The genome-scale metabolic model iIN800 of Saccharomyces cerevisiae and its validation: a scaffold to query lipid metabolism

Nookaew, Intawat; Jewett, Michael Christopher; Meechai, Asawin; Thammarongtham, Chineae; Laoteng, Kobkul; Cheevadhanarak, Supapon; Nielsen, Jens; Bhumiratana, Sakarindr

Published in:
BMC Systems Biology

Link to article, DOI:
10.1186/1752-0509-2-71

Publication date:
2008

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Nookaew, I., Jewett, M. C., Meechai, A., Thammarongtham, C., Laoteng, K., Cheevadhanarak, S., ... Bhumiratana, S. (2008). The genome-scale metabolic model iIN800 of Saccharomyces cerevisiae and its validation: a scaffold to query lipid metabolism. BMC Systems Biology, 2, 71. DOI: 10.1186/1752-0509-2-71
The genome-scale metabolic model *iIN800* of *Saccharomyces cerevisiae* and its validation: a scaffold to query lipid metabolism

Intawat Nookaew1, Michael C Jewett5,6, Asawin Meechai1, Chinae Thammarongtham2, Kobkul Laoteng2, Supapon Cheevadhanarak3, Jens Nielsen*5,7 and Sakarindr Bhumiratana*1,2,4

Address: 1Department of Chemical Engineering, Faculty of Engineering, King Mongkut's University of Technology Thonburi, Bangkok 10140, Thailand, 2Biochemical Engineering and Pilot Plant Research and Development Unit, National Center for Genetic Engineering and Biotechnology (BIOTEC) at King Mongkut's University of Technology Thonburi, Bangkhuntien, Bangkok 10150, Thailand, 3School of Bioresources and Technology, King Mongkut's University of Technology Thonburi, Bangkok 10140, Thailand, 4National Science and Technology Development Agency, Ministry of Science and Technology, Thailand Science Park, Klong Luang, Pathumthani 12120, Thailand, 5Center for Microbial Biotechnology, Biocentrum, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark, 6Department of Genetics, Harvard Medical School, Boston, MA 02115, USA and 7Department of Chemical and Biological Engineering, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden

Email: Intawat Nookaew - intawat.in@gmail.com; Michael C Jewett - mcjewett@genetics.med.harvard.edu; Asawin Meechai - asawin.mee@kmutt.ac.th; Chinae Thammarongtham - chinae@biotec.or.th; Kobkul Laoteng - kobkul@biotec.or.th; Supapon Cheevadhanarak - supapon.che@kmutt.ac.th; Jens Nielsen* - nielsenj@chalmers.se; Sakarindr Bhumiratana* - sakarindr@exchange.nstda.or.th

* Corresponding authors

Published: 7 August 2008
BMC Systems Biology 2008, 2:71 doi:10.1186/1752-0509-2-71

Abstract

**Background:** Up to now, there have been three published versions of a yeast genome-scale metabolic model: *iFF708, iND750* and *iLL672*. All three models, however, lack a detailed description of lipid metabolism and thus are unable to be used as integrated scaffolds for gaining insights into lipid metabolism from multilevel omic measurement technologies (e.g. genome-wide mRNA levels). To overcome this limitation, we reconstructed a new version of the *Saccharomyces cerevisiae* genome-scale model, *iIN800* that includes a more rigorous and detailed description of lipid metabolism.

**Results:** The reconstructed metabolic model comprises 1446 reactions and 1013 metabolites. Beyond incorporating new reactions involved in lipid metabolism, we also present new biomass equations that improve the predictive power of flux balance analysis simulations. Predictions of both growth capability and large scale *in silico* single gene deletions by *iIN800* were consistent with experimental data. In addition, 13C-labeling experiments validated the new biomass equations and calculated intracellular fluxes. To demonstrate the applicability of *iIN800*, we show that the model can be used as a scaffold to reveal the regulatory importance of lipid metabolism precursors and intermediates that would have been missed in previous models from transcriptome datasets.

**Conclusion:** Performing integrated analyses using *iIN800* as a network scaffold is shown to be a valuable tool for elucidating the behavior of complex metabolic networks, particularly for identifying regulatory targets in lipid metabolism that can be used for industrial applications or for understanding lipid disease states.
Background
The yeast *Saccharomyces cerevisiae* is widely used for production of many different commercial compounds such as food, feed, beverages and pharmaceuticals [1]. It also serves as a model eukaryotic organism and has been the subject of more than 40,000 research publications [2,3]. After the complete genome sequence for yeast was released in 1996 [4], about 4,600 ORFs were characterized [3] and yeast contains many genes with human homologs [2]. This has allowed for comparative functional genomics and comparative systems biology between yeast and human. Yeast, for example, has been used to understand the function of complex metabolic pathways that are related to the development of human diseases [5-7].

Several human diseases (e.g. cancer, atherosclerosis, Alzheimer's disease, and Parkinson's disease) are associated with disorders in lipid metabolism [8-10]. The emergence of lipidomics has allowed analysis of lipid metabolism at the systems level [8,11]. Lipidomics promises to make a significant impact in our understanding of lipid related disease development [12]. As with other high-throughput techniques, however, we hypothesize that one of the main challenges for utilization of lipidome data will be our ability to develop appropriate frameworks to integrate and map data for studying relations between lipid metabolism and other cellular networks.

Previous work has shown that genome-scale metabolic models provide an excellent scaffold for integrating data into single, coherent models [13]. The calculation of Reporter Metabolites using genome-scale metabolic models is an example of how metabolic models can be used to upgrade the information content of omics data [14]. This approach allows mapping of key metabolites and reactions in large metabolic networks when combined with transcriptome [14] or metabolome data [15]. However, pathways, reactions, and genes that are not included in the metabolic network cannot be queried. Therefore, the Reporter Metabolite algorithm requires a reliable and global genome scale-model to achieve precise and accurate data interpretation.

So far, three yeast genome-scale metabolic models, *iFF708*, *iND750* and *iLL672*, have been published. All three models, however, lack a detailed description of the lipid metabolism. The first model, *iFF708* [16], consists of 1175 reactions linked to 708 ORFs. *iFF708* shows good predictions of many different cellular functions [17] and gene essentiality predictions [18]. However, almost all intermediate reactions in lipid metabolism were either lumped or neglected. The second model published was *iND750* [19]. *IND750* is fully compartmentalized, consisting of 1498 reactions linked to 750 ORFs. The model was validated by a large-scale gene deletion study and metabolic phenotypes [20] and was expanded to include regulation for predicting gene expression and phenotypes of different transcription factor mutants [21]. *iND750* contains more reactions and metabolites in lipid metabolism than *iFF708*, but still lacks a comprehensive description of lipid metabolism. The third published model is *iLL672*, which is derived from *iFF708* and comprises 1038 reactions. Several dead-end reactions of *iFF708* were eliminated leading to an improved accuracy of the single gene deletion prediction [22]. However, only minor improvements were made to reactions involved in lipid metabolism. The model was validated using 13C-labeling experiments to study the robustness of different yeast mutants [23].

Here our objective was to expand the genome-scale metabolic model of yeast to include a detailed description of lipid metabolism for use as a scaffold to integrate omics data. We used *iFF708* as a template for building a model based on recent literature that contains new reactions in lipid metabolism and transport relative to all previous models. The new model named *iIN800* includes 92 additional ORFs and provides a more detailed structure of lipid metabolism, tRNA synthesis and transport processes than previous models. The biomass composition, which is very important for flux balance analysis and predicting lethality, was also recalculated and improved. *iIN800* was validated with large-scale gene deletion data and growth simulation predictions. Simulated intracellular fluxes were also supported by 13C-labeling flux experimental data. Finally, we show that the transcriptome data of yeast cultivated under various growth conditions can be integrated with *iIN800* to identify lipid related Reporter Metabolites. We anticipate that *iIN800* will be useful as a scaffold for integrating multilevel omic data and that this new model will have a significant impact in the emerging field of lipidomics.

