Chapter 27

Use of Genome-Scale Metabolic Models in Evolutionary Systems Biology

Balázs Papp, Balázs Szappanos, and Richard A. Notebaart

Abstract

One of the major aims of the nascent field of evolutionary systems biology is to test evolutionary hypotheses that are not only realistic from a population genetic point of view but also detailed in terms of molecular biology mechanisms. By providing a mapping between genotype and phenotype for hundreds of genes, genome-scale systems biology models of metabolic networks have already provided valuable insights into the evolution of metabolic gene contents and phenotypes of yeast and other microbial species. Here we review the recent use of these computational models to predict the fitness effect of mutations, genetic interactions, evolutionary outcomes, and to decipher the mechanisms of mutational robustness. While these studies have demonstrated that even simplified models of biochemical reaction networks can be highly informative for evolutionary analyses, they have also revealed the weakness of this modeling framework to quantitatively predict mutational effects, a challenge that needs to be addressed for future progress in evolutionary systems biology.

Key words: Flux balance analysis (FBA), constraint-based modeling, gene essentiality, genetic interaction, genome evolution, fitness landscape, metabolic network, *Saccharomyces cerevisiae*.

1. Introduction

Addressing many important questions in evolutionary biology relies on our understanding of the mapping between genotype and phenotype. Although evolutionary genetics analyses often use highly simplified and abstract genotype-fitness maps, recent advances in systems biology provide an unprecedented opportunity to calculate genotype–phenotype relationships using realistic mathematical models of molecular systems (1). With an increasing potential to computationally predict evolutionary relevant

J.I. Castrillo, S.G. Oliver (eds.), *Yeast Systems Biology*, Methods in Molecular Biology 759,
DOI 10.1007/978-1-61779-173-4_27,

parameters, such as the distribution of mutational effects and genetic interactions, molecular systems biology approaches begin to offer mechanistic insights into various topics from mutational robustness to genome evolution (2–6). Mathematical models that are most suited to address evolutionary questions can either directly calculate phenotypes that serve as fitness correlates (e.g., growth rate) or infer gene–gene relationships based on certain calculated phenotypes (e.g., identifying gene sets with correlated activities). These models include detailed kinetic models of metabolic pathways (7), regulatory circuits (cell cycle, circadian clock, etc.) (8), logical models of signaling networks (9), and constraint-based models of genome-scale metabolic networks (10). While most modeling approaches focus on small-scale biochemical systems (i.e., individual pathways) and characterize the mechanism of each enzymatic step, constraint-based models aim to calculate the metabolic behavior of relatively large systems (i.e., 600–1,300 genes) with relatively low data requirements. Moreover, these models are available for a number of microbial species, thereby providing a rigorous way to test evolutionary hypotheses.

The constraint-based framework uses mass balance and capacity constraints to define the space of all feasible steady-state flux distributions of the metabolic network leading from input (i.e., nutrient uptake) to output (an objective function, for instance, biomass production). Optimal network states are then identified within this space by maximizing or minimizing a certain metabolic objective function, an approach called flux balance analysis (FBA) (10, 11). However, the large size and comprehensive nature of these metabolic network models comes at a price as the framework lacks mechanistic details (e.g., kinetic rate constants and regulatory mechanisms), can only calculate steady-state patterns, and assumes that cells are fine-tuned from an evolutionary point of view. Furthermore, these models are restricted to simulate the effects of "large" mutations only (i.e., complete gene deletions or gene additions). Ultimately, the utility of genome-scale metabolic models for evolutionary analyses depends on how accurately they predict fitness correlates (e.g., growth phenotypes) and evolutionary relevant gene–gene relationships, a question that needs empirical investigation.

This chapter starts by discussing the use of constraint-based metabolic modeling in *Saccharomyces cerevisiae* to predict the effect of single and multiple mutations and hence to explore fitness landscapes (**Fig. 27.1a**). Fitness landscapes (or adaptive landscapes) visualize the relationship between genotypes and fitness and allow evolutionary biologists to investigate how mutations interact (epistasis) and which particular trajectories are taken during evolution. For example, the presence of multiple peaks on a fitness landscape indicates that some of the mutational paths to higher fitness alleles are selectively inaccessible (12) (**Fig. 27.1a**).

