Modeling Lipid Metabolism in Yeast

Eduard J. Kerkhoven

Contents

1 Introduction ................................................................................... 2
2 Genome-Scale Metabolic Models ...................................................... 2
   2.1 Generation of Genome-Scale Model ............................................. 2
   2.2 Computational Analysis ........................................................... 3
   2.3 Omics Data Analysis ................................................................. 5
3 Models of Lipid Metabolism in Yeast ................................................ 6
   3.1 Saccharomyces cerevisiae ........................................................... 6
   3.2 Oleaginous Yeasts ..................................................................... 9
4 Use of Yeast Models ...................................................................... 10
   4.1 Target Prediction ....................................................................... 10
   4.2 Omics Data Analysis ................................................................. 11
5 Research Needs ............................................................................ 12

References ....................................................................................... 13

Abstract

The various pathways and mechanisms behind fatty acid, lipid, and membrane biosynthesis together form a complex network. Moreover, lipid metabolism does not operate in isolation but rather functions in the context of a whole cell, surrounded by all its other metabolic pathways, a situation that results in additional connectivity and complexity. Computational models can aid to provide understanding of these complex networks and to make sense of interactions on a whole cell or genomic scale. In particular, these models have proven to be valuable to answer biotechnological questions, such as how to increase the biosynthesis of fatty acids. This chapter discusses the development and usage of genome-scale metabolic models in the light of lipid biosynthesis. A special
focus is placed on baker’s yeast *Saccharomyces cerevisiae* and the oleaginous yeast *Yarrowia lipolytica* and the use of genome-scale metabolic models to answer biotechnological questions.

1 Introduction

Cellular metabolism forms a complex network, where many metabolic pathways are functioning through enzymes that perform many different catalytic functions. All metabolic pathways are interconnected, by branching off each other, producing the same products or consuming the same substrates, or by using similar cofactors such as ATP and NADH. In order to gain understanding of these complex and highly connected networks, computational models can aid as valuable tools. A recurrent approach to modeling of metabolism is constraint-based modeling (Price et al. 2004), where metabolic fluxes are predicted from the stoichiometries of the metabolic reactions. When the genome is available of an organism of interest, one can identify all enzymes that make up the metabolic network – depending on the quality of the annotation – and generate a genome-scale constraint-based model.

Also the metabolic pathways constituting lipid biosynthesis are part of the complex network of cellular metabolism, connecting to other parts of metabolism through, e.g., precursors and cofactors such as acetyl-CoA and NADPH. By investigating lipid biosynthesis in the genome-scale context, rather than studying it in isolation, we preserve the functional interactions that take place between lipid biosynthesis and the remainder of cellular metabolism.

2 Genome-Scale Metabolic Models

2.1 Generation of Genome-Scale Model

Genome-scale metabolic models (GEMs) describe the whole metabolic network that is present in a particular cell and as such represents its metabolic capabilities: what reactions it can perform, which substrates it can use, and which by-products it creates. As the majority of the metabolic reactions are catalyzed by enzymes, knowledge about all the enzymes encoded in the organism’s genome is imperative to generate a new GEM. From the genome, it can be identified what proteins are present and what chemical reactions these proteins can catalyze (Fig. 1). These reactions together constitute a network, where substrates and products are nodes and reactions are edges. In addition to reactions identified from the genome sequence, unannotated and spontaneous reactions are included to generate a fully connected metabolic network. A wide range of computational tools have been developed that can assist in the construction of a GEM, such as SEED (Henry et al. 2010), RAVEN (Agren et al. 2013), and CoReCo (Pitkänen et al. 2014); however, the generation of a high-quality genome-scale model remains an iterative
process requiring intensive manual curation. The process of model construction is described in detail here (Thiele and Palsson 2010).

While the genome is used to reconstruct the metabolic network by identifying what reactions are present, also the association between genes, proteins, and reactions is stored as gene-protein-reaction (GPR) relationships. Among other things, these GPR relationships facilitate the integration of omics-level datasets with the GEM, as discussed in more detail below. The metabolic network itself can be represented in a stoichiometric matrix, where the coefficients correspond to the reaction coefficients (Fig. 1). This combination of an S-matrix and GPR relationships renders the GEM ready for further computational analysis.