Results and discussion

*Model reconstruction and characteristics of iIN800*
Due to the complexity of compartmentalization used in *iND750* and the smaller scope of *iLL672*, the metabolic model *iFF708* was selected as a template for the development of the model *iIN800*. Pathway and reaction databases (e.g. KEGG), online resources (e.g. SGD), and literature were used to expand *iFF708*, with particular focus on lipid metabolism. *iIN800* contains 340 total reactions in lipid metabolism, more than at least 143 reactions greater than previous models (Table 1).

To compare metabolic characteristics of the different *in silico* models, lipid metabolism was classified into unique sub-categories (e.g. mitochondrial fatty acid synthesis, ergosterol biosynthesis) (Table 1). Fatty acid synthesis and elongation accounted for three of these sub-catego-
Sphingolipid synthesis reactions were added to iIN800 according to a recently reported model [25], resulting in more sphingolipid reactions than the template iFF708. Sphingolipid synthesis is the only sub-category in iIN800 with a significantly lower reaction tally than iND750. This is because iND750 incorporated both C24:0 and C26:0 as very long-chain fatty acids (the back bone of sphingolipids) to produce ceramides. Because the amount of very long chain fatty acids in *S. cerevisiae* is so low relative to other fatty acid species (<2% of total fatty acid pool) [24,26], iIN800 treats very long chain fatty acids as a single metabolite. As a result, fewer reactions are present in sphingolipid synthesis.

Related to other models, only minor changes in the biosynthesis of phospholipids and triacylglycerides as well as ergosterol were introduced in iIN800. However, esterification of sterols and degradation of lipids, which were not included in all other previous models, are present in iIN800 (Table 1). Finally, 26 ORFs encoding for tRNA synthetases were also included in iIN800. The additionally included ORFs and their related references as well as detailed comparisons of reactions in lipid metabolism of all reported models are given in Additional files 1 and 2, respectively.

In summary, iIN800 was reconstructed from 17.2% of the characterized ORFs in yeast and contains 1446 metabolic reactions and 1013 metabolites in total. This model is relatively more comprehensive as compared with previously described models (Table 2). The network characteristics of iIN800 and the starting model iFF708 are shown in Table 3. Within lipid metabolism, we have incorporated many new reactions in mitochondrial fatty acid synthesis, cytosolic fatty acid synthesis, fatty acid elongation, fatty acid activation and beta-oxidation. These ORFs were missing from previous models, which prevented simulation of mitochondrial fatty acid synthesis in both mitochondria and the cytosol. Fatty acid synthesis, which involves iterative malonyl-CoA condensations that result in a growing chain of fatty acids, is catalyzed by four major enzymes: β-ketoacyl-ACP synthase (a condensing enzyme), β-ketoacyl-ACP reductase, β-dehydroxyacyl-ACP dehydratase and enoyl-ACP reductase. In the cytosol, these enzymes are encoded by *FAS1* and *FAS2*. In the mitochondria, however, fatty acid synthesis is carried out by the products encoded by *CEM1*, *OAR1*, *HTD2* and *ETR1*. These ORFs were missing from previous models, which prevented simulation of mitochondrial fatty acid synthesis. Fatty acid elongation, which leads to the production of other fatty acid species (<2% of total fatty acid pool) [24,26], is so low relative to other fatty acid species (<2% of total fatty acid pool) [24,26], iIN800 treats very long chain fatty acids as a single metabolite. As a result, fewer reactions are present in sphingolipid synthesis.

### Table 1: Comparison of the number of lipid metabolism reactions among yeast genome-scale metabolic models

| Model                              | iFF708 | iLL672 | iND750 | iIN800 |
|------------------------------------|--------|--------|--------|--------|
| Mitochondrial fatty acid synthesis | 14     | 0      | 13     | 45     |
| Cytosolic fatty acid synthesis     | 17     | 18     | 27     | 48     |
| Fatty acid elongation              | 0      | 4      | 2      | 33     |
| Fatty acid activation and beta-oxidation | 9     | 19     | 53     | 65     |
| Sphingolipid synthesis            | 18     | 22     | 37     | 27     |
| Phospholipid and TAG synthesis     | 37     | 37     | 35     | 68     |
| Ergosterol biosynthesis            | 31     | 28     | 30     | 29     |
| Ergosterol esterification          | 0      | 0      | 0      | 2      |
| Lipid degradation                  | 0      | 0      | 0      | 23     |
| **Total**                          | 126    | 129    | 197    | 340    |

### Table 2: Structure comparison of *S. cerevisiae* genome-scale metabolic models

| Model | Genes | Reactions | Metabolites |
|-------|-------|-----------|-------------|
| iFF708| 708(15.2%)* | 1175 | 825 |
| iLL672| 672(14.1%)* | 1038 | 636 |
| iND750| 750(16.1%)* | 1489 | 972 |
| iIN800| 800(17.2%)* | 1446 | 1013 |

*percentage of associated ORFs in the model relative to characterized ORFs in the yeast genome
Table 3: Network characteristics of the reconstructed metabolic network of S. cerevisiae strain iFF708 and iN800

| Model                        | iFF708 | iN800 |
|------------------------------|--------|-------|
| Metabolites                  |        |       |
| Cytosolic metabolites        | 925    | 1013  |
| Mitochondrial metabolites    | 518    | 631   |
| Extracellular metabolites    | 170    | 228   |
| Excretion fluxes             | 1175   | 1446  |
| Mitochondrial reactions      | 104    | 161   |
| Cytosolic reactions          | 723    | 906   |
| Exchange fluxes              | 348    | 379   |
| Cytosolic exchange fluxes    | 286    | 304   |
| Mitochondrial exchange fluxes| 62     | 75    |
| Reactions with ORF assignments| 1075   | 1209  |
| Biochemical and Physical considerations | 140  | 237   |

Improved biomass equation

The biomass equation is crucial for using genome-scale metabolic models to simulate growth using flux balance analysis (FBA). Therefore, an important consideration in the development of iN800 was to address the concern that the biomass composition of S. cerevisiae changes under different growth conditions. For example, during growth on excess glucose the carbohydrate content increases and during growth on excess ammonium the protein content increases. To assess the sensitivity of flux simulations using iN800 towards changes in the macro-molecular composition, we performed constraint-based simulations by varying the protein, RNA, carbohydrate and lipid content of the biomass in physiological relevant ranges based on previous experimental reports [27-29], from 35–65%, 3.5–12%, 15–50% and 2–15%, respectively. Specifically, glucose and ammonium uptake rates were minimized for both glucose- and ammonium-limited growth conditions, respectively, using different macromolecular compositions at fixed growth rates, (note: this is the same mathematical problem as fixing uptake rates and maximizing growth rate). In this way, we could compare the differences between glucose- and ammonium-limited growth conditions. The results are illustrated in Figure 2. An interesting finding was that the protein content strongly affects the uptake rates at both glucose- and ammonium-limited conditions, albeit to a greater extent in ammonium-limited conditions (Fig. 2A). The carbohydrate content on the other hand does not have an impact on the ammonium uptake rate, it strongly impacts the glucose uptake rate (Fig. 2C). The RNA content and the lipid content have only a minor impact on growth (Figures 2B and 2D).

