essential for sphingolipid production, either directly or through the metabolism of very long chain fatty acids. Similarly, the YMR013C (SEC59) gene was a false negative prediction because dolichol species were not considered in the biomass composition.

Interestingly, by imposing a minimum flux value on at least three of the defined biosynthetic reactions for lipid species (namely Bipc_a, Bmipc_a, and Bdolp), it was possible to attain the correct prediction of true positive for four out of five of these genes, with the exception of LCB1 (Figure 4c). The imposed flux value was equal to 10^{-6} mmol·gDW^{-1}·h^{-1}, which corresponds to the smallest BBB flux contributing to biomass in the network. This shows the significance and value of the addition of these reactions and highlights the importance of a consistently defined cell lipid composition. When comparing the results of the two models for their common genes (Figure 4d), redLips performed better in predicting experimentally essential genes; iMM904 predicted 45 true positives which were a subset of the 50 predicted by redLips. The five additional genes encode enzymes which catalyze reactions belonging to lipid pathways, serine metabolism and glycolysis (Figure 4f). Correspondingly, the 27 false

negative genes for redLips were a subset of iMM904's 32. Four of them could be turned into true positives if the sphingolipid requirements of the biomass were modified as mentioned previously. Other false negative genes included YDR208W (MSS4), YLR240W (VSP34), and YNL267W (PIK1), all related to PI synthesis. This part of the network is fairly complex and in possession of multiple alternative biosynthetic reaction routes. Furthermore, PI derivative species have been known to be especially active in signaling and membrane trafficking (Downes, et al. 2005, Krauss and Haucke 2007), the mechanisms of which are either unknown or not included in the model.

redLips and iMM904 each predicted nine false positive genes, four of which were common. In redLips, these predictions occurred mostly due to alternative pathways missing from the network; while this was expected due to redLips being a reduced model, about half of these genes were false positives for iMM904 as well (Figure 4e). The rest of the false negatives stemmed from the inclusion of ergosterol in the biomass composition of the model. While ergosterol is essential to yeast cells, mutants incapable of synthesizing it are viable by accumulating ergosterol precursors in their membranes (Kato and Wickner 2001, Liu, et al. 2017), an effect that was not included in either of the models.

In conclusion, the gene essentiality analysis and comparison of redLips opposite iMM904 showcases the ability of redLips to make accurate predictions, and in most cases performing better than the GEM. Genes that are not part of iMM904's annotation were included in the gene-reaction relationships of redLips, and the vast majority of genes that encode enzymes which catalyze lipid related reactions were predicted correctly as essential or non-essential.

Accepted manuscript to FEMS Yeast Research, Volume 20, Issue 2, March 2020, foaa006, https://doi.org/10.1093/femsyr/foaa006 Published: 18 February 2020