### 2.2 Computational Analysis

The constraint-based modeling approach has given rise to a wide range of algorithms (Lewis et al. 2012). Underlying constraint-based modeling is the quasi-steady-state assumption (QSSA), which specifies that the system is in a state where the intracellular metabolite levels and metabolic fluxes are not changing (Fig. 2a). The QSSA can be assumed because metabolic rates are typically much faster than other cellular processes, such as cell division or depletion of nutrients from the environment.
Specifically, cells growing exponentially in a continuous culture can readily be presumed to be at steady state, but some caution is recommended when conditions are changing, such as during a batch fermentation. Notwithstanding, the QSSA allows balancing of metabolic fluxes in a network: if metabolite levels are constant, then the sum of the fluxes going through each metabolite has to sum up to zero (Fig. 2a). In addition to the S-matrix, upper and lower boundaries can be set for each reaction, e.g., based on experimentally determined uptake rates of nutrients, and reaction reversibility (Price et al. 2004). Methods have been developed to include additional constraints, such as the total cellular volume occupied by enzymes (Beg et al. 2007), or their kinetic efficiencies (Chakrabarti et al. 2013), but these will not be discussed here. Notwithstanding the sophistication of the applied constraints, it is now possible to determine how the flux is distributed through the metabolic network by balancing the fluxes, a method called flux balance analysis (Orth et al. 2010).

A complication with flux balance analysis (FBA) is that the metabolic network is typically large and of high complexity. This implies that instead of one unique solution, the metabolic fluxes can be balanced in numerous ways, all abiding to the same flux balancing rules and constraints that are imposed. Rather than one solution, a solution space can be defined (Fig. 2b). While all of the flux distributions within this space are feasible, not all are evenly likely. To select the most likely flux distribution, an objective function needs to be defined, and the most optimal flux distribution for this objective can then be calculated. Evolutionarily, it can be argued that microorganisms have evolved to grow as fast as possible in order to outcompete other microorganisms that compete for the same nutrients, and biomass production is
therefore often selected as an objective function in models of microorganisms. With growth as an objective function, it becomes important to describe the constitutive components of the organism’s biomass. It is worth noting that this equation will change from species to species, but potentially also from condition to condition (Feist and Palsson 2010). Other often used objective functions are ATP generation and product formation, while a combination of different objective functions seems most suitable to describe microbial metabolism (Schuetz et al. 2012).

Even when an objective function has been selected, there are typically still a wide range of flux distributions that all abide by the same constraints and satisfy the same objective. This can be analyzed by flux variability analysis (Mahadevan and Schilling 2003) and random sampling (Bordel et al. 2010), two methods that attempt to describe all possible flux values given the constraints and objective function. From this, mean and standard deviations can be described for each reaction, providing an overview of likely metabolic fluxes in a particular condition. The importance of exploring the solution space through such methods should be stressed, identifying whether fluxes change solely based on single FBA simulations ignores the remainder of the solution space, full of equally valid flux distributions.

Now able to estimate metabolic fluxes and their uncertainty, the power of in silico modeling is that such calculations can be done extremely rapidly in a computer. This opens the way to in silico experimentation, determining the effects of changes in the metabolic network that would otherwise take days to weeks to implement in the lab. Rather, the GEM can be queried for thousands of different changes, in less than a minute. For instance, this allows the GEM to predict targets for metabolic engineering, and for this various algorithms have been developed.

A classical approach predicting metabolic engineering targets is OptKnock (Burgard and Maranas 2003), where in silico gene knockouts are performed to identify their effect on a bioengineering objective, such as increased production of a compound of interest. In silico gene knockouts can easily be implemented in the GEM by restricting the flux through the associated reaction(s) to zero. With OptKnock, solutions are found that are maximizing both the cell growth and production of target compound. From this principle, a wide array of different optimization methods has been developed to aid strain design in metabolic engineering (Lewis et al. 2012).

2.3 Omics Data Analysis

As a GEM does not only contain a description of the network stoichiometry but also GPR relationships, GEMs can readily be used as scaffolds for omics data analysis. Changes in transcript and protein levels, determined through transcriptomics and proteomics techniques, are representing regulatory changes in the cell, and the GPR relationships allow to compare these regulatory changes with changes in metabolic fluxes that are estimated from FBA. Correlation between different omics levels and metabolic fluxes, or the absence of correlation, is then indicative of whether a
reaction is regulated on a transcriptional, posttranscriptional, or metabolic level (Bordel et al. 2010).