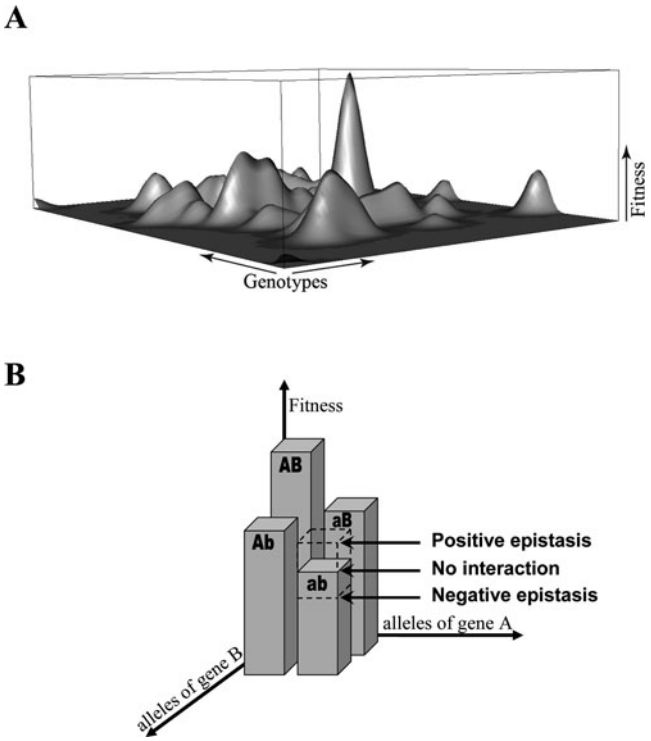

Fig. 27.1. Fitness landscape and epitasis studies. (**a**) An imaginary fitness landscape visualizing the relationship between genotype and fitness. The plane of the landscape contains all possible genotypes in such a way that similar genotypes are located close to each other on the plane and the height of the landscape reflects the fitness of the corresponding genotype. (**b**) Positive and negative epistases on a two-locus, two-allele genotype plane. Independence of gene action (no epistasis) is defined by a multiplicative model (i.e., when the fitness of the double mutant equals to the product of the fitnesses of the two single mutants).

Thus, to fully understand why particular evolutionary trajectories are realized and to what extent systems-level properties constrain the evolution of biochemical networks, we need detailed fitness landscapes of molecular systems. Next, we ask whether microbial metabolic network models have the potential to predict the outcome of evolutionary change, at least on short timescales, a question that has been addressed by laboratory evolution experiments in bacteria. Although most prior studies on evolutionary outcomes focused on *Escherichia coli* and other bacteria, we believe that similar approaches could also be adopted to analyze metabolic network evolution in baker's yeast. Finally, we analyze the shortcomings of constraint-based metabolic models to calculate weak fitness effects and genetic interactions, and discuss the utility of incorporating additional biological knowledge to increase their predictive performance. Although it is beyond the scope of the present review, we note that besides evolutionary

analyses, genome-scale metabolic network reconstructions have also been employed to draw ecological inferences, for example, to infer the habitable environments of different species (13) and to investigate the ecological strategies of bacteria (14).

2. Interrogating the Fitness Landscape: Predicting Mutational Effects

A straightforward systems biology strategy to begin to explore fitness landscapes is to computationally predict the effects of single mutations and their pair-wise epistatic interactions (i.e., the non-independence between the phenotypic effects of two mutations) (*see* **Fig. 27.1b**). Besides providing an intelligible abstraction of the high-dimensional adaptive landscape, accurate prediction of single- and double-mutant phenotypes can be used, among others, to investigate the nature and evolution of genetic robustness (3, 15) and can be harnessed to identify potential novel antimicrobial drug targets (16, 17).

2.1. Computing the Growth Effect of Single-Gene Deletions

The popular approach of flux balance analysis uses optimization principles to find one particular solution (i.e., flux distribution) among all possible metabolic network states that satisfy the governing physicochemical constraints (10). The objective function of the optimization protocol may be the rate of biomass formation (growth rate) or usually the biomass yield (i.e., the rate of biomass production divided by the rate of nutrient uptake), given limiting nutrients from the environment. Thus, the phenotype of wild-type and mutant strains can be computationally characterized by their (optimal) growth rates, a phenotype that can be easily measured in laboratory experiments. This strategy formed the basis of one of the first applications of the yeast genome-scale metabolic model (18), which systematically compared in silico growth of single-gene deletant strains with in vivo growth phenotypes on a qualitative scale (i.e., lethal or viable) (19). Taking essential genes (i.e., those whose disruption leads to lethality under standard laboratory conditions) as a reference, different versions of the yeast model have been reported to predict essential and non-essential genes with 83–90% accuracy (19–21). However, this high percentage of consistent phenotypes obscures the fact that essential genes are both less frequent in vivo and more difficult to predict than non-essential ones (21–23). Plotting the true-positive and false-positive rates of essentiality predictions (**Fig. 27.2**) reveals that one model (iLL672) is clearly superior to another (iND750) in terms of its capacity to discriminate between lethal and viable knockout phenotypes across 16