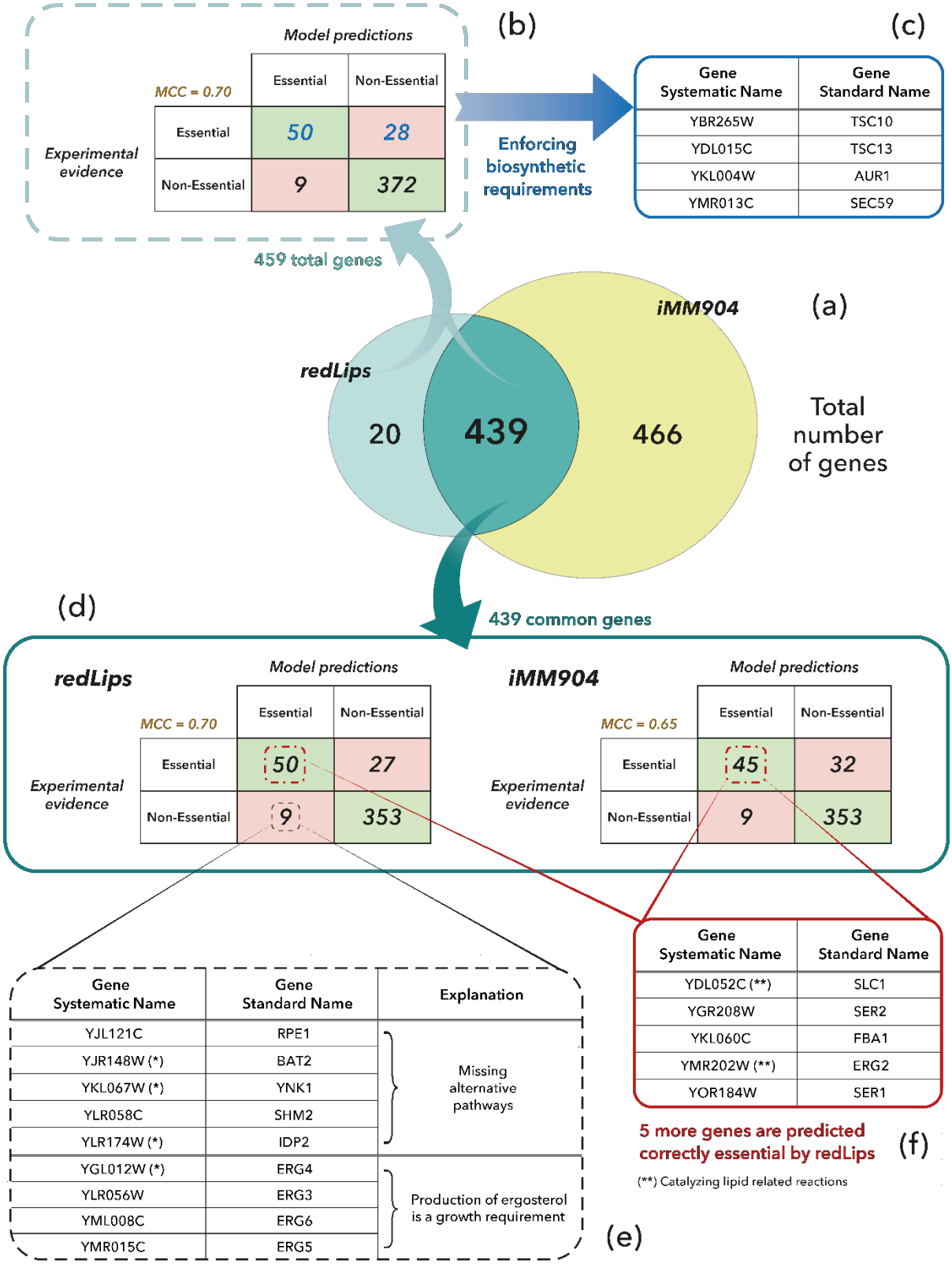


Figure 4. (a) Venn diagram of the genes included in redLips and iMM904. (b) Gene essentiality analysis in redLips and comparison with experimental evidence. (c) Improvements that can made to the predictions by enforcing lipid biosynthetic requirements. (d) Gene essentiality analysis in both redLips and iMM904 for the enzymes they have in common; and comparison with experimental evidence. (e)

The nine genes that correspond to false positive predictions of redLips and explanations of the occurrence. (e) The five true positive predicted genes of redLips that iMM904 predicts falsely negative. The Matthews Correlation Coefficient (MCC)(Matthews 1975) is also reported for each case.

Comparison with yeast GEMs

Comparing redLips with the available yeast GEMs provided insight into the network in terms of comprehensiveness and pathway connectivity. As mentioned previously, our model includes all the major lipid species and respective biosynthetic pathways.

One common difference between models that can lead to very dissimilar numbers in reactions and species is the considered fatty acyl chains. As stated in the Materials & Methods section, glycerophospholipids possess two fatty acyl chains (four in the case of CL). In our model, we defined a fatty acid pool that participates in the formation of these species, thus considering only one metabolite of each class with an attached, generic fatty acyl (the assumed length is still C16:0 for thermodynamic calculations). Some other GEMs, namely Yeast7 and Yeast8, consider four individual fatty acid species as reactants in these biotransformations: C16:0, C18:0, C16:1, and C18:1. Naturally, this leads to a very large combinatory number of reactions and species. For example, if we consider a species with two fatty acyl tails, there are ten possible combinations leading to ten model metabolites. This number can grow exponentially if one considers the large number of fatty acyl derivatives and remodeling reactions. The same issue arises with fatty acyl-CoAs. Since the aim of redLips is to be used as a scaffold for omics integration and its nature is not context limited to specific species or pathways in order to enable versatility, this is a pitfall we aimed to avoid in order to preserve a concise representation. As mentioned earlier, a model including all the combinatoric occurrences will be difficult to curate and handle. It is important to note that, in the case of available experimental data or focused studies, separately considering these species can be beneficial for the accuracy of predictions (Sanchez, et al. 2019), and should be taken into account by expanding the associated parts of the network species and reactions accordingly. To this end, one can make use of the lipid pools we have defined and follow the same procedure as we did for the expansion of previously defined lumped/pooled reactions of iMM904, as described in the Materials and Methods section.