In an alternative approach, the network stoichiometry is used to identify so-called reporter metabolites (Patil and Nielsen 2005). Given a set of differentially expressed genes, and their GPR relationship to a metabolic network, metabolites can be identified around which the most significant transcriptional changes occur. Metabolites themselves can also be used as the input for analysis in the context of the metabolic network. The Kiwi package is capable to take a list of metabolites, which, for instance, have been measured to differ between experimental conditions, and combine this list with a GEM in order to extract an interaction network between these metabolites (Väremo et al. 2014). Various other methods have been developed to integrate omics data with GEMs (Sánchez and Nielsen 2015).

3 Models of Lipid Metabolism in Yeast

3.1 Saccharomyces cerevisiae

The first GEM for S. cerevisiae was published by Förster et al. (2003). This model, iFF708, contained information from 708 genes and a total of 1,175 reactions. In the years following, this model has been updated and improved in a series of models, a process that is discussed in detail here (Heavner and Price 2015). A major improvement was the inclusion of a more detailed description of lipid metabolism in iIN800 (Nookaew et al. 2008). This improvement increased the number of genes to 800 and the number of reactions to 1,446 and was primarily done through an extensive manual curation of the lipid metabolic pathways using various online databases and research papers. Specifically, iIN800 introduced mitochondrial fatty acid synthesis, fatty acid elongation, and updated beta-oxidation, sphingolipid biosynthesis, and biosynthesis of phospholipids, triacylglyceroles, and ergosterol. As such, with iIN800 it was now possible to query yeast lipid metabolism using a genome-scale model.

Development of yeast GEMs has not remained idle after this. In an attempt to combine the various modeling efforts and to assure consistency between future models, a consensus metabolic network was created: Yeast 1 (Herrgård et al. 2008). Again, this model was improved in various updates, while the latest major expansion in lipid metabolism was obtained with the introduction of Yeast 7 (Aung et al. 2013). In previous models, fatty acids and acyl-CoAs of different lengths and saturation were pooled, according to experimentally measured ratios, implemented in various ways in different model versions (Table 1). With Yeast 7, each unique acyl group, regarding length and saturation, was modeled explicitly. This has resulted in a massive expansion of the description of lipid metabolism. As example, where diacylglycerol acyltransferase used to be modeled as one reaction in the form of diacylglycerol + acyl-CoA \( \rightarrow \) triacylglycerol + CoA, these are now modeled with 32 reactions in the form of palmitoyl-CoA + diglyceride (1-16:0, 2-16:1) \( \rightarrow \) triglyceride (1-16:0, 2-16:1, 3-16:0) + CoA (Table 1). Nonetheless, all 32 reactions are still
Table 1  Evolution of lipid metabolism in *S. cerevisiae* genome-scale models. For a number of models is indicated how the last step of triacylglycerol biosynthesis is described. A increased complexity is observed with newer models. ACP: acyl-carrier protein; l: lipid droplet; r: endoplasmic reticulum. Square brackets indicate localization