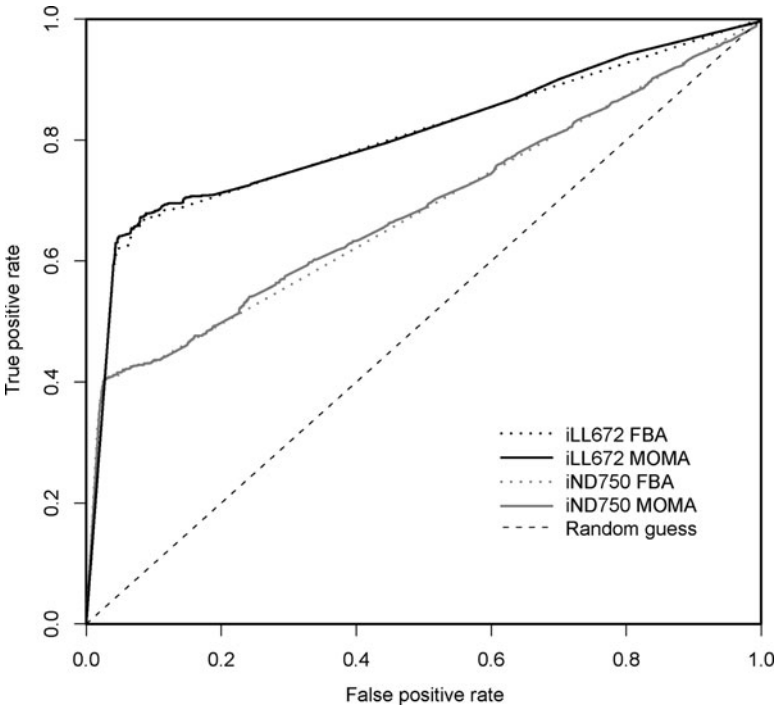

Fig. 27.2. Comparison of gene essentiality prediction performances of two different yeast metabolic network models (iND750 and iLL672) and two different optimization algorithms (FBA and MOMA) using receiver-operating characteristic (ROC) curves. ROC curves visualize classifier performance and can be employed to explore the trade-off between true-positive and false-positive rates at all possible cutoff levels (i.e., at different predicted growth rate cutoffs below which a strain is considered lethal). The closer the ROC plot is to the *upper left corner*, the higher the overall accuracy of the prediction. The *horizontal axis* represents the false-positive rate (number of true non-essential genes predicted as lethal/number of true non-essential genes), whereas the *vertical axis* represents the true-positive rate (number of correctly predicted true essential genes/number of true essential genes). We compiled gene essentiality data from (22), which measured growth phenotypes under 16 metabolically relevant conditions, and from (40), which identified genes that are essential even under nutrient-rich conditions. Computational predictions were taken from (22). Only genes present in both models were used for the comparison. We note that approximately one-third of the advantage of the iLL672 model against the iND750 model can be explained by differences in the biomass composition of the models (the area under the ROC curve of iLL672, iND750, and iND750 supplemented with the biomass composition of iLL672 is 0.8176, 0.6822, and 0.7241, respectively).

metabolically relevant growth conditions (*see* **Fig. 27.2** legend for details).

One conceptual reason why flux balance analysis might mispredict in vivo gene essentiality is that it assumes optimal network behavior even in gene knockout strains. To overcome this difficulty, other optimization criteria have been proposed that assume minimal flux reorganization in gene deletant strains with respect to the wild-type flux distribution (minimization of

metabolic adjustment, MOMA, and regulatory on/off minimization, ROOM, algorithms, (24, 25)). Although this approach yields a more accurate prediction of mutant flux distributions (26), it only slightly improves gene essentiality predictions (22) (*see* continuous and dotted lines in **Fig. 27.2**). Apparently, successful prediction of essential genes more critically depends on the proper formulation of biomass compositions (i.e., model output) (21) and on the completeness of our knowledge of the metabolic processes associated with these genes (23).

While constraint-based metabolic models predict the presence or the absence of mutant growth with good accuracy, one might wonder whether these models could also capture quantitative growth differences of viable mutants. In theory, FBA or MOMA analysis of mutant strains gives growth yields as outputs, i.e., the rate of biomass production divided by the rate of limiting nutrient uptake, therefore providing quantitative predictions. The availability of large-scale competitive fitness (27) and growth curve measurement data (28) for viable yeast knockouts offers an opportunity to contrast predicted and experimentally determined growth parameters for hundreds of metabolic gene deletants. In agreement with a prior small-scale study (26), we generally find weak correlations between in vivo competitive fitness or growth rate and *in silico* biomass production (predicted biomass yield) (**Fig. 27.3a–c**). Here, it should be noted that constraint-based metabolic models compute biomass yields (29) and a comparison of growth rate data to predicted biomass yields (**Fig. 27.3c**) might not be fully descriptive. Therefore, we also plotted experimentally measured growth efficiency (a proxy for growth yield) against in silico-predicted biomass yield (**Fig. 27.3d**) which resulted in an even weaker association. This suggests that the constraint-based modeling approach fails to capture quantitative growth differences in yeast mutants, at least in batch cultures with glucose-minimal (SD) or -rich (YPD) media. It should be noted, however, that *S. cerevisiae* displays repressed respiration when grown aerobically in excess glucose (30, 31). Indeed, this regulatory effect varies across gene deletion backgrounds (32, 33) and it could strongly affect mutant growth in a way that cannot be easily captured by a stoichiometric model. It remains to be seen whether model predictions better match empirical data when mutants are grown on non-fermentable carbon sources.

2.2. Predicting Genetic Interactions

The fitness effect of mutation in one gene might be modulated by mutations in other genes, a phenomenon called genetic interaction or epistasis (**Fig. 27.1b**). Negative genetic interaction occurs when two mutations decrease fitness more than would be expected based on their individual effects (the most drastic form is referred to as synthetic lethality), and for positive interactions, the opposite is true. Epistatic relations reveal functional associations