In summary, the sensitivity analysis shows that the biomass composition can significantly impact predictions made with genome-scale metabolic models to varying degrees based on different growth conditions. We therefore present new biomass equations to be used under C-limited and N-limited growth conditions, respectively. These compositions result from previous studies and our own measurements of lipids and fatty acids across multiple N-limited and C-limited growth conditions (data not shown). Using a separate biomass composition for N-limited cultures has not been proposed previously. The N-limited biomass equation is therefore new. Relative to previous C-limited biomass compositions, the most dramatic changes in our here proposed biomass equation is with respect to the lipids and fatty acids (Table 4). While previous C-limited biomass compositions, the most dramatic changes in our here proposed biomass equation is with respect to the lipids and fatty acids (Table 4). While the sensitivity analysis suggests that these components will most likely only lead to a small improvement in the accuracy of C-limited flux simulations, they may play an important role in lethality prediction by the model, as the addition of extra components in the biomass equation will give a higher resolution.

Growth simulation capability

In silico genome-scale models are most generally used to predict various phenotypes. These include growth rates and extracellular secretion rates of metabolite products, as well as uptake rates of nutrients. In addition, models can be employed to explore active route(s) in metabolic pathways under certain growth conditions as illustrated for a genome-scale metabolic model of E. coli [30-32] as well as for one of the S. cerevisiae genome-scale metabolic models [17].

To validate iN800, we first investigated the model's ability to simulate aerobic and anaerobic growth in glucose- or ammonium-limited conditions. Several published chemostat datasets were used as experimental references. As shown in Figure 3, the results from the computational growth prediction agreed with experimental measurements. Less than 10% relative error was observed (Figure 3). The details of the simulations and the corresponding reference data are given in Additional file 3. Intracellular
fluxes can be easily visualized using the ReMapper software and our model (Additional files 4 and 5).

Since the new biomass equations would be expected to impact the overall flux distributions, we used 13C-flux analysis data to further confirm the computed intracellular fluxes. Specifically, fluxes in the central carbon metabolism at two different growth conditions were both measured by 13C-labeling experiments and calculated by FBA using iIN800. The model validation is shown in Figure 4. There is a high degree of agreement between the predicted and experimental fluxes in the central metabolism, with the exception of fluxes through the pentose phosphate pathway (PPP). Using FBA, the flux through the PPP is largely determined by the requirement for NADPH, and it has earlier been shown difficult to balance NADPH production and consumption [33]. This may explain why the FBA simulations under-predict the flux through this pathway.

**Evaluation of large-scale gene deletion**

To verify further iIN800, we investigated the ability of the model to predict for growth viability due to a single gene deletion. *In silico* deletion phenotype predictions were examined for the new model with cells grown in both minimal media with a sole carbon source (glucose, galactose, glycerol and ethanol) and with rich media (YPD). iIN800 was assessed for its ability to make correct predictions based on experimental data [22,34]. A summary of the *in silico* single gene deletion predictions are given in Table 5. The overall prediction rate of iIN800, derived from 3392 total predictions, was 89.36%, with 95.50% sensitivity and 38.69% selectivity. The evaluation of the mean of a confusion matrix as the geometric mean of iIN800 equals 60.79%. The performance of the iIN800 model has improved by ~2% and ~7% in terms of overall prediction rate compared with the models iFF708 and iND750, respectively. We believe that the improvement is mainly due to upgrades in the biomass equation, which is

---

Figure 1
The reconstructed *S. cerevisiae* genome-scale metabolic model iIN800.
consistent with results from Kuepfer et al. demonstrating that more accurate biomass compositions lead to improved lethality predictions [22]. The false predictions might be due to missing information in gene regulation, biomass compositions, dead-end reactions and medium composition, especially in the rich medium [18,19].

**Integration of transcriptome data with genome-scale metabolic models**

Genome-scale metabolic models have shown promise for identifying Reporter Metabolites, defined as metabolites whose neighboring genes in a bipartite metabolic graph are most significantly affected and respond as a group to genetic or environmental perturbations [14]. Such an approach has previously been used to reveal important regulatory hot-spots in metabolism from genome-wide expression data and has demonstrated promise for integrating omic data using network topology. To highlight the importance and utility of having a more complete metabolic model in this integrated analysis, the genome-scale models *iN800* and *iFF708* were used to calculate Reporter Metabolites. Multiple sets of transcriptome data were used for analysis. Lists of the top thirty most significant Reporter Metabolites for several perturbations are compared between *iN800* and *iFF708* in Table 6, and Reporter Metabolites unique to *iN800* are marked in bold.

First, transcriptome data from the yeast metabolic cycle [35] were analyzed. Notably, the reporter algorithm identified unique Reporter Metabolites with *iN800* that would have been missed if *iFF708* was used as the scaffold (Table 6). The most dramatic difference was observed for the reductive charging phase of the metabolic cell cycle.

---

**Figure 2**

*Sensitivity analysis shows the influence of macromolecular composition on the simulated growth rate using iN800.* The simulations were performed for aerobic glucose- and ammonium-limited cultivations by varying (A) the protein content (35–65%), (B) the RNA content (3.5–12%), (C) the carbohydrate content (15–50%) and (D) the lipid content (2 – 15%).
Table 4: Biomass composition

| Metabolites    | Amount (mmol/gDW) | Carbon-limited | Nitrogen-limited |
|----------------|-------------------|----------------|------------------|
| **Amino acids** |                   |                |                  |
| L-Alanine      | 0.357             | 0.252          |
| L-Arginine     | 0.136             | 0.098          |
| L-Asparagine   | 0.172             | 0.153          |
| L-Aspartate    | 0.172             | 0.153          |
| L-Cysteine     | 0.043             | 0.044          |
| L-Glutamate    | 0.268             | 0.231          |
| L-Glutamine    | 0.268             | 0.231          |
| Glycine        | 0.325             | 0.278          |
| L-Histidine    | 0.075             | 0.071          |
| L-Isoleucine   | 0.172             | 0.142          |
| L-Leucine      | 0.250             | 0.207          |
| L-Lysine       | 0.239             | 0.204          |
| L-Methionine   | 0.050             | 0.044          |
| L-Phenylalanine| 0.114             | 0.092          |
| L-Proline      | 0.129             | 0.118          |
| L-Serine       | 0.254             | 0.225          |
| L-Threonine    | 0.197             | 0.160          |
| L-Tryptophan   | 0.027             | 0.028          |
| L-Tyrosine     | 0.096             | 0.068          |
| **Carbohydrates** |               |                |                  |
| Glycogen       | 0.519             | 0.667          |
| alpha,alpha-Trehalose | 0.023 | 0.085 |
| Mannan         | 0.821             | 0.994          |
| 1,3-beta-D-Glucan | 1.136  | 0.963 |
| **RNA**        |                   |                |                  |
| AMP            | 0.051             | 0.040          |
| GMP            | 0.051             | 0.040          |
| CMP            | 0.050             | 0.039          |
| UMP            | 0.067             | 0.052          |
| **DNA**        |                   |                |                  |
| dAMP           | 0.004             | 0.004          |
| dCMP           | 0.002             | 0.003          |
| dTMP           | 0.004             | 0.004          |
| dGMP           | 0.002             | 0.003          |
| **Lipids**     |                   |                |                  |
| Phosphatidylcholine | 0.002884 | 0.001660 |
| 1-Phosphatidy-D-myoinositol | 0.001531 | 0.001656 |
| Phosphatidylserine | 0.000373 | 0.000302 |
| Phosphatidylethanolamine | 0.000697 | 0.00083 |
| Acyl_acids     | 0.000206          | 0.000723       |
| Triacylglycerol| 0.000781          | 0.003618       |
| Ergosterol-ester | 0.000812 | 0.004632 |
| Ergosta-5,7,22,24(28)-tetraenol | 0.000125 | 0.000167 |
| Ergosterol     | 0.005603          | 0.005155       |
| Zymosterol     | 0.000015          | 0.000051       |
| Episterol      | 0.000096          | 0.000062       |
| Fecosterol     | 0.000114          | 0.000068       |
| Lanosterol     | 0.000032          | 0.000074       |
| 4,4-Dimethylzymosterol | 0.000056 | 0.000046 |
| Ceramide-I     | 0.000351          | 0.000075       |
| Ceramide-II    | 0.000066          | 0.000009       |
Further demonstrations of the applicability of iIN800 as a scaffold to integrate omic data were performed by analyzing transcriptome data derived from nutrient-limited [36], oxygen-limited [36] and temperature stress conditions [37]. Previously, mRNA and protein levels of genes and enzymes in fatty acid catabolism have been shown to be significantly different between carbon-limited and nitrogen limited growth [38]. When comparing these conditions, only iIN800 was able to identify fatty acids as Reporter Metabolites (Table 6). In anaerobic yeast cultivation, oleic acid has to be added to the medium because unsaturated fatty acids synthesis is not possible; therefore, the expression of genes in this pathway is induced by the function of the ORE element [39]. Consistent with this observed cellular response, only iIN800, with identified Reporter Metabolites involved in β-oxidation (Table 6). Similarly, iIN800 was able to highlight the importance of unsaturated fatty acids when comparing high and low temperature cultivations (Table 6), which is known to be important for maintaining proper membrane structure and fluidity [40].