<table>
<thead>
<tr>
<th>Model</th>
<th>Reaction</th>
<th>Equation</th>
<th>Gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>iFF708</td>
<td>U69_</td>
<td>Diacylglycerol + 0.017000 Decanoyl-[ACP] + 0.062000 Dodecanoyl-[ACP] + 0.100000 Myristoyl-[ACP] + 0.270000 Hexadecanoyl-[ACP] + 0.169000 Palmitoyl-[ACP] + 0.055000 Stearoyl-[ACP] + 0.235000 Oleoyl-[ACP] + 0.093000 Linolenoyl-[ACP] $\rightarrow$ Triacylglycerol + ACP</td>
<td></td>
</tr>
<tr>
<td>iN800</td>
<td>DGA1</td>
<td>Acyl-CoA + diacylglycerol $\rightarrow$ triacylglycerol + CoA</td>
<td>YOR245C</td>
</tr>
<tr>
<td>Pool_Acyl</td>
<td></td>
<td>0.008290 Decanoyl-CoA + 0.008277 Dodecanoyl-CoA + 0.019114 Tetradecanoyl-CoA + 0.005707 Tetradecanoyl-9-ene-CoA + 0.154233 Hexadecanoyl-9-ene-CoA + 0.037531 Octadecanoyl-CoA + 0.161896 Octadecanoyl-9-ene-CoA $\rightarrow$ Acyl-CoAs</td>
<td></td>
</tr>
<tr>
<td>1282</td>
<td>L-3-hydroxyhexacosanoyl-CoA[r] $\rightarrow$ acyl-CoA[r]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1283</td>
<td>L-3-hydroxypalmityl-CoA[r] $\rightarrow$ 0.78723 acyl-CoA[r]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1284</td>
<td>3-hydroxyoctadecanoyl-CoA[r] $\rightarrow$ 0.82979 acyl-CoA[r]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1285</td>
<td>L-3-hydroxytetradecanoyl-CoA[r] $\rightarrow$ 0.74468 acyl-CoA[r]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1288</td>
<td>3-oxohexacosanoyl-CoA[r] $\rightarrow$ acyl-CoA[r]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1289</td>
<td>3-oxopalmityl-CoA[r] $\rightarrow$ 0.78723 acyl-CoA[r]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1290</td>
<td>3-oxooctadecanoyl-CoA[r] $\rightarrow$ 0.82979 acyl-CoA[r]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1291</td>
<td>3-oxotetradecanoyl-CoA[r] $\rightarrow$ 0.74468 acyl-CoA[r]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1298</td>
<td>Hexacosanoyl-CoA[r] $\rightarrow$ acyl-CoA[r]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1302</td>
<td>Oleoyl-CoA[r] $\rightarrow$ 0.82979 acyl-CoA[r]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1305</td>
<td>Palmitoyl-CoA[r] $\rightarrow$ 0.78723 acyl-CoA[r]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1310</td>
<td>Tetraicosanoyl-CoA[r] $\rightarrow$ 0.95745 acyl-CoA[r]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1356</td>
<td>Palmitoleoyl-CoA[r] $\rightarrow$ 0.78723 acyl-CoA[r]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model</td>
<td>Reaction</td>
<td>Equation</td>
<td>Gene</td>
</tr>
<tr>
<td>-------</td>
<td>----------</td>
<td>----------</td>
<td>------</td>
</tr>
<tr>
<td>2416</td>
<td>Stearoyl-CoA[l] + diglyceride (1–16:0, 2–16:1)[l] → triglyceride (1–16:0, 2–16:1, 3–18:0)[l]</td>
<td>YOR245C</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
annotated to the same gene DGA1, as they are still all catalyzed by the same enzyme. Yeast 7 implemented this extension of lipid metabolism in combination with further manual curation and fixing localization of various pathways. The result of this expansion is that Yeast 7 includes 3,493 reactions and 2,218 metabolites (cf. 1,882 and 1,454 for Yeast 6, prelipid expansion), while the number of genes has increased only from 901 to 916, mainly due to other alterations.

While this more expansive description of lipid metabolism is a more accurate representation of what the metabolic network looks like in vivo, the expansion can be argued to have limited benefit in silico. The expansion increases the complexity of the model, and knowledge about the specific composition of each lipid is often unknown. Rather, experimental measurements typically measure lipid compositions in classes such as triacylglycerols, phosphatidylinositols, and sterol esters, and fatty acid composition is determined after hydrolyzation of the whole sample, giving an overview of all fatty acids in the cell.

Furthermore, the expansion of lipid metabolism introduced many alternative pathways within lipid metabolism, which leads to a large expansion of the solution space, resulting in more uncertainty about fluxes calculated by FBA. Interestingly, inclusion of kinetics and protein level constraints on Yeast 7 has shown to result to large decreases in flux variability in lipid metabolism (Sanchez et al. submitted). Combining the lipid expansion with such constraints is therefore preferred.