Another difference between redLips and other models is that some other models, for example iIN800 (Nookaew, et al. 2008), include multiple identical reactions if their respective associated enzymes are encoded by multiple genes or in cases of multiple enzyme paralogs. This practice is acceptable according to genome annotation, but it leads to misleading computations; the mechanistic representation of the network calculates the net flux through each reaction irrespective of which enzyme is catalyzing it. This means that the resulting net flux value for this particular reaction will be the sum of all the discreet flux values for each reaction copy. To resolve this point, we considered only unique reaction occurrences in redLips that represent the net reaction rate for each biotransformation. The exception to this rule is the case where the same

biotransformation occurs in different cellular compartments. Since the metabolites in each compartment are modeled separately, this consideration does not result in duplicate reactions—in mathematical terms, the stoichiometric matrix will not have duplicate columns. In circumstances where enzymatic or kinetic properties are relevant for a study and require a separate consideration for these instances, the model can simply be modified to incorporate them.

We also present a detailed comparison of our network to the other yeast networks in Table 4. Included in this table are the number of lipid-related reactions, species, and cellular compartments considered in each model. To ensure accurate comparability, we curated the number of reactions and species of interest for all considered models. The criteria we used were as follows: (i) We considered only one generic instance of metabolites possessing one or more fatty acyl chains. This applies both to species and reactions. As discussed previously, each model considers a different number of fatty acyl chain lengths, and in combination, this can lead to misleadingly different statistics. (ii) We didn't consider metabolite pools or pooling reactions. Similarly, each model considers various diverse metabolite pools that can be heavily connected to the network by a large number of pooling reactions. (iii) We didn't consider duplicate reactions unless they occurred in different compartments. (iv) We didn't consider transport and boundary reactions. Since each model considers a different number of cellular compartments, the number of transport reactions varies accordingly. (v) We didn't consider disconnected reactions; there were rare occurrences of reactions in which both reactants and products did not participate in any other reaction in the network. These reactions serve for annotation purposes and most probably will be gap-filled in the future, but they do not contribute to the functionality of the model.

Using this comparison, we can see in Table 4 that redLips covers at least as many species as the other GEMs and more reactions than most of them. The major differences in the non-curated numbers of species and reactions can be attributed to the reasons listed above as well as the number of compartments of each model. If we go through the reactions per pathway, we can see that the majority of differences stem from the biosynthetic routes for PI derivatives, such as glycosyl-phosphatidylinositol (GPI) anchors for proteins and inositol and PI polyphosphates. These molecules play a major role in cell signaling, which was beyond the scope of redLips at this time. Signaling cascades in lipid metabolism is a vast area of study on its own, and we feel that it would be best served with a dedicated model. Another difference, especially concerning the Yeast7 and Yeast8 models, was in phospholipid biosynthesis. The larger numbers in these models are due to the consideration of five additional cellular compartments, including membranes, and the assignment of reactions occurring in more than one of them.

Accepted manuscript to FEMS Yeast Research, Volume 20, Issue 2, March 2020, foaa006, https://doi.org/10.1093/femsyr/foaa006 Published: 18 February 2020

Table 4. Detailed comparison between redLips and other yeast GEMs in terms of reactions, species, and cellular compartments. The curated numbers of lipid reactions and species are given in parentheses next to the non-curated numbers if these numbers differ.