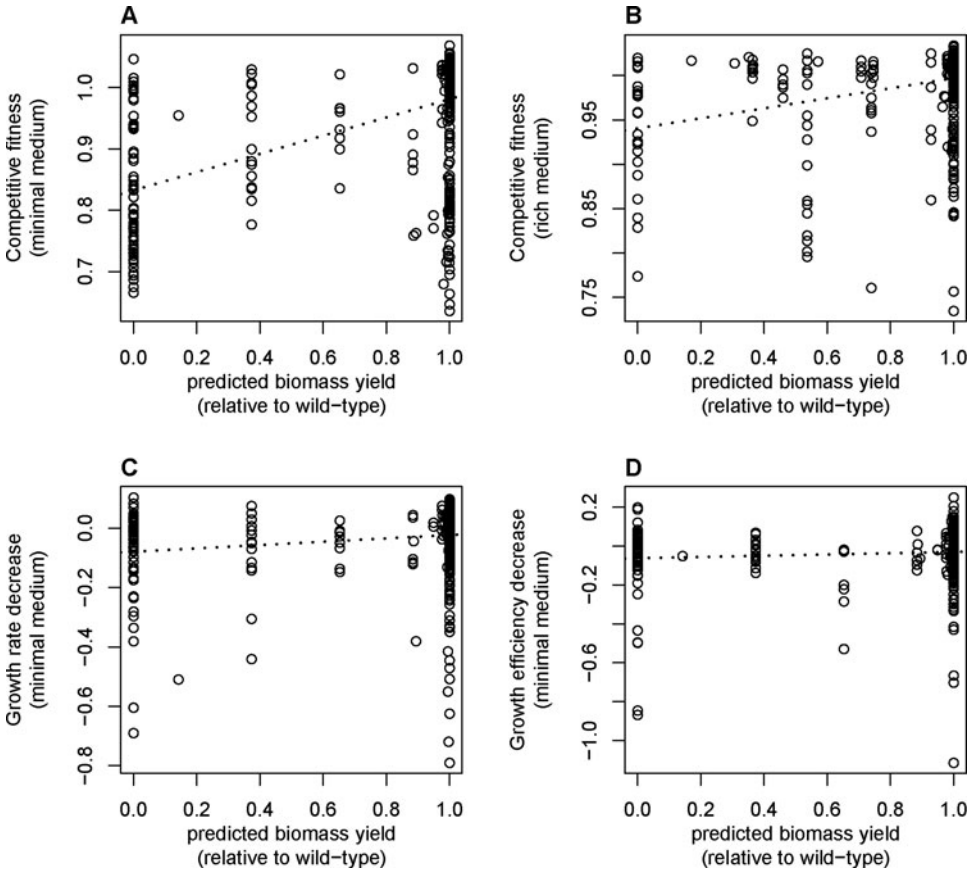

Fig. 27.3. Comparisons of in silico-predicted and experimentally determined growth parameters for viable metabolic gene deletion strains ($n = 499$). Computational predictions were taken from (22) and are based on the iLL672 model and MOMA algorithm (conclusions remain unchanged if prediction from the iND750 model or the FBA algorithm is used). (**a**) Comparison to competitive fitness data measured on glucose-minimal (SD) medium (27), Spearman's rho = 0.46, $p < 10^{-27}$. (**b**) Comparison to competitive fitness data measured on glucose-rich (YPD) medium (27), Spearman's rho = 0.26, $p < 10^{-8}$. (**c**) Comparison to growth rate data derived from growth curve measurements on SD (minimal medium) (28), Spearman's rho = 0.14, $p = 0.002$. (**d**) Comparison to growth efficiency data derived from growth curve measurements on SD (minimal medium) (28), Spearman's rho = 0.05, $p = 0.27$.

between genes (34) and influence many evolutionary processes (35), therefore it is of great importance to understand the molecular mechanisms underlying them and to develop reliable computational tools to predict them. Genome-scale metabolic models can rapidly calculate growth phenotypes for arbitrary sets of gene deletions, therefore, in principle, could be applied to systematically compute epistatic interactions between double or higher order gene knockouts. Indeed, yeast FBA models have been applied to compute both positive and negative pair-wise genetic interactions (36), and to identify synthetic lethality among multiple gene knockouts (37). However, relatively little is known about the accuracy of FBA models to capture different forms

and magnitudes of in vivo genetic interactions, and one might expect that if predictions of single-gene deletion phenotypes are not perfect, then those of multiple gene deletions would be even less accurate. A small-scale experimental validation of model-predicted synthetic lethal pairs (synthetic lethals) in *S. cerevisiae* showed that almost 50% of them were correct (15), which is much higher than would be expected by chance (<1%, based on (38)). However, the FBA model missed more than 75% of published synthetic lethals in the metabolic network, suggesting that the constraint-based framework underestimates the prevalence of negative genetic interactions. Furthermore, given the apparent failure of FBA to capture quantitative growth differences of single mutants, one might speculate that positive and weak genetic interactions would be predicted with even lower success rates than synthetic lethals. Future studies using quantitative epistatic data from large-scale genetic interaction screens (39) would be needed to rigorously assess the performance and limitations of constraint-based metabolic models to predict epistasis.

2.3. Understanding Gene Dispensability and Mutational Robustness

Large-scale, single-gene deletion screens have revealed that almost 80% of protein coding genes in *S. cerevisiae* seem not to be essential for viability under standard laboratory conditions (40), an observation that tallies with results from similar analyses performed in other organisms (41). This finding raises the questions of what the mechanistic basis of gene dispensability is and whether it is the result of an evolved capacity of genetic networks to compensate for mutations. It has been suggested that the high fraction of non-essential genes might reflect mutational robustness, i.e., the capacity to compensate for mutations by using either redundant gene duplicates or alternative biochemical pathways (42). A second possibility is that seemingly dispensable genes are simply not active in the tested environmental condition(s), although they have important fitness contributions under special conditions (3). Computational systems biology models that can reliably predict the viability of single-gene deletants hold the promise to provide mechanistic explanations for gene dispensability. Indeed, a flux balance analysis model of yeast metabolism showed that a large fraction of non-essential enzymatic genes catalyze reactions that are inactive under the tested condition (i.e., carry zero flux), hence there might be no need to invoke any compensatory mechanism to explain their dispensability (3, 43). Furthermore, according to the model, genes that are active but not essential are mostly compensated by redundant gene duplicates and not by alternative pathways. It should be noted, however, that FBA models assume optimal network behavior and are therefore likely to overestimate the number of in vivo inactive reactions (i.e., suboptimal pathways are completely silenced in model solutions) (44). Nevertheless, some of the above computational predictions have

been confirmed experimentally: ^{13}C-flux analysis of viable yeast knockouts showed that flux rerouting through alternative pathways explains only the minority of non-essential genes (43). Furthermore, simulating gene deletions under a number of different nutrient-limiting conditions predicted that many functionally inactive genes would become essential under other conditions (3). A large-scale chemical genomic assay in yeast provided strong support for this notion: 97% of gene deletions exhibited a measurable growth phenotype in at least one of hundreds of tested conditions compared to only 34% in rich medium (45).