### 3.2 Oleaginous Yeasts

While there has always been much interest on lipid biosynthesis of *S. cerevisiae* as a model organism, from a biotechnological perspective, there are other yeasts that are arguably more suitable for the production of oleochemicals. While *S. cerevisiae* lipid content ranges from 5% to 20% in an engineered strain, oleaginous yeasts accumulate over 20% in the right conditions, while lipid contents of over 70% are also observed. The ascomycete *Yarrowia lipolytica* has been studied as model oleaginous yeast and as a promising microbial cell factory for the production of oleochemicals (Ledesma-Amaro and Nicaud 2016). To aid the study into oleaginous lipid metabolism, GEMs have also been constructed for *Y. lipolytica*. The first GEM was built.
using the then-existing *S. cerevisiae* GEMs as scaffolds (Loira et al. 2012), and also the *Y. lipolytica* GEMs have gone through a number of versions where improvements have been made. The most recent version (Kerkhoven et al. 2016) used Yeast 7 as a scaffold.

In addition to *Y. lipolytica*, GEMs have been generated for other oleaginous fungi, such as *Mortierella alpina* (Ye et al. 2015) and *Mucor circinelloides* (Vongsangnak et al. 2016). While all of these models are functional, capable of simulating the production of biomass and biosynthesis of lipids, a challenge with models of nonconventional yeasts is that the quality of their genome annotation is far lower than of model organisms such as *S. cerevisiae*. The functions of most of the genes in these organisms are solely annotated based on homology to other sequences, the presence of Pfam and other protein domains, and other computational approaches. In contrast, for *S. cerevisiae*, many of the enzymes have been characterized in vitro, generating proof of their catalytic capabilities. Due to the lower level of annotation, and the higher number of genes with unknown function in nonconventional yeasts, it is to be expected that the nonconventional yeast models are poorer representatives of reality in comparison to *S. cerevisiae* models. Annotated genes might have different substrate specificities, while unannotated genes might have uncharacterized and currently unrepresented catalytic capabilities. Notwithstanding these challenges in investigating lesser-studied organisms, GEMs of nonconventional yeasts such as *Y. lipolytica* have proven to be useful for further analysis, as detailed below.

4 Use of Yeast Models

Different yeast models have been used to investigate lipid biosynthesis using the different methods detailed above. While computational target prediction for metabolic engineering, such as with OptKnock, appears like a promising method to increase the microbial production of oleochemicals, there are de facto only a few examples of successful application. Rather, GEMs have been used to integrate omics data in investigations of lipid metabolism.

4.1 Target Prediction

The first *S. cerevisiae* model iFF708 was used to investigate how to increase the production of sesquiterpene, an isoprenoid that is produced in a competing pathway branching off ergosterol biosynthesis (Asadollahi et al. 2009). Previous strategies in increasing sesquiterpene production relied on the manipulation of classical targets within the mevalonate pathway that generates the precursor for isoprenoid biosynthesis. In contrast, Asadollahi et al. used the OptGene algorithm (Patil et al. 2005), an extension of OptKnock, to identify targets anywhere in the metabolic network that would increase sesquiterpene production. The intervention suggested by OptGene entailed deletion of glutamate dehydrogenase *GDH1*, which in silico resulted in a tenfold increase in sesquiterpene production alongside a 15% reduction in growth.
GDH1 catalyzes the reaction from alpha-ketoglutarate to glutamate, fixing ammonium while using NADPH as a cofactor. In an alternative pathway, ammonium can be fixed via the two-step GS-GOGAT pathway, mediated by glutamine synthetase (GS) and glutamate synthase (GOGAT), using NADH as a cofactor. Knocking out GDH1 redirects flux through the GS-GOGAT pathway, making NADPH available for other processes, such as NAPDH requiring reactions in the mevalonate pathway. Experimental confirmation of this approach indeed gave an 85% increase in final sesquiterpene titer.

Predicting interventions to increase product formation does not necessarily mean genetic engineering. For Y. lipolytica, analysis of batch fermentation data using a GEM gave rise to a hypothesis how lipid accumulation could be increased using a different feeding strategy (Kavšček et al. 2015). It was observed that in the end of the fermentation, when nitrogen had been depleted, the major product formed was citrate rather than lipids. Initially the model was only able to replicate this behavior if the citrate excretion was explicitly constrained to the measured rate. From this, it was hypothesized that lipid biosynthesis was at its maximum rate in these experimental conditions and any additional citrate that was formed but could not be used in lipid biosynthesis was rather excreted. While fixing the lipid biosynthesis rate in the model, the glucose uptake rate was decreased until citrate was no longer excreted. In silico, citrate excretion was abolished at a glucose uptake rate of 0.152 mmol (g·h)^{-1}, while the previous experiment had a glucose uptake rate of 0.350 mmol (g·h)^{-1}. Therefore, a fed-batch strategy was conceived, where the glucose feed was set to match the in silico calculated optimal feeding rate. Indeed, citrate excretion was completely abolished, while the TAG biosynthesis quadrupled to 0.203 g TAG per g glucose, whereas further increasing the glucose feeding rate reinstated citrate excretion.