Model :	**redLips**	**iMM904** (Mo, et al. 2009)	**iIN800** (Heavner and Price 2015, Nookaew, et al. 2008)	**iTO977** (Heavner and Price 2015, Osterlund, et al. 2013)	**Yeast7** (Aung, et al. 2013)	**Yeast8** (Lu, et al. 2019)
# of lipid reactions (Curated # of lipid reactions)	481 (451)	291 (264)	348 (281)	331 (324)	1758 (442)	1787 (460)
Cardiolipin Biosynthesis	7	4	2	3	270 (4)	270 (4)
Carnitine Shuttle	4	3	1	1	3	3
Fatty Acid Biosynthesis	67	33	71 (63)	64	25	25
Fatty Acid Biosynthesis Mitochondrial	39	13	34	34	15	15
Fatty Acid Elongation	28	6	33 (28)	28	30	30
Fatty Acid Degradation	99	51	46 (43)	47	100	100
Glycerolipid Metabolism	7	9	15 (14)	7	29 (14)	31 (16)
Glycoprotein Metabolism	5	7	15 (4)	4	3	3
GPI Biosynthesis	0	0	0	8	9	19
Isoprenoid Biosynthesis	19	1	0	0	22	22
Mevalonate Pathway	14	14	13 (12)	11	14	14
Phospholipid Biosynthesis	60	50	51 (35)	44	686 (78)	690 (82)
PI Signaling System	0	0	0	8	15	15
Sphingolipid Biosynthesis	58	63 (36)	36 (25)	25	257 (63)	257 (63)
Sterol Metabolism	41	36	22 (19)	25 (23)	43 (32)	43 (32)
TAG Decomposition	3	1	4 (1)	1	48 (4)	48 (4)
Other	0	0	0	15	11	13
Pooling Reactions	30	0	5	5	178	189
# of lipid species (Curated # of lipid species)	241 (237)	156 (143)	184 (178)	231 (220)	500 (233)	523 (237)
# of involved compartments	7 + ex	7 + ex	2 + ex	3 + ex	12 + ex	12 + ex

Conclusions

In conclusion, redLips is a metabolic model that captures the complexity of lipid metabolism by preserving and uniting the vast majority of known lipid reactions and pathways while avoiding the pitfall of excessive—and often times redundant—detail. It was created by gathering, merging, and upgrading existing lipid metabolic pathways, integrating them into the iMM904 GEM of *S. cerevisiae* and subsequently reducing this model around the major lipid-related subsystems using the redGEM and lumpGEM frameworks. Additionally, it is consistent with the organism biochemistry as well as thermodynamic principles and can be further constrained through lipidomics measurements, applied both as flux and concentration bounds. redLips could be used as a concise platform for studying lipid metabolism across different species, and is a valuable tool for health or industry related research. We believe that this model will continue to accommodate future discoveries through the incorporation of new reactions and species as well as providing a coherent base to link cell signaling routes and building kinetic models.

Conflict of Interest

The authors declare no financial or commercial conflict of interest.

Funding

S.T. was supported by the Swiss National Science Foundation grant [315230_163423]. V.H. was supported by the Ecole Polytechnique Fédérale de Lausanne (EPFL).

Supplementary Material

Table S1. Enzyme essentiality consistency check between the integrated GEM and redLips.

Table S2. Reaction variability consistency check between the integrated GEM and redLips. (Excel)

Table S3. List of reactions and thermodynamic information for network metabolites and reactions. (Excel)

Table S4. Gene essentiality analysis comparison between redLips and iMM904. (Excel)

File S1. The full redLips metabolic network map (.json)

Accepted manuscript to FEMS Yeast Research, Volume 20, Issue 2, March 2020, foaa006, https://doi.org/10.1093/femsyr/foaa006 Published: 18 February 2020

704 File S2. The redLips model (.mat)

705 File S3. The redLips model in SBML format – without thermodynamic constraints (.xml)

706

References

Ataman M. Navigating and Managing the Complexity of Genome Scale Metabolic Networks for Studies in Cellular Physiology and Industrial Biotechnology*Chemistry and Chemical Engineering* volume PhD. Lausanne: EPFL, 2016, 178.

Ataman M, Gardiol DFH, Fengos G *et al.* redGEM: Systematic reduction and analysis of genome-scale metabolic reconstructions for development of consistent core metabolic models. *Plos Computational Biology* 2017;**13**.

Ataman M, Hatzimanikatis V. lumpGEM: Systematic generation of subnetworks and elementally balanced lumped reactions for the biosynthesis of target metabolites. *Plos Computational Biology* 2017;**13**.

Aung HW, Henry SA, Walker LP. Revising the Representation of Fatty Acid, Glycerolipid, and Glycerophospholipid Metabolism in the Consensus Model of Yeast Metabolism. *Ind Biotechnol (New Rochelle N Y)* 2013;**9**: 215-28.

Cherry JM, Adler C, Ball C *et al.* SGD: Saccharomyces Genome Database. *Nucleic Acids Res* 1998;**26**: 73-9.

Cohen AE, Venkatachalam V. Bringing bioelectricity to light. *Annu Rev Biophys* 2014;**43**: 211-32.

da Silveira Dos Santos AX, Riezman I, Aguilera-Romero MA *et al.* Systematic lipidomic analysis of yeast protein kinase and phosphatase mutants reveals novel insights into regulation of lipid homeostasis. *Mol Biol Cell* 2014;**25**: 3234-46.