Without the expanded model, the importance of cellular regulation stemming from lipid metabolism would be missed in analyses where metabolic topology is used for integrating data. As an illustration, we integrated results from our Reporter Metabolite analysis with known protein-protein and protein-DNA interaction networks to infer regulatory structure. First, genes associated to Reporter Metabolites in lipid metabolism unique to iIN800 and determined when comparing carbon- and nitrogen-limited growth (decanoyl-CoA, dodecanoyl-CoA, trans-2-C141-CoA, trans-2-C161-CoA, trans-2-C181-CoA) were identified. These genes were then used to search for highly regulated subnetworks within a protein-protein and protein-DNA interaction network. By applying a p-value threshold of 0.01 to filter for genes with significant gene expression, we inferred a regulatory network controlling the expression of lipid metabolism genes associated to the Reporter Metabolites (Figure 5). Strikingly, regulators at the top of this hierarchy are consistent with those previously known to be significantly changed between carbon- and nitrogen-limited growth. These include: SNF1, SNF4, MIG1 and ADR1 (glucose repression), OAF1 (β-oxidation), and INO1 and INO4 (phospholipid synthesis), among others. Previously reported genome-scale models are not capable of being used as scaffolds for implicating the conditional response of these lipid metabolism regulators because they lack a detailed description of lipid metabolism.

**Conclusion**

Genome-scale metabolic models have emerged as a valuable tool in the post-genomic era for illustrating whole-cell functions based on the complete network of biochemical reactions. An iterative reconstruction process is required to achieve a comprehensive *S. cerevisiae* genome-scale metabolic model. In this work, we focused on improving the formulation of lipid metabolism relative to previously published *S. cerevisiae* genome-scale metabolic models. Validating the model and new biomass equations, the constraint-based simulation of iIN800 showed accurate predictions of cellular growth and is also consistent with $^{13}$C-labeling experiments. Furthermore, *in silico* gene essentialness predictions were found to be in high agreement with *in vivo* results. Finally, we show that iIN800, being more complete, is a better network scaffold for integration of multilevel omics data.
In conclusion, by incorporating a more complete description of lipid metabolism, \textit{iIN800} is positioned to have a broader impact than previously described yeast models. Its capability of predictions were consistent with a number of experimental data both quantitatively (growth rate) and qualitatively (gene essentialness). Moreover, the new model is positioned to be used for studying the regulation and role of lipid metabolism during different growth conditions. With the high degree of homology in lipid metabolism between yeast and humans and emergence of lipidomics, this is expected to allow for new insights into the connection between lipid metabolism and overall cellular function for industrial and medical applications.

\textbf{Methods}

\textbf{Model reconstruction and visualization}

Reconstruction of the \textit{S. cerevisiae} genome-scale metabolic model was done by expanding \textit{iFF708} [16]. The additional ORFs included in the expansion procedure were involved in lipid metabolism, tRNA synthesis and lipoamide dehydrogenase. These ORFs were added based on publications listed in Additional file 1. Online resources related to \textit{S. cerevisiae}, such as SGD [41], MIPS [42] and YPD [43], were also used to confirm the existence of the ORFs and their function. Pathway and reaction databases including KEGG [44], ExPaSy [45], and Reactome [46], were used together with research papers to identify relevant information of the additional reactions and metabolites, e.g. stoichiometry and co-factor usage. The expanded \textit{iFF708}, called \textit{iIN800}, was visualized by Adobe Illustrator software (Adobe Systems), and then converted to EPS format (Adobe Systems) format which is downloadable as Additional file 6. In this visualization file, it is possible to overlay information about transcription, fluxes etc. A detailed list of metabolic reactions in \textit{iIN800} is provided as Additional file 7.

\textbf{Metabolic modeling and simulations}

The reaction set in \textit{iIN800} was used for construction of a stoichiometric matrix \(S\ (m \times n)\). In the stoichiometric matrix, \(m = 1013\), which is the number of metabolites, and \(n = 1446\), which is the number of metabolic reactions. With an assumption of steady state for all metabolite pools, a linear equation constraining the fluxes in the metabolic network is obtained [30,47]:

\[
S \cdot v = 0 \tag{1}
\]

Here \(v\) is a vector that contains all the fluxes in the model. Equation (1) has a large number of degrees of freedom, i.e. it is an underdetermined problem, and linear programming was employed to solve the equation system by maximizing an objective function \(Z\) (equal to the growth rate), an approach generally referred to as flux balance analysis (FBA) [30,47]. The problem formulation is described below.

Maximize:

\[
Z = \omega \cdot v
\]

Subject to:

\[
S \cdot v = 0
\]

\[
\alpha \leq v \leq \beta
\]

where \(\alpha\) and \(\beta\) are lower and upper bounds of fluxes respectively, \(\omega\) is a weight vector indicating an amount of desired metabolites for biomass synthesis. For irreversible fluxes semi-positive infinite boundary was applied as \(0 \leq v < \infty\), and fully infinite boundaries was applied as \(-\infty \leq v \leq \infty\) for reversible fluxes. The problem was solved by using the commercial linear programming software package LINDO (Lindo systems Inc., Chicago, IL, USA). The calculated intracellular fluxes were overlaid on the visualized genome-scale map as described previously by the ReMapper software (The software has been developed for visualization of multilevel omics data onto a metabolic map.).

\textbf{Calculation of biomass composition and sensitivity analysis}

The biomass composition was re-calculated in order to improve the prediction of the model during growth at different nutrition-limitations, i.e. carbon- and nitrogen-lim-
ited growth condition. The contents of macro-molecules were extracted from the thesis of Schulze [27] who measured the biomass composition at a dilution rate of 0.1 h\(^{-1}\). The calculations were performed as described previously [16]. The calculation of protein precursors, i.e. amino acids, and carbohydrate precursors, i.e. trehalose, glycogen, manna and glucan, were adopted from Schulze's work [27]. Deoxyribonucleotide and ribonucleotide compositions were calculated from the study of Vaughan-Martini and co-workers [48]. Lipid compositions were calculated from our own measurements of structural lipidomics, which contains phospholipids, triacylglycerol, sterols, sterol-esters, sphingolipids, free fatty acids and fatty acids composition of all measured lipid classes (unpublished data). The impact of the macromolecular composition on biomass yield was explored in aerobically glucose- and ammonium-limited conditions by fixing the specific growth rate and then minimizing the glucose and ammonium uptake rates at both glucose- and ammonium-limited growth conditions. Four parameters were evaluated, namely the protein, RNA, carbohydrate and lipid content of the biomass.