4.2 Omics Data Analysis

The S. cerevisiae GEM iIN800 has been used to study the interaction of inositol-choline and Snf1 in controlling lipid metabolism (Chumnanpuen et al. 2012). In S. cerevisiae, many genes that are involved in lipid biosynthesis are under transcriptional control, and this has been known to be influenced by protein kinase Snf1, inositol-choline levels, and nutrient limitation. In this investigation, a multifactorial experimental setup was used to enquire the various interactions. The first factor was the level of inositol-choline, either high or low; the second factor was the absence (snf1Δ) and presence (WT) of SNF1; and the third factor was the presence of either carbon or nitrogen as limiting nutrient. Chemostat fermentations exploring all eight possible combinations of the three factors were performed and sampled for transcriptomics and lipid content. The transcriptomics data was integrated with iIN800, and the reporter metabolite analysis indicated where in the metabolic network the transcriptional regulation was concentrated, while reporter transcription factor analysis revealed that Snf1 and inositol-choline exert transcriptional control on lipid metabolism through transcription factors such as Ino2, Opi1, and Mga2.
In *Y. lipolytica*, the response of lipid accumulation to nitrogen limitation has been investigated by integrating transcriptomics data with the GEM iYali4 overexpressing *Y. lipolytica* strain, known to result in high lipid accumulation (Tai and Stephanopoulos 2013), was cultivated in both nitrogen-limiting and carbon limitation chemostats and sampled for transcriptomics and lipid content (Kerkhoven et al. 2016). Interestingly, while nitrogen limitation results in an increase in lipid accumulation, this was not mediated through an increased expression of lipid biosynthetic enzymes. Contrastingly, amino acid biosynthetic pathways and beta-oxidation were transcriptionally downregulated upon nitrogen limitation. A downregulation of amino acid biosynthesis can be rationalized as the cell attempting to reduce its nitrogen usage, while a lower expression of beta-oxidation makes sense in the light of nitrogen-limiting medium being rich in carbons, such that there is no reason to mobilize storage lipids as carbon source. These results on transcription level were then correlated with metabolic flux estimated from the GEM. The random sampling framework (Bordel et al. 2010) was used to support that the downregulation of amino acid biosynthetic enzymes resulted in a decreased flux. This could then be linked to the involvement of Snf1 but also TORC1.

### 5 Research Needs

While the argument for studying lipid metabolism in the context of the whole metabolic network is compelling, with current genome-scale metabolic models able to simulate metabolic fluxes through the complex lipid metabolic pathways and omics data sets being ever easier to generate, there are a number of issues that would thrive this area of research further.

One of the major issues is the accurate estimation of metabolic fluxes from GEMs, which is impeded by the vast solution space in purely stoichiometric models. There have been efforts to incorporate enzyme kinetics and protein constraints in the models, and these approaches have already resulted in large reductions of the solution space. However, any additional development in this direction, further tightening the flux predictions, would make the methods more valuable.

Another issue is the annotation and characterization of lipid biosynthetic pathways in different species. Especially from biotechnological viewpoint, there is a good argument to use nonconventional organisms, but this also means the absence of high-confidence knowledge on their metabolic networks and constituent enzymes. While computational approaches can go a long way to annotate genes, there remains a need for biochemical studies where enzyme activities are measured in vitro.

Lastly, while the incorporation of kinetics and metabolic constraints add more biological information into the models, there is another layer of information that can currently only be integrated poorly with GEMs. Besides the metabolic network, there is a complex regulatory network behind much of the behaviors of (lipid) metabolism. Improvements in these three areas would enhance the applicability of genome-scale models and as such would be able to answer questions in lipid metabolism.
Acknowledgments The author would like to acknowledge support from the US Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomic Science program (DE-SC0008744), and the Novo Nordisk Foundation.

References