Daum G, Tuller G, Nemec T *et al.* Systematic analysis of yeast strains with possible defects in lipid metabolism. *Yeast* 1999;**15**: 601-14.

de Kroon AIPM. Lipidomics in research on yeast membrane lipid homeostasis. *Bba-Mol Cell Biol L* 2017;**1862**: 797-9.

Dickinson JR, Schweizer M. *Metabolism and Molecular Physiology of Saccharomyces Cerevisiae*: CRC Press, 2004.

Emanuelsson O, Brunak S, von Heijne G *et al.* Locating proteins in the cell using TargetP, SignalP and related tools. *Nature Protocols* 2007;**2**: 953-71.

Fahy E, Subramaniam S, Brown HA *et al.* A comprehensive classification system for lipids. *J Lipid Res* 2005;**46**: 839-61.

Fahy E, Subramaniam S, Murphy RC *et al.* Update of the LIPID MAPS comprehensive classification system for lipids. *J Lipid Res* 2009;**50**: S9-S14.

Forster J, Famili I, Fu P *et al.* Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network. *Genome Res* 2003;**13**: 244-53.

Guan XL, Souza CM, Pichler H *et al.* Functional Interactions between Sphingolipids and Sterols in Biological Membranes Regulating Cell Physiology. *Mol Biol Cell* 2009;**20**: 2083-95.

Hadadi N, Cher Soh K, Seijo M *et al.* A computational framework for integration of lipidomics data into metabolic pathways. *Metab Eng* 2014;**23**: 1-8.

Han XL. Lipidomics for studying metabolism. *Nat Rev Endocrinol* 2016;**12**: 668-79.

Harayama T, Riezman H. Understanding the diversity of membrane lipid composition. *Nature Reviews Molecular Cell Biology* 2018;**19**: 281-96.

Holthuis JC, Menon AK. Lipid landscapes and pipelines in membrane homeostasis. *Nature* 2014;**510**: 48-57.

Ivanova PT, Milne SB, Myers DS *et al.* Lipidomics: a mass spectrometry based systems level analysis of cellular lipids. *Curr Opin Chem Biol* 2009;**13**: 526-31.

753 Joyce AR, Palsson BO. The model organism as a system: integrating 'omics' data sets. *Nature*
754 *Reviews Molecular Cell Biology* 2006;**7**: 198-210.
755 Klose C, Surma MA, Gerl MJ *et al.* Flexibility of a Eukaryotic Lipidome - Insights from Yeast
756 Lipidomics. *Plos One* 2012;**7**.
757 Kontush A, Chapman MJ. Lipidomics as a Tool for the Study of Lipoprotein Metabolism. *Curr*
758 *Atheroscler Rep* 2010;**12**: 194-201.
759 Lopes H, Rocha I. Genome-scale modeling of yeast: chronology, applications and critical
760 perspectives. *Fems Yeast Research* 2017;**17**.
761 Lu HZ, Li FR, Sanchez BJ *et al.* A consensus S. cerevisiae metabolic model Yeast8 and its
762 ecosystem for comprehensively probing cellular metabolism. *Nature*
763 *Communications* 2019;**10**.
764 Ataman M. Navigating and Managing the Complexity of Genome Scale Metabolic Networks
765 for Studies in Cellular Physiology and Industrial Biotechnology*Chemistry and*
766 *Chemical Engineering* volume PhD. Lausanne: EPFL, 2016, 178.
767 Ataman M, Gardiol DFH, Fengos G *et al.* redGEM: Systematic reduction and analysis of
768 genome-scale metabolic reconstructions for development of consistent core
769 metabolic models. *Plos Computational Biology* 2017;**13**.
770 Ataman M, Hatzimanikatis V. lumpGEM: Systematic generation of subnetworks and
771 elementally balanced lumped reactions for the biosynthesis of target metabolites.
772 *Plos Computational Biology* 2017;**13**.
773 Aung HW, Henry SA, Walker LP. Revising the Representation of Fatty Acid, Glycerolipid, and
774 Glycerophospholipid Metabolism in the Consensus Model of Yeast Metabolism. *Ind*
775 *Biotechnol (New Rochelle N Y)* 2013;**9**: 215-28.
776 Cherry JM, Adler C, Ball C *et al.* SGD: Saccharomyces Genome Database. *Nucleic Acids*
777 *Research* 1998;**26**: 73-9.
778 Cohen AE, Venkatachalam V. Bringing bioelectricity to light. *Annu Rev Biophys* 2014;**43**: 211-
779 32.
780 da Silveira Dos Santos AX, Riezman I, Aguilera-Romero MA *et al.* Systematic lipidomic
781 analysis of yeast protein kinase and phosphatase mutants reveals novel insights into
782 regulation of lipid homeostasis. *Mol Biol Cell* 2014;**25**: 3234-46.
783 Daum G, Tuller G, Nemec T *et al.* Systematic analysis of yeast strains with possible defects in
784 lipid metabolism. *Yeast* 1999;**15**: 601-14.
785 de Kroon AIPM. Lipidomics in research on yeast membrane lipid homeostasis. *Bba-Mol Cell*
786 *Biol L* 2017;**1862**: 797-9.
787 Dickinson JR, Schweizer M. *Metabolism and Molecular Physiology of Saccharomyces*
788 *Cerevisiae*: CRC Press, 2004.
789 Downes CP, Gray A, Lucocq JM. Probing phosphoinositide functions in signaling and
790 membrane trafficking. *Trends in cell biology* 2005;**15**: 259-68.
791 Emanuelsson O, Brunak S, von Heijne G *et al.* Locating proteins in the cell using TargetP,
792 SignalP and related tools. *Nat Protoc* 2007;**2**: 953-71.
793 Fahy E, Subramaniam S, Brown HA *et al.* A comprehensive classification system for lipids. *J*
794 *Lipid Res* 2005;**46**: 839-61.
795 Fahy E, Subramaniam S, Murphy RC *et al.* Update of the LIPID MAPS comprehensive
796 classification system for lipids. *J Lipid Res* 2009;**50**: S9-S14.
797 Forster J, Famili I, Fu P *et al.* Genome-scale reconstruction of the Saccharomyces cerevisiae
798 metabolic network. *Genome Res* 2003;**13**: 244-53.