**Growth simulations**

The metabolic capabilities of iIN800 were evaluated by using FBA and linear programming to simulate the biomass flux representing the \textit{in silico} growth rate, which were derived by maximizing the biomass production. Data from various carbon-limited and nitrogen-limited chemostat experiments performed at either aerobic or anaerobic growth condition were taken from the literature for comparisons (see references in Additional file 3). These data were used to validate the metabolic capabilities of the model by comparing \textit{in silico} biomass yields with \textit{in vivo} biomass yields. The \textit{in silico} biomass yields were calculated by fixing measurable uptake rates of extracellular metabolites, such as glucose, ammonium and oxygen, as well as secretions rates of acetate, glycerol, ethanol, succinate, pyruvate and carbon dioxide. The biomass equation (or flux), which was the objective function, was changed depending on the growth conditions evaluated according to the data provide in Table 4.

**Large-scale gene essentiality simulations**

The impact of individual gene deletions on cell growth of \textit{iIN800} was evaluated by eliminating the reaction(s) corresponding to each gene in the model from the stoichiometric matrix \(S\) and then simulating growth of the mutant by FBA. The \textit{in silico} gene essentialities were simulated for growth on rich- and minimal-medium. For minimal media, different carbon sources (glucose, galactose, glycerol and ethanol), ammonium, sulphate and phosphate were evaluated. For rich media, the uptake fluxes of amino acids, purines and pyrimidines were added as additional constraints as previously described [18]. The \textit{in silico} simulations were compared to experimental data available in the MIPS and SGD databases and from competitive growth assays [34] as well as yeast mutant array experiments [22]. The power of \textit{iIN800} to predict gene essentiality was evaluated based on the criteria defined as follows:

\[
\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)}
\]

\[
\text{Sensitivity} = \frac{TP}{(TP + FN)}
\]

\[
\text{Specificity} = \frac{TN}{(TN + FP)}
\]

\[
\text{Positive predictive value} = \frac{TN}{(TP + FP)}
\]

\[
\text{Negative predictive value} = \frac{TN}{(TN + FN)}
\]

\[
\text{Geometric mean} = \left(\frac{\text{Sensitivity} \cdot \text{Specificity}}{1}\right)^{1/2}
\]

* = Aerobic growth, ** = Anaerobic growth

TP = true positive, TN = true negative, FP = false positive, FN = false negative.

| Evaluation | Glucose* | Glucose** | Galactose | Glycerol | Ethanol | YPD | Total |
|------------|---------|----------|-----------|----------|---------|-----|-------|
| TP         | 468     | 469      | 461       | 461      | 463     | 567 | 2889  |
| TN         | 23      | 23       | 20        | 17       | 21      | 38  | 142   |
| FP         | 37      | 37       | 42        | 45       | 43      | 21  | 225   |
| FN         | 14      | 13       | 19        | 19       | 15      | 56  | 136   |
| Number of deletions | 542 | 542    | 542       | 542      | 542     | 542 | 682   |
| Positive prediction rate | 92.67 | 92.69   | 91.65     | 91.11    | 91.50   | 96.43 | 92.77 |
| Negative prediction rate | 62.16 | 63.89   | 51.28     | 47.22    | 58.33   | 40.43 | 51.08 |
| Accuracy   | 90.59   | 90.77    | 88.75     | 88.19    | 89.30   | 88.71 | 89.36 |
| Sensitivity| 97.10   | 97.30    | 96.04     | 96.04    | 96.86   | 91.01 | 95.50 |
| Selectivity| 38.33   | 38.33    | 32.26     | 27.42    | 32.81   | 64.41 | 38.69 |
| Geometric mean | 61.01 | 61.07    | 55.66     | 51.32    | 56.38   | 76.56 | 60.79 |

\(* = \text{Aerobic growth, ** = Anaerobic growth}

TP = \text{true positive, TN = true negative, FP = false positive, FN = false negative.}
| Metabolites                          | iIN800 | iFF708 | iIN800 | iFF708 | iIN800 | iFF708 |
|------------------------------------|--------|--------|--------|--------|--------|--------|
| **Oxidative phase**                |        |        |        |        |        |        |
| AMP                                | IMP    | IMP    | AMPM   | AMPM   | AMPP   | AMPP   |
| IMP                                | Xanthosine 5'-phosphate | Xanthosine 5'-phosphate | PyrophosphateM | PyrophosphateM | tRNA(M) | tRNA(M) |
| Pyrophosphate                       | L-Methionine | L-Methionine | ATPM   | PyrophosphateM | tRNA(M) | tRNA(M) |
| L-Methionine                        | 5-Phospho-alpha-D-ribose 1-PP | 5-Phospho-alpha-D-ribose 1-PP | tRNA(M) | PyrophosphateM | tRNA(M) | tRNA(M) |
| Xanthosine 5'-phosphate             | L-Aspartate | L-Aspartate | H+M    | L-TryptophanM | H+M    | H+M    |
| ATP                                | Sulfate | Sulfate | NADHM  | L-Tryptophan-tRNA(Trp)M | NADHM  | L-Tryptophan-tRNA(Trp)M |
| 5-Phospho-alpha-D-ribose 1-PP      | Homocysteine | Homocysteine | Porphobilinogen | Porphobilinogen | Dolichyl beta-D-mannosyl-P | Mannan |
| L-Serine                           | AMP    | AMP    | AMP    | H+M    | AMP    | H+M    |
| L-Aspartate                        | H+EXT  | H+EXT  | L-Tryptophanyl-tRNA(Trp)M | Xanthine | L-Asparaginy-tRNA(Asn)M | H+M    |
| Homocysteine                       | N6-(L-1,3-Dicarboxypropyl)-L-lysine | N6-(L-1,3-Dicarboxypropyl)-L-lysine | tRNA(Ile)M* | Dolichyl phosphate | tRNA(Ile)M* | Dolichyl phosphate |
| Sulfate                            | 5,10-Methylene tetrahydrofolic acid | 5,10-Methylene tetrahydrofolic acid | L-Isoleucyl-tRNA(Ile)M* | L-Isoleucyl-tRNA(Ile)M* | L-Isoleucyl-tRNA(Ile)M* | L-Isoleucyl-tRNA(Ile)M* |
| L-Glutamine                        | Aminoimidazole ribotide | Aminoimidazole ribotide | tRNA(Thr)M* | all-trans-Nonaprenyl-PP | tRNA(Thr)M* | all-trans-Nonaprenyl-PP |
| L-Cysteine                         | L-Cystathionine | L-Cystathionine | L-Threonyl-tRNA(Thr)M* | NADHM  | L-Threonyl-tRNA(Thr)M* | NADHM  |
| L-Asparagine                       | L-Serine | L-Serine | Xanthine | ATPM   | Xanthine | ATPM   |
| S-Adenosyl-L-methionine            | Uracil  | Uracil  | Dolichyl phosphate | D-Mannose 6-phosphate | Dolichyl phosphate | D-Mannose 6-phosphate |
| Sulfite                            | Sulfite | Sulfite | all-trans-Nonaprenyl-PP | all-trans-Nonaprenyl-PP | all-trans-Nonaprenyl-PP | all-trans-Nonaprenyl-PP |
| 5,10-Methylene tetrahydrofolic acid | 5-amino-4-imidazolecarboxylate | 5-amino-4-imidazolecarboxylate | L-Asparaginyl-tRNA(Asn)M | L-Asparaginyl-tRNA(Asn)M | L-Asparaginyl-tRNA(Asn)M | L-Asparaginyl-tRNA(Asn)M |
| 3-Phospho-oxyxypyrurate            | 2-Hydroxybutane-1,2,4-tricarboxylate | 2-Hydroxybutane-1,2,4-tricarboxylate | tRNA(Phen)M* | CO2M    | tRNA(Phen)M* | CO2M    |
| N6-(L-1,3-Dicarboxypropyl)-L-lysine| S-Adenosyl-L-methionine | S-Adenosyl-L-methionine | L-Phenylalaninyl-tRNA(Phen)M* | Guanosine | L-Phenylalaninyl-tRNA(Phen)M* | Guanosine |
| L-Cystathionine                    | L-Asparagine | L-Asparagine | Intermediate_Methylzymosterol_II | IsocitrateM | Intermediate_Methylzymosterol_II | IsocitrateM |
| NH₃                                | tRNA(Phen)M* | tRNA(Phen)M* | 5-Phosphoribosylamine | Intermediated_Zymosterol_II | 5-Phosphoribosylamine | Intermediated_Zymosterol_II |
| L-Phenylalaninyl-tRNA(Phen)M*      | GlycineM | GlycineM | GlycineM | GTPM    | GlycineM | GTPM    |
| Tetrahydrofolate                   | Guanine | Guanine | UbiquinoM | GDPM    | UbiquinoM | GDPM    |
| Guanine                            | N1-(5'-phosphoribosyl)acetamide | N1-(5'-phosphoribosyl)acetamide | D-Mannose 6-phosphate | D-Mannose 6-phosphate | D-Mannose 6-phosphate | D-Mannose 6-phosphate |
| Sulfite                            | Tetrahydrofolate | Tetrahydrofolate | Gamma GLucose | Tetrahydrofolate | Gamma GLucose | Tetrahydrofolate |
| L-Histidine                        | L-Histidine | L-Histidine | UbiquinoM-9M | IDPM    | UbiquinoM-9M | IDPM    |
| Sulfite                            | L-Histidine | L-Histidine | L-Histidine | IDPM    | L-Histidine | IDPM    |
| 5-amino-4-imidazolecarboxylate     | Homoisoctamate | Homoisoctamate | L-Aspartyl-tRNA(Asp)M* | Dodecanoyl-CoA* | L-Aspartyl-tRNA(Asp)M* | Dodecanoyl-CoA* |
| Glycine M                          | GMP    | GMP    | Pyrophosphate | Phosphatidate | C100ACPm | IDP    |
| **Reductive building phase**       |        |        |        |        |        |        |
| **Reductive charging Phase**       |        |        |        |        |        |        |