Although the above computational and experimental studies suggest that gene dispensability is only apparent and is mostly explained by condition-specific gene functions, other empirical works showed that most non-essential genes display synthetic lethal interactions with some other genes (23). As synthetic lethal genetic interactions indicate compensation between two genes (i.e., mutational robustness), this raises the question as to how these seemingly contradictory findings can be reconciled. A flux balance analysis study of synthetic genetic interactions under a large number of environmental conditions demonstrated that the capacity to compensate null mutations varies substantially between different nutritional environments (15). More specifically, it has been shown computationally, and confirmed by double deletion experiments, that synthetic lethal interactions are often restricted to particular environmental conditions, partly because genes that are compensated in one condition make an essential fitness contribution in another condition. Further empirical studies on yeast gene duplicates corroborated the widespread condition dependency of mutational compensation (46, 47).

The above findings also offer indirect insights into the selective forces shaping metabolic network evolution. Instead of regarding apparent redundancies as adaptations against harmful mutations (42), the presence of distinct but functionally overlapping metabolic pathways more likely reflects the outcome of an evolutionary adaptation characterized by selection for growth in varying environments (i.e., various different nutrients). As a correlated response, some of these pathways may also increase mutational resilience under some conditions (15, 48).

3. Predicting Evolutionary Outcomes: FBA as an Evolutionary Optimization Model

How predictable is evolution? Evolutionary change is often considered to be contingent on initial conditions and chance events and therefore unique on the one hand, and replicable owing to predictable adaptive changes on the other hand (49). Systems

biology modeling offers novel ways to investigate the predictability of evolutionary outcomes and organismal diversity. In particular, metabolic network models have recently been employed to test various hypotheses regarding the end state of evolutionary process. First, flux balance analysis of metabolic networks gives specific predictions on the steady-state behavior of evolutionary adapted metabolic systems. Second, by accounting for systems-level gene functions and relationships between genes, constraint-based metabolic models also have the power to predict interspecies differences in metabolic gene content, therefore to explain comparative genomics patterns. The latter includes predicting genes most likely undergoing loss and horizontal transfer events (50), asymmetric gain or loss of enzyme pairs (4), and gene contents of reduced genome endosymbiotic bacteria based on knowledge of its distant ancestors and its current lifestyle (5). Here, we restrict our attention to the use of FBA models to generate testable hypotheses on the outcome of short-term adaptive evolution.

As mentioned above, flux balance analysis uses optimization principles to find one particular network state that maximizes biomass production, that is, cellular growth. Because growth can be considered as a fitness correlate in microbes, FBA models can be seen as models about adaptation (51) in which in vivo metabolic states are sought that maximize organismal fitness. Microbial metabolism is optimized by the process of adaptive evolution, therefore FBA has, in principle, the potential to predict the outcome of evolutionary adaptation and give insight into the constraints that influence adaptation. An essential step in optimality approaches is to test the model predictions against empirical observations to reveal the particular selective forces and constraints that might have played significant roles during the evolutionary history of the organism under study.

Various experimental works have been performed to evaluate the power of FBA to predict the outcome of both natural and laboratory evolution. For example, it has been demonstrated that in vivo and *in silico* flux distributions are consistent under certain environmental conditions in *E. coli* (11, 25), suggesting that maximizing biomass production might have been an important selective force in the history of *E. coli*. However, simple FBA models fail to explain the metabolic behavior of microbes that do not metabolize nutrients most efficiently (29, 48). For instance, *S. cerevisiae* uses a mixture of respiration (high-yield route) and fermentation (low-yield route) to utilize glucose even under aerobic conditions when glucose is abundant in the medium (29). Applying alternative objective functions instead of biomass yield (52) or using game-theoretical approaches (53, 54), that is, formulating the optimization problem as frequency dependent instead of frequency independent, could help to resolve discrepancies between model predictions and experimental observations.

Suboptimal metabolic behavior of wild-type strains might also stem from incomplete adaptedness to the tested conditions and experiments have been designed to adapt strains under specific growth selection pressures in the laboratory to test whether evolved strains display *in silico*-predicted growth properties. One such adaptive evolution experiment has been performed in *E. coli*, with growth rate as the selection criterion and glycerol uptake as the main carbon source to produce biomass components (55). Cells initially grew sub-optimally on glycerol, but adapted over a period of ~60 days (1,000 generations), toward the FBA-predicted optimal behavior. One would expect that such an adaptation event should be reflected in the genotype and this has been demonstrated by whole-genome resequencing of evolved strains. In particular, mutations were identified in the glycerol kinase gene, which is clearly associated with the growth environment (56). Adaptation to utilize glycerol has also been predicted, and experimentally verified, for a lactic acid bacteria, showing extremely low initial growth rates in a glycerol environment (57). Growth of lactic acid bacteria with glycerol as the main carbon source has never been demonstrated experimentally before, even though the metabolic model predicted this output phenotype. Similarly, FBA can be used to predict the result of evolutionary adaptation in response to gene deletions, that is, the outcome of compensatory evolution. For example, adaptive evolution of *E. coli* strains carrying metabolic gene deletions resulted in increased growth rates that were similar to those predicted by FBA (in 78% of the strains tested) (58). Taken together, these studies clearly demonstrate that FBA has the potential to predict the outcome of adaptive evolution at the phenotype level.