Guan XL, Souza CM, Pichler H *et al.* Functional Interactions between Sphingolipids and Sterols in Biological Membranes Regulating Cell Physiology. *Mol Biol Cell* 2009;**20**: 2083-95.

Hadadi N, Cher Soh K, Seijo M *et al.* A computational framework for integration of lipidomics data into metabolic pathways. *Metab Eng* 2014;**23**: 1-8.

Han XL. Lipidomics for studying metabolism. *Nat Rev Endocrinol* 2016;**12**: 668-79.

Harayama T, Riezman H. Understanding the diversity of membrane lipid composition. *Nat Rev Mol Cell Bio* 2018;**19**: 281-96.

Heavner BD, Price ND. Comparative Analysis of Yeast Metabolic Network Models Highlights Progress, Opportunities for Metabolic Reconstruction. *PLoS Comput Biol* 2015;**11**: e1004530.

Holthuis JC, Menon AK. Lipid landscapes and pipelines in membrane homeostasis. *Nature* 2014;**510**: 48-57.

Ivanova PT, Milne SB, Myers DS *et al.* Lipidomics: a mass spectrometry based systems level analysis of cellular lipids. *Curr Opin Chem Biol* 2009;**13**: 526-31.

Joyce AR, Palsson BO. The model organism as a system: integrating 'omics' data sets. *Nat Rev Mol Cell Bio* 2006;**7**: 198-210.

Kato M, Wickner W. Ergosterol is required for the Sec18/ATP-dependent priming step of homotypic vacuole fusion. *The EMBO journal* 2001;**20**: 4035-40.

King ZA, Drager A, Ebrahim A *et al.* Escher: A Web Application for Building, Sharing, and Embedding Data-Rich Visualizations of Biological Pathways. *PLoS Comput Biol* 2015;**11**: e1004321.

Klose C, Surma MA, Gerl MJ *et al.* Flexibility of a Eukaryotic Lipidome - Insights from Yeast Lipidomics. *Plos One* 2012;**7**.

Kontush A, Chapman MJ. Lipidomics as a Tool for the Study of Lipoprotein Metabolism. *Curr Atheroscler Rep* 2010;**12**: 194-201.

Krauss M, Haucke V. Phosphoinositide-metabolizing enzymes at the interface between membrane traffic and cell signalling. *EMBO reports* 2007;**8**: 241-6.

Liu G, Chen Y, Faergeman NJ *et al.* Elimination of the last reactions in ergosterol biosynthesis alters the resistance of Saccharomyces cerevisiae to multiple stresses. *Fems Yeast Res* 2017;**17**.