* Metabolite is contained in iIN800 only
1 = data from Tu, B. P., A. Kudlicki, et al. (2005)
Table 6: Top thirty Reporter Metabolites calculated from various perturbations. The Reporter Metabolite algorithm was performed with *iIN800* and *iFF708*. (Continued)

| Carbon- and Nitrogen-limited<sup>2</sup> | Aerobic and Anaerobic<sup>2</sup> | Temperature(30°C and 15°C)<sup>3</sup> |
|----------------------------------------|---------------------------------|----------------------------------|
| **iIN800** | **iFF708** | **iIN800** | **iFF708** | **iIN800** | **iFF708** |
| Glyoxylate | Glyoxylate | Oxygen | Ferricytochrome cM | IMP | IMP |
| GLUxt | L-Phenylalanine | Ferricytochrome cM | Tetrahydrofolate | alpha, alpha-Trehalose | D-Erythro 4-phosphate |
| Isocitrate | GLUxt | Ferricytochrome cM | Ferrocyanocobalamin | alpha, alpha-Trehalose | D-Erythro 4-phosphate |
| ALAx3 | Isocitrate | Ubiquinone-9M | Oxygen | IMP | IMP |
| Malate | ALAx3 | UbiquinolM | UbiquinolM | IMP | IMP |
| Allantoate | Allantoate | ADPM | ADPM | IMP | IMP |
| SERxt | Malate | H+M | H+M | IMP | IMP |
| L-Alanine | L-Alanine | Dodecanoyl-CoA<sup>*</sup> | N6-(L-1,3-Dicarboxypropyl)-L-lysine | N6-(L-1,3-Dicarboxypropyl)-L-lysine | N6-(L-1,3-Dicarboxypropyl)-L-lysine |
| Decanoyl-CoA<sup>*</sup> | SERxt | Dodecanoyl-CoA<sup>*</sup> | FumarateM | N6-(L-1,3-Dicarboxypropyl)-L-lysine | N6-(L-1,3-Dicarboxypropyl)-L-lysine |
| ASNxt | ASNxt | FumarateM | NADH | URlx | Homocysteine |
| GLNxt | GLNxt | Hexadecanoyl-9-ene-CoA<sup>*</sup> | N6-(L-1,3-Dicarboxypropyl)-L-lysine | N6-(L-1,3-Dicarboxypropyl)-L-lysine | N6-(L-1,3-Dicarboxypropyl)-L-lysine |
| ILExt | ILExt | Hexadecanoyl-9-ene-CoA<sup>*</sup> | Sphinganine 1-phosphate | Homocysteine | N-Acetyl-L-glutamateM |
| VALxt | VALxt | Octadecanoyl-9-ene-CoA<sup>*</sup> | Fumarate | Octadecanoyl-CoA | N-Acetyl-L-glutamateM |
| Trans-2-C161-CoA* | Ferricytochrome cM | Octadecanoyl-9-ene-CoA<sup>*</sup> | Dihydrofolate | N-Acetyl-L-glutamateM | Dihydrofolate |
| Trans-2-C181-CoA* | Ferrocyanocobalamin | Sphinganine 1-phosphate | Octadecanoyl-CoA | N2-Acetyl-L-ornithinM | N2-Acetyl-L-ornithinM |
| Trans-2-C141-CoA* | PHExt | Isocitrate | Dihydrofolate | Anthranilate | Anthranilate |
| PHExt | L-Asparagine | Tetradeanoyl-Co | ZYMStxt | Anthranilate | Anthranilate |
| Ferricytochrome cM | Allantoin | Fumarate | NAD<sup>+</sup> | S-Adenosyl-L-homocysteine | S-Adenosyl-L-homocysteine |
| Ferrocyanocobalamin | LEUxt | Trans-2-C161-CoA* | FADM | Octadecanoyl-9-ene acid* | Octadecanoyl-9-ene acid* |
| FRUxt | FRUxt | Trans-2-C181-CoA* | 6-Phospho-D-gluconate | L-Aspartate | L-Aspartate |
| Allantoic | Succinate | Trans-2-C141-CoA* | 1,3-Diaminopropane | S-Adenosyl-L-homocysteine | S-Adenosyl-L-homocysteine |
| Succinate | TYRxt | Glyoxylate | sn-Glycerol 3-phosphate | NADH | NADH |
| HISxt | METxt | sn-Glycerol 3-phosphate | O-Acetylaminocitrate | UREAxt | UREAxt |
| Dodecanoyl-CoA<sup>*</sup> | GLYxt | ERGOSTxt | Ethanol | 1-Phosphatidyl-D-myoinositol-3-P | 1-Phosphatidyl-D-myoinositol-3-P |
| PROxt | ASPxt | DPIPExt | N(p)-Methyl-L-histidine | L-Asparagine | L-Asparagine |
| alpha-D-Mannose | GLCxt | Dipetide | N(p)-Methyl-L-histidine | Adenosine 3',5'-bisphosphate | Adenosine 3',5'-bisphosphate |
| Trans-3-C16-CoA* | L-Tyrosine | Oligopeptide | 4-imidazolecarboxylate | Adenosine 3',5'-bisphosphate | Adenosine 3',5'-bisphosphate |
| Trans-3-C18-CoA* | PROxt | PEPTxt | 3-Methyl-2-oxobutanoateM | L-Asparagine | L-Asparagine |
| Trans-3-C14-CoA* | alpha-D-Mannose | Trans-3-C16-CoA* | Sphinganine | S-Adenosyl-L-homocysteine | S-Adenosyl-L-homocysteine |

<sup>*</sup> Metabolite is contained in *iIN800* only

2 = Tai, S. L., V. M. Boer, et al. (2005)
3 = Pizarro, F., M.C. Jewett, et al. (2008)
where TP = true positive, TN = true negative, FP = false positive, FN = false negative. Positive and negative values referred to viable and lethal phenotype, respectively.