4. Future Challenges

Constraint-based models of microbial genome-scale metabolic reconstructions present a simple computational framework to explore the metabolic capacity of wild-type and mutant strains under different environmental conditions, thereby providing a mapping between genotype and metabolic phenotype. Despite its simplicity and dependence on optimality principles, these models proved successful to compute the viability of mutant strains, to make testable predictions on gene loss and gene gain patterns on the phylogenetic tree, and, in some cases, to predict the phenotypic outcome of short-term adaptive evolution. However, the same approach appears to perform poorly in predicting weak growth effects of mutations in yeast, at least on glucose media. Furthermore, it remains to be seen how accurately in vivo genetic

interactions can be captured within this framework. Given that weak fitness effects and epistatic interactions are especially important to understand the process of evolution and are more frequent than strong effects, there is a great need to develop computational approaches that can accurately describe mutations with weak phenotypic impacts (1). We now briefly discuss some possible strategies to improve the predictive power of the constraint-based framework. First, constraint-based metabolic models can be made more realistic by imposing additional relevant constraints to decrease the solution space. Some of the proposed extra constraints include thermodynamic constraints (59) (i.e., elimination of thermodynamically infeasible solutions) and regulatory constraints (60) (i.e., elimination of reactions that are repressed under a given condition). With the rapid ongoing development of high-throughput techniques that accumulate data on intracellular metabolite concentrations, mRNA, protein expression levels, and reaction fluxes (61, 62), these additional constraints could be routinely applied in future studies (*see* (63, 64)). Second, new algorithms to compute the immediate physiological effect of mutations need to be developed and tested. Clearly, the assumption of optimal growth is not tenable for mutants and some methods have been put forward to describe metabolic states after gene removal (24, 25). However, these modified optimization algorithms are largely ad hoc, and it remains to be seen whether more realistic alternative methods can be developed based on empirical data on physiological changes following gene deletions, including high-throughput data on alterations in growth properties (28), metabolic footprints (65), intracellular fluxes (43), and mRNA expression (66). Furthermore, by assuming maximal biomass production in the wild type, the FBA approach is unable to capture beneficial loss-of-function mutations (only the addition of new reactions could increase *in silico* growth in this framework). This particular shortage of FBA could be alleviated only by developing new algorithms to calculate wild-type growth behavior. Third, there are efforts underway to reconcile constraint-based and kinetic modeling approaches in order to build large-scale dynamic models of cellular metabolism without the need for extensive experimental data (67, 68). Such hybrid frameworks would allow the piecewise incorporation of both additional flux constraints and information on enzyme kinetics when they become available and would hopefully improve the predictive power of large-scale models in determining metabolic responses to perturbations. Given the tremendous efforts put into generating functional genomics and comparative data and inferring cellular networks in yeasts, we expect that *S. cerevisiae* would be at the forefront of developing new generations of genome-scale metabolic models and applying them to evolutionary questions.

Acknowledgments

We thank Csaba Pál for suggestions on the manuscript and Kiran Patil and Juan I. Castrillo for comments on the issue of quantitative fitness predictions with FBA. B.P. is supported by The International Human Frontier Science Program Organization, the Hungarian Scientific Research Fund (OTKA), the "Lendület Program," and the Bolyai Fellowship of the Hungarian Academy of Sciences. R.N. is supported by The Netherlands Genomics Initiative (NGI – Horizon grant) and The Netherlands Organisation for Scientific Research (NWO – VENI Grant).

References

1. Loewe, L. (2009) A framework for evolutionary systems biology. *BMC Syst. Biol.* **3**, 27.
2. Endy, D., You, L., Yin, J., and Molineux, I. J. (2000) Computation, prediction, and experimental tests of fitness for bacteriophage T7 mutants with permuted genomes. *Proc. Natl. Acad. Sci. USA* **97**, 5375–5380.
3. Papp, B., Pál, C., and Hurst, L. D. (2004) Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast. *Nature* **429**, 661–664.
4. Notebaart, R. A., Kensche, P. R., Huynen, M. A., and Dutilh, B. E. (2009) Asymmetric relationships between proteins shape genome evolution. *Genome Biol.* **10**, R19.
5. Pál, C., Papp, B., Lercher, M. J., Csermely, P., Oliver, S. G., and Hurst, L. D. (2006) Chance and necessity in the evolution of minimal metabolic networks. *Nature* **440**, 667–670.
6. Loewe, L., and Hillston, J. (2008) The distribution of mutational effects on fitness in a simple circadian clock. *Lect. Notes Bioinf.* **5307**, 156–175.
7. Teusink, B., Walsh, M. C., van Dam, K., and Westerhoff, H. V. (1998) The danger of metabolic pathways with turbo design. *Trends Biochem. Sci.* **23**, 162–169.
8. Chen, K. C., Calzone, L., Csikasz-Nagy, A., Cross, F. R., Novak, B., and Tyson, J. J. (2004) Integrative analysis of cell cycle control in budding yeast. *Mol. Biol. Cell* **15**, 3841–3862.
9. Christensen, T. S., Oliveira, A. P., and Nielsen, J. (2009) Reconstruction and logical modeling of glucose repression signaling pathways in *Saccharomyces cerevisiae*. *BMC Syst. Biol.* **3**, 7.
10. Price, N. D., Reed, J. L., and Palsson, B. O. (2004) Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat. Rev. Microbiol.* **2**, 886–897.
11. Kauffman, K. J., Prakash, P., and Edwards, J. S. (2003) Advances in flux balance analysis. *Curr. Opin. Biotechnol.* **14**, 491–496.
12. Poelwijk, F. J., Kiviet, D. J., Weinreich, D. M., and Tans, S. J. (2007) Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* **445**, 383–386.
13. Borenstein, E., Kupiec, M., Feldman, M. W., and Ruppin, E. (2008) Large-scale reconstruction and phylogenetic analysis of metabolic environments. *Proc. Natl. Acad. Sci. USA* **105**, 14482–14487.
14. Freilich, S., Kreimer, A., Borenstein, E., et al. (2009) Metabolic-network-driven analysis of bacterial ecological strategies. *Genome Biol.* **10**, R61.
15. Harrison, R., Papp, B., Pal, C., Oliver, S. G., and Delneri, D. (2007) Plasticity of genetic interactions in metabolic networks of yeast. *Proc. Natl. Acad. Sci. USA* **104**, 2307–2312.
16. Raman, K., Rajagopalan, P., and Chandra, N. (2005) Flux balance analysis of mycolic acid pathway: targets for anti-tubercular drugs. *PLoS Comput. Biol.* **1**, e46.
17. Lee, D. S., Burd, H., Liu, J., et al. (2009) Comparative genome-scale metabolic reconstruction and flux balance analysis of multiple *Staphylococcus aureus* genomes identify novel antimicrobial drug targets. *J. Bacteriol.* **191**, 4015–4024.
18. Forster, J., Famili, I., Fu, P., Palsson, B. O., and Nielsen, J. (2003) Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res.* **13**, 244–253.