Lopes H, Rocha I. Genome-scale modeling of yeast: chronology, applications and critical perspectives. *Fems Yeast Research* 2017;**17**.

Lu HZ, Li FR, Sanchez BJ *et al.* A consensus S. cerevisiae metabolic model Yeast8 and its ecosystem for comprehensively probing cellular metabolism. *Nature Communications* 2019;**10**.

Matthews BW. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim Biophys Acta* 1975;**405**: 442-51.

Mavrovouniotis ML. Group Contributions for Estimating Standard Gibbs Energies of Formation of Biochemical-Compounds in Aqueous-Solution. *Biotechnol Bioeng* 1990;**36**: 1070-82.

Mavrovouniotis ML. Estimation of Standard Gibbs Energy Changes of Biotransformations. *J Biol Chem* 1991;**266**: 14440-5.

Mo ML, Palsson BO, Herrgard MJ. Connecting extracellular metabolomic measurements to intracellular flux states in yeast. *Bmc Systems Biology* 2009;**3**.

Mountain HA, Bystrom AS, Larsen JT *et al.* Four major transcriptional responses in the methionine/threonine biosynthetic pathway of Saccharomyces cerevisiae. *Yeast* 1991;**7**: 781-803.

Nielsen J. Systems biology of lipid metabolism: From yeast to human. *Febs Lett* 2009;**583**: 3905-13.

Nookaew I, Jewett MC, Meechai A *et al.* The genome-scale metabolic model iIN800 of Saccharomyces cerevisiae and its validation: a scaffold to query lipid metabolism. *BMC Syst Biol* 2008;**2**: 71.

Orij R, Postmus J, Ter Beek A *et al.* In vivo measurement of cytosolic and mitochondrial pH using a pH-sensitive GFP derivative in Saccharomyces cerevisiae reveals a relation between intracellular pH and growth. *Microbiology* 2009;**155**: 268-78.

Orij R, Urbanus ML, Vizeacoumar FJ *et al.* Genome-wide analysis of intracellular pH reveals quantitative control of cell division rate by pH(c) in Saccharomyces cerevisiae. *Genome Biol* 2012;**13**.

Osterlund T, Nookaew I, Bordel S *et al.* Mapping condition-dependent regulation of metabolism in yeast through genome-scale modeling. *BMC Syst Biol* 2013;**7**: 36.

Osterlund T, Nookaew I, Nielsen J. Fifteen years of large scale metabolic modeling of yeast: Developments and impacts. *Biotechnol Adv* 2012;**30**: 979-88.

Paroutis P, Touret N, Grinstein S. The pH of the secretory pathway: measurement, determinants, and regulation. *Physiology (Bethesda)* 2004;**19**: 207-15.

Petranovic D, Tyo K, Vemuri GN *et al.* Prospects of yeast systems biology for human health: integrating lipid, protein and energy metabolism. *Fems Yeast Res* 2010;**10**: 1046-59.

Preston RA, Murphy RF, Jones EW. Assay of vacuolar pH in yeast and identification of acidification-defective mutants. *Proc Natl Acad Sci U S A* 1989;**86**: 7027-31.

Salvy P, Fengos G, Ataman M *et al.* pyTFA and matTFA: a Python package and a Matlab toolbox for Thermodynamics-based Flux Analysis. *Bioinformatics* 2019;**35**: 167-9.

Sanchez BJ, Li F, Kerkhoven EJ *et al.* SLIMEr: probing flexibility of lipid metabolism in yeast with an improved constraint-based modeling framework. *BMC Syst Biol* 2019;**13**: 4.

Sanchez BJ, Nielsen J. Genome scale models of yeast: towards standardized evaluation and consistent omic integration. *Integr Biol-Uk* 2015;**7**: 846-58.

Santos AXS, Riezman H. Yeast as a model system for studying lipid homeostasis and function. *Febs Lett* 2012;**586**: 2858-67.

Savoglidis G, dos Santos AXD, Riezman I *et al.* A method for analysis and design of metabolism using metabolomics data and kinetic models: Application on lipidomics using a novel kinetic model of sphingolipid metabolism. *Metab Eng* 2016;**37**: 46-62.

Schneiter R, Brugger B, Sandhoff R *et al.* Electrospray ionization tandem mass spectrometry (ESI-MS/MS) analysis of the lipid molecular species composition of yeast subcellular membranes reveals acyl chain-based sorting/remodeling of distinct molecular species en route to the plasma membrane. *J Cell Biol* 1999;**146**: 741-54.