**Reporter Metabolite determination**

Published microarray data were retrieved from Gene Expression Omnibus (GEO) [49]. The CEL files were normalized by the dChip software [50] in order to minimize overall intensity variation among a set of chips. The statistical test of significance was done by ANOVA or student t-test for p-value calculation.

Briefly, we describe the Reporter Metabolite calculations. The genome-scale model was converted to a bipartite undirected graph. In this graph each metabolite node has as neighbors the enzymes catalyzing the formation and consumption of the metabolite. The transcriptome data were mapped on the enzyme nodes using the significant values of gene expression. The normal commutative distribution was used to convert the p-values to a Z-score for further calculations. To identify an importance of metabolites in the metabolic network of the particular experimental conditions, the reporter algorithm was applied as described earlier [14].

**Inferring regulatory modules from Reporter Metabolites**

The interactome network was initially constructed with data obtained from YPD [43], ChIP-chip databases [51] (protein-DNA interaction) and BioGRID [52] (protein-protein interaction). The candidate genes of high scoring Reporter Metabolites were retrieved from the bipartite metabolite-gene encoding enzyme interaction graph. They were then used to identify subnetworks from the interactome network [53]. Significantly changing p-values from microarray data were mapped on the subnetwork and then also genes having a p-value < 0.01 directly connected with the Reporter Metabolites. The module was visualized by Cytoscape software [54].

**Authors’ contributions**

IN designed the study, performed the metabolic reconstruction and validation, and contributed to manuscript writing. MCJ carried out the C\(^{13}\)-labeling flux experiments, helped curate the model and contributed to manuscript writing. AM, CT, KL and SC contributed to the manuscript preparations, JN and SB participated in the concept and design of the study. All authors read and approved the final manuscript.

**Additional material**

| Additional file 1 | Additional ORFs. List of additional ORFs and their references containing in iIN800. Click here for file [http://www.biomedcentral.com/content/supplementary/1752-0509-2-71-S1.pdf] |
| Additional file 2 | Lipid metabolism reactions and comparison. Comparison of lipid metabolism reactions of all S. cerevisiae genome-scale models. Click here for file [http://www.biomedcentral.com/content/supplementary/1752-0509-2-71-S2.zip] |
| Additional file 3 | Growth simulation results. Growth simulations and comparison with experimental measurements. Click here for file [http://www.biomedcentral.com/content/supplementary/1752-0509-2-71-S3.pdf] |
| Additional file 4 | Aerobic Flux distribution. Visualization of flux distribution of aerobic growth mapping on Figure 1. Click here for file [http://www.biomedcentral.com/content/supplementary/1752-0509-2-71-S4.pdf] |
Additional file 5

Anaerobic Flux distribution. Visualization of flux distribution of anaerobic growth mapping on Figure 1. Click here for file [http://www.biomedcentral.com/content-supplementary/1752-0509-2-71-S5.pdf]

Additional file 6

High resolution file of Figure 1. High resolution of S. cerevisiae metabolic map is provided as EPS format. Click here for file [http://www.biomedcentral.com/content-supplementary/1752-0509-2-71-S6.eps]

Additional file 7

iIn800 model. List of all participated reactions in iIn800 model. Click here for file [http://www.biomedcentral.com/content-supplementary/1752-0509-2-71-S7.pdf]

Acknowledgements

The authors gratefully thank Mikael Rørdam Andersen and Kiran Raosaheb Patil for providing the ReMapper and the Reporter software, respectively. This work is supported by a grant from the National Center for Genetic Engineering and Biotechnology (BIOTEC) (grant number BT-B-06-NG-B5-4602). Intawat Nookaew gratefully acknowledges financial support by Thai Graduate Student Institute Science and Technology (TGIST). Michael C. Jewett is grateful to the NSF International Research Fellowship Program for supporting his work.

References

1. Nielsen J, Jewett MC: Impact of systems biology on metabolic engineering of Saccharomyces cerevisiae. FEBS Lett 2007.
2. Bottstein D, Chervitz SA, Cherry JM: Yeast as a model organism. Science 1997, 277(5330):1259-1260.
3. Pena-Castillo L, Hughes TR: Why are there still over 1000 uncharacterized yeast genes? Genetics 2007, 176(1):7-14.
4. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG: Life with 6000 genes. Science 1996, 274(5287):546-563.
5. Bassett DE Jr., Boguski MS, Hieter P: Yeast genes and human disease. Nature 1996, 379(6656):589-590.
6. Fourny F: Human genetic diseases: a cross-talk between man and yeast. Gene 1997, 195(1):1-10.
7. Steinmetz LM, Scharfe C, Deutschbauer AM, Mokranjac D, Herman ZS, Jones T, Chu AM, Giaever G, Prokisch H, Qfeiner P, Davis RW: Systematic screen for human disease genes in yeast. Nat Genet 2002, 31(4):400-404.
8. German JB, Gillies LA, Smillowitz JT, Zivkovic AM, Watkins SM: Lipidomics and lipid profiling in metabolomics. Curr Opin Lipidol 2007, 18(1):46-71.
9. Vigh L, Eszriba V, Sonnleitner A, Sonnleitner M, Piotto S, Maresca B, Liebundguth N, Lockhart DJ, Lucau-Danila A, Lussier M, M’Rabet N, Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Jewett MC: Impact of systems biology on metabolic engineering of Saccharomyces cerevisiae. FEBS Lett 2007.
10. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG: Life with 6000 genes. Science 1996, 274(5287):546-563.
11. Schirmer CR, Feany MB: Yeast genetics targets lipids in Parkinson’s disease. Trends Genet 2004, 20(7):273-277.
12. Mutch DM, Fauconnot L, Grigorov M, Fay LB: Putting the ‘Ome’ in lipid metabolism. Biotechnol Annu Rev 2006, 12:67-84.
13. Paul KR, Akesson M, Nielsen J: Use of genome-scale microbial models for metabolic engineering. Curr Opin Biotechnol 2004, 15(1):64-69.
14. Paul KR, Nielsen J: Uncovering transcriptional regulation of metabolism by using metabolic network topology. Proc Natl Acad Sci U S A 2005, 102(8):2685-2689.
15. Cakir T, Patil KR, Onsan Z, Ullgen KO, Kirdar B, Nielsen J: Integration of metabolome data with metabolic networks reveals reporter reactions. Mol Syst Biol 2006, 2:50.
16. Forster J, Famiili I, Fu P, Palsson BO, Nielsen J: Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network. Genome Res 2003, 13(2):244-253.
17. Famiili I, Forster J, Nielsen J, Palsson BO: Saccharomyces cerevisiae phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network. Proc Natl Acad Sci U S A 2003, 100(23):13134-13139.
18. Forster J, Famiili I, Palsson BO, Nielsen J: Large-scale evaluation of in vivo gene deletions in Saccharomyces cerevisiae. Omics 2003, 7(2):193-202.
19. Duarte NC, Herrgard MJ, Palsson BO: Reconstruction and validation of Saccharomyces cerevisiae IND750, a fully compartmentalized genome-scale metabolic model. Genome Res 2004, 14(12):1298-1309.
20. Duarte NC, Palsson BO, Fu P: Integrated analysis of metabolic phenotypes in Saccharomyces cerevisiae. BMC Genomics 2004, 5(1):93.
21. Herrgard MJ, Lee BS, Portnoy V, Palsson BO: Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in Saccharomyces cerevisiae. Genome Res 2006, 16(5):627-635.
22. Kuepfer L, Sauer U, Blank L: Metabolic functions of duplicate genes in Saccharomyces cerevisiae. Genome Res 2005, 15(10):1421-1430.
23. Blank LM, Kuepfer L, Sauer U: Large-scale 13C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. Genome Biol 2005, 6(6):r49.
24. Han G, Gable K, Kohlwein SD, Beaudoin F, Gipser JA, Dunn TM: The Saccharomyces cerevisiae YBR139w gene encodes the 3-ketoar dehydrogenase of the microsomal fatty acid elongase. J Biol Chem 2002, 277(38):35440-35449.
25. Alvareza-Vasquez F, Sims KJ, Cowart LA, Okamoto Y, Voit EO, Hannun YA: Quantitative analysis of modelled sphingolipid metabolism in Saccharomyces cerevisiae. Nature 2005, 433(7024):425-430.
26. Welch JW, Burlingame AL: Very long-chain fatty acids in yeast. J Bacteriol 1973, 115(1):464-466.
27. Schulze LJ: Anaerobic physiology of Saccharomyces cerevisiae. Lyngby. Technical University of Denmark; 1995.
28. Dyer JM, Chapatil DC, Kuan JW, Mullen RT, Pepperman AB: Metabolic engineering of Saccharomyces cerevisiae for production of novel lipid compounds. Appl Microbiol Biotechnol 2002, 59(2):224-230.
29. Jollow D, Kellerman GM, Linnane AW: The biogenesis of mitochondria. 3. The lipid composition of aerobically and anaerobically grown Saccharomyces cerevisiae as related to the membrane systems of the cells. J Cell Biol 1968, 37(2):221-230.
30. Edwards JS, Ibarra RU, Palsson BO: In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. Nat Biotechnol 2001, 19(2):125-130.
31. Feist AM, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, Broadbent LJ, Hatzimanikatis V, Palsson BO: A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. Mol Syst Biol 2007, 3:121.
32. Reed JL, Vo TD, Schilling CH, Palsson BO: An expanded genome-scale model of metabolism of Escherichia coli K-12 (iJR904 GSP/GPR). Genome Biol 2003, 4(9):R54.
33. 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. J Bacteriol 2001, 183(4):1441-1451.
34. Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connolly C, Davis K, Dietrich F, Dow SW, El Bakkoury M, Foury F, Friend SH, Gentelen E, Giaever G, Hegernhan J, Jones T, Laub M, Liao H, Liebenduth N, Lockhart DJ, Lucau-Danila A, Lussier M, M’Rabet N, Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connolly C, Davis K, Dietrich F, Dow SW, El Bakkoury M, Foury F, Friend SH, Gentelen E, Giaever G, Hegernhan J, Jones T, Laub M, Liao H, Liebenduth N, Lockhart DJ, Lucau-Danila A, Lussier M, M’Rabet N, Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H, Chu AM, Connolly C, Davis K, Dietrich F, Dow SW, El Bakkoury M, Foury F, Friend SH, Gentelen E, Giaever G, Hegernhan J, Jones T, Laub M, Liao H, Liebenduth N, Lockhart DJ, Lucau-Danila A, Lussier M, M’Rabet N,
Menard P, Mittmann M, Pai C, Reibischung C, Revuelta JL, Riles L, Rober-
ts C, Ross-Macdonald P, Sherens B, Snyder M, Sookhai-Mahadeo S, Stormer K, Vorontsova S, Voet M, Volckaert G, Ward TR, Wolocki R, Yen GS, Yu K, Zimmermann K, Philipsen P, Johnston M, Davis RW: Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science 1999, 285(5429):901-906.