19. Forster, J., Famili, I., Palsson, B. O., and Nielsen, J. (2003) Large-scale evaluation of in silico gene deletions in *Saccharomyces cerevisiae*. *Omics* **7**, 193–202.
20. Duarte, N. C., Herrgard, M. J., and Palsson, B. O. (2004) Reconstruction and validation of *Saccharomyces cerevisiae* iND750, a fully compartmentalized genome-scale metabolic model. *Genome Res.* **14**, 1298–1309.
21. Kuepfer, L., Sauer, U., and Blank, L. M. (2005) Metabolic functions of duplicate genes in *Saccharomyces cerevisiae*. *Genome Res.* **15**, 1421–1430.
22. Snitkin, E. S., Dudley, A. M., Janse, D. M., Wong, K., Church, G. M., and Segre, D. (2008) Model-driven analysis of experimentally determined growth phenotypes for 465 yeast gene deletion mutants under 16 different conditions. *Genome Biol.* **9**, R140.
23. Becker, S. A., and Palsson, B. O. (2008) Three factors underlying incorrect in silico predictions of essential metabolic genes. *BMC Syst. Biol.* **2**, 14.
24. Segrè, D., Vitkup, D., and Church, G. M. (2002) Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl. Acad. Sci. USA* **99**, 15112–15117.
25. Shlomi, T., Berkman, O., and Ruppin, E. (2005) Regulatory on/off minimization of metabolic flux changes after genetic perturbations. *Proc. Natl. Acad. Sci. USA* **102**, 7695–7700.
26. Snitkin, E. S., and Segre, D. (2008) Optimality criteria for the prediction of metabolic fluxes in yeast mutants. *Genome Inform.* **20**, 123–134.
27. Deutschbauer, A. M., Jaramillo, D. F., Proctor, M., et al. (2005) Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. *Genetics* **169**, 1915–1925
28. Warringer, J., Ericson, E., Fernandez, L., Nerman, O., and Blomberg, A. (2003) High-resolution yeast phenomics resolves different physiological features in the saline response. *Proc. Natl. Acad. Sci. USA* **100**, 15724–15729.
29. Schuster, S., Pfeiffer, T., and Fell, D. A. (2008) Is maximization of molar yield in metabolic networks favoured by evolution? *J. Theor. Biol.* **252**, 497–504.
30. Gancedo, J. M. (1998) Yeast carbon catabolite repression. *Microbiol. Mol. Biol. Rev.* **62**, 334–361.
31. Sonnleitner, B., and Kappeli, O. (1986) Growth of *Saccharomyces cerevisiae* is controlled by its limited respiratory capacity: formulation and verification of a hypothesis. *Biotechnol. Bioeng.* **28**, 927–937.
32. Schuurmans, J. M., Boorsma, A., Lascaris, R., Hellingwerf, K. J., and Teixeira de Mattos, M. J. (2008) Physiological and transcriptional characterization of *Saccharomyces cerevisiae* strains with modified expression of catabolic regulators. *FEMS Yeast Res.* **8**, 26–34.
33. Usaite, R., Jewett, M. C., Oliveira, A. P., Yates, J. R., 3rd, Olsson, L., and Nielsen, J. (2009) Reconstruction of the yeast Snf1 kinase regulatory network reveals its role as a global energy regulator. *Mol. Syst. Biol.* **5**, 319.
34. Boone, C., Bussey, H., and Andrews, B. J. (2007) Exploring genetic interactions and networks with yeast. *Nat. Rev. Genet.* **8**, 437–449.
35. Wolf, J. B., Brodie, E. D., and Wade, M. J. (2000) *Epistasis and the Evolutionary Process.* New York, NY: Oxford University Press.
36. Segrè, D., Deluna, A., Church, G. M., and Kishony, R. (2005) Modular epistasis in yeast metabolism. *Nat. Genet.* **37**, 77–83.
37. Deutscher, D., Meilijson, I., Kupiec, M., and Ruppin, E. (2006) Multiple knockout analysis of genetic robustness in the yeast metabolic network. *Nat. Genet.* **38**, 993–998.
38. Tong, A. H., Lesage, G., Bader, G. D., et al. (2004) Global mapping of the yeast genetic interaction network. *Science* **303**, 808–813.
39. Schuldiner, M., Collins, S. R., Thompson, N. J., et al. (2005) Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. *Cell* **123**, 507–519.
40. Giaever, G., Chu, A. M., Ni, L., et al. (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**, 387–391.
41. Hurst, L. D., and Pál, C. (2007) Genomic redundancy and dispensability. In: Pagel, M., and Pomiankowski, A. (eds.), *Evolutionary Genomics and Proteomics* (pp. 141–160). Sunderland, MA: Sinauer Associates Inc.
42. Wagner, A. (2000) Robustness against mutations in genetic networks of yeast. *Nat. Genet.* **24**, 355–361.
43. Blank, L. M., Kuepfer, L., and Sauer, U. (2005) Large-scale 13C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome Biol.* **6**, R49.
44. Nishikawa, T., Gulbahce, N., and Motter, A. E. (2008) Spontaneous reaction silencing in metabolic optimization. *PLoS Comput. Biol.* **4**, e1000236.
45. Hillenmeyer, M. E., Fung, E., Wildenhain, J., et al. (2008) The chemical genomic