Thomas D, Barbey R, Henry D *et al.* Physiological analysis of mutants of Saccharomyces cerevisiae impaired in sulphate assimilation. *J Gen Microbiol* 1992;**138**: 2021-8.

Wenk MR. The emerging field of lipidomics. *Nat Rev Drug Discov* 2005;**4**: 594-610.

Zomorrodi AR, Maranas CD. Improving the iMM904 S. cerevisiae metabolic model using essentiality and synthetic lethality data. *BMC Syst Biol* 2010;**4**: 178.

Mavrovouniotis ML. Group Contributions for Estimating Standard Gibbs Energies of Formation of Biochemical-Compounds in Aqueous-Solution. *Biotechnology and Bioengineering* 1990;**36**: 1070-82.

Mavrovouniotis ML. Estimation of Standard Gibbs Energy Changes of Biotransformations. *Journal of Biological Chemistry* 1991;**266**: 14440-5.

Mo ML, Palsson BO, Herrgard MJ. Connecting extracellular metabolomic measurements to intracellular flux states in yeast. *Bmc Systems Biology* 2009;**3**.

Nielsen J. Systems biology of lipid metabolism: From yeast to human. *Febs Lett* 2009;**583**: 3905-13.

Nookaew I, Jewett MC, Meechai A *et al.* The genome-scale metabolic model iIN800 of Saccharomyces cerevisiae and its validation: a scaffold to query lipid metabolism. *BMC Syst Biol* 2008;**2**: 71.

Orij R, Postmus J, Ter Beek A *et al.* In vivo measurement of cytosolic and mitochondrial pH using a pH-sensitive GFP derivative in Saccharomyces cerevisiae reveals a relation between intracellular pH and growth. *Microbiology* 2009;**155**: 268-78.

Orij R, Urbanus ML, Vizeacoumar FJ *et al.* Genome-wide analysis of intracellular pH reveals quantitative control of cell division rate by pH(c) in Saccharomyces cerevisiae. *Genome Biol* 2012;**13**.

Osterlund T, Nookaew I, Nielsen J. Fifteen years of large scale metabolic modeling of yeast: Developments and impacts. *Biotechnology Advances* 2012;**30**: 979-88.

Paroutis P, Touret N, Grinstein S. The pH of the secretory pathway: measurement, determinants, and regulation. *Physiology (Bethesda)* 2004;**19**: 207-15.

Petranovic D, Tyo K, Vemuri GN *et al.* Prospects of yeast systems biology for human health: integrating lipid, protein and energy metabolism. *Fems Yeast Research* 2010;**10**: 1046-59.

Preston RA, Murphy RF, Jones EW. Assay of vacuolar pH in yeast and identification of acidification-defective mutants. *Proc Natl Acad Sci U S A* 1989;**86**: 7027-31.

Salvy P, Fengos G, Ataman M *et al.* pyTFA and matTFA: a Python package and a Matlab toolbox for Thermodynamics-based Flux Analysis. *Bioinformatics* 2019;**35**: 167-9.

Sanchez BJ, Li F, Kerkhoven EJ *et al.* SLIMEr: probing flexibility of lipid metabolism in yeast with an improved constraint-based modeling framework. *BMC Syst Biol* 2019;**13**: 4.

Sanchez BJ, Nielsen J. Genome scale models of yeast: towards standardized evaluation and consistent omic integration. *Integr Biol-Uk* 2015;**7**: 846-58.

Santos AXS, Riezman H. Yeast as a model system for studying lipid homeostasis and function. *Febs Lett* 2012;**586**: 2858-67.

Savoglidis G, dos Santos AXD, Riezman I *et al.* A method for analysis and design of metabolism using metabolomics data and kinetic models: Application on lipidomics using a novel kinetic model of sphingolipid metabolism. *Metabolic Engineering* 2016;**37**: 46-62.

Schneiter R, Brugger B, Sandhoff R *et al.* Electrospray ionization tandem mass spectrometry (ESI-MS/MS) analysis of the lipid molecular species composition of yeast subcellular membranes reveals acyl chain-based sorting/remodeling of distinct molecular species en route to the plasma membrane. *J Cell Biol* 1999;**146**: 741-54.

Wenk MR. The emerging field of lipidomics. *Nat Rev Drug Discov* 2005;**4**: 594-610.