Tu BP, Kudlicki A, Rowicka M, McKnight SL: Logic of the yeast metabolic cycle: temporal compartmentalization of cellular processes. Science 2005, 310(5751):152-158.

Tai SL, Boer VM, Daran-Lapujade P, Walsh MC, de Winde JH, Daran JM, Pronk JT: Two-dimensional transcriptome analysis in chemostat cultures. Combinatorial effects of oxygen availability and macronutrient limitation in Saccharomyces cerevisiae. J Biol Chem 2005, 280(1):437-447.

Pizarro F, Jewett MC, Nielsen J, Agosin E: Physiological and transcriptional mapping of evolutionary differences between laboratory and commercial Saccharomyces cerevisiae strains (submitted). 2008.

Kollman A, Daran-Lapujade P, Fullaondo A, Olsthoorn MM, Pronk JT, Sliper M, Heck AJ: Proteome analysis of yeast response to various nutrient limitations. Mol Syst Biol 2006, 2:2006. 0026.

Guttman A, Murasala AM, Firzinger A, Hamilton B, Klpelainen SH, Hartig A, Ruis H, Hiltunen JK, Rottensteiner H: Peroxisomal Delta3-
cis-Delta2-trans-enoyl-CoA isomerase encoded by ECI1 is required for growth of the yeast Saccharomyces cerevisiae on unsaturated fatty acids. J Biol Chem 1998, 273(47):31366-31374.

Rodriguez-Yargas S, Sanchez-Garcia A, Martinez-Rivas JM, Priezo JA, Randez-Gil F: Fluidization of membrane lipids enhances the tolerance of Saccharomyces cerevisiae to freezing and salt stress. Appl Environ Microbiol 2007, 73(1):10-116.

Hirschman JE, Balakrishnan R, Christie KR, Costanzo MC, Dwight SS, Engel SR, Fisk DG, Hong EL, Livstone MS, Nash R, Park J, Oughtred D, Skrzypek MS, Brooks J, Garrels JI: Yeast Genome Database. Nucleic Acids Res 2005, 33(Database issue):D442-5.

Guldener U, Munsterkotter M, Kastenmüller G, Strack N, van Helden JM, Pronk JT: BioGRID: a general repository for interaction datasets. Nucleic Acids Res 2006, 34(Database issue):D636-8.

Csank C, Costanzo MC, Hirschman J, Hodges P, Kranz JE, Mangan M, O’Neill K, Robertson LS, Skrzypek MS, Brooks J, Garrels JI: Three yeast proteome databases:YPD, PombePD, and CalPD (MycopathPD). Methods Enzymol 2002, 350:347-373.

Arakawa K, Kono N, Yamada Y, Mori H, Tomita M: KEGG-based pathway visualization tool for complex omics data. In Silico Biol 2005, 5(4):419-423.

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A: ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 2003, 31(13):3784-3788.

Vastrik I, D’Eustachio P, Schmidt E, Joshii-Tope G, Gopinath G, Croft D, de Bono B, Gillespie M, Jassal B, Lewis S, Matthews L, Wu G, Birney E, Stein L: Reactome: a knowledge base of biologic pathways and processes. Genome Biol 2007, 8(3):R39.

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(10):5522-5533.

Vaughan-Martini A, Martini A, Cardinali G: Electrophoretic karyotyping as a taxonomic tool in the genus Saccharomyces. Antonie Van Leeuwenhoek 1999, 73(2):145-156.

Barrett T, Edgar R: Gene expression omnibus: microarray data storage, submission, retrieval, and analysis. Methods Enzymol 2006, 411:352-369.

Schadt EE, Li C, Ellis B, Wong WH: Feature extraction and normalization algorithms for high-density oligonucleotide gene expression array data. J Cell Biochem Suppl 2001, Suppl 37:120-125.

Harbison CT, Gordon DB, Lee TI, Rinaldi NJ, Macisaac KD, Danford TW, Hannett NM, Tagne JB, Reynolds DB, Yoo J, Jennings EG, Zeitlinger J, Pokholok DK, Kellis M, Rolfe PA, Takusagawa KT, Lander ES, Gifford DK, Fraenkel E, Young RA: Transcriptional regulatory code of a eukaryotic genome. Nature 2004, 431(7004):99-104.

Jarvik C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M: BioGRID: a general repository for interaction datasets. Nucleic Acids Res 2006, 34(Database issue):D535-9.

Ideler TE, Thorsson V, Karp RM: Discovery of regulatory interactions through perturbation: inference and experimental design. Pac Symp Biocomput 2000:305-316.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T: Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003, 13(11):2498-2504.