portrait of yeast: uncovering a phenotype for all genes. *Science* **320**, 362–365.
46. Ihmels, J., Collins, S. R., Schuldiner, M., Krogan, N. J., and Weissman, J. S. (2007) Backup without redundancy: genetic interactions reveal the cost of duplicate gene loss. *Mol. Syst. Biol.* **3**, 86.
47. Musso, G., Costanzo, M., Huangfu, M., et al. (2008) The extensive and condition-dependent nature of epistasis among whole-genome duplicates in yeast. *Genome Res.* **18**, 1092–1099.
48. Papp, B., Teusink, B., and Notebaart, R. A. (2009) A critical view of metabolic network adaptations. *HFSP J.* **3**, 24–35.
49. Travisano, M., Mongold, J. A., Bennett, A. F., and Lenski, R. E. (1995) Experimental tests of the roles of adaptation, chance, and history in evolution. *Science* **267**, 87–90.
50. Pál, C., Papp, B., and Lercher, M. J. (2005) Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nat. Genet.* **37**, 1372–1375.
51. Parker, G. A., and Smith, J. M. (1990) Optimality theory in evolutionary biology. *Nature* **348**, 27–33.
52. Schuetz, R., Kuepfer, L., and Sauer, U. (2007) Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Mol. Syst. Biol.* **3**, 119.
53. Pfeiffer, T., and Schuster, S. (2005) Game-theoretical approaches to studying the evolution of biochemical systems. *Trends Biochem. Sci.* **30**, 20–25.
54. MacLean, R. C. (2008) The tragedy of the commons in microbial populations: insights from theoretical, comparative and experimental studies. *Heredity* **100**, 471–477.
55. Ibarra, R. U., Edwards, J. S., and Palsson, B. O. (2002) Escherichia coli K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* **420**, 186–189.
56. Herring, C. D., Raghunathan, A., Honisch, C., et al. (2006) Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nat. Genet.* **38**, 1406–1412.
57. Teusink, B., Wiersma, A., Jacobs, L., Notebaart, R. A., and Smid, E. J. (2009) Understanding the adaptive growth strategy of *Lactobacillus plantarum* by in silico optimisation. *PLoS Comput. Biol.* **5**, e1000410.
58. Fong, S. S., and Palsson, B. O. (2004) Metabolic gene-deletion strains of *Escherichia coli* evolve to computationally predicted growth phenotypes. *Nat. Genet.* **36**, 1056–1058.
59. Beard, D. A., Liang, S. D., and Qian, H. (2002) Energy balance for analysis of complex metabolic networks. *Biophys. J.* **83**, 79–86.
60. Covert, M. W., Schilling, C. H., and Palsson, B. (2001) Regulation of gene expression in flux balance models of metabolism. *J. Theor. Biol.* **213**, 73–88.
61. Sauer, U. (2006) Metabolic networks in motion: 13C-based flux analysis. *Mol. Syst. Biol.* **2**, 62.
62. Ishii, N., Nakahigashi, K., Baba, T., et al. (2007) Multiple high-throughput analyses monitor the response of *E. coli* to perturbations. *Science* **316**, 593–597.
63. Akesson, M., Forster, J., and Nielsen, J. (2004) Integration of gene expression data into genome-scale metabolic models. *Metab. Eng.* **6**, 285–293.
64. Hoppe, A., Hoffmann, S., and Holzhutter, H. G. (2007) Including metabolite concentrations into flux balance analysis: thermodynamic realizability as a constraint on flux distributions in metabolic networks. *BMC Syst. Biol.* **1**, 23.
65. Allen, J., Davey, H. M., Broadhurst, D., et al. (2003) High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. *Nat. Biotechnol.* **21**, 692–696.
66. Hughes, T. R., Marton, M. J., Jones, A. R., et al. (2000) Functional discovery via a compendium of expression profiles. *Cell* **102**, 109–126.
67. Smallbone, K., Simeonidis, E., Broomhead, D. S., and Kell, D. B. (2007) Something from nothing: bridging the gap between constraint-based and kinetic modelling. *FEBS J* **274**, 5576–5585.
68. Covert, M. W., Xiao, N., Chen, T. J., and Karr, J. R. (2008) Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. *Bioinformatics* **24**, 2044–2050.