45
Metabolic Network Modeling for Computer-Aided Design of Microbial Interactions

Hyun-Seob Song, William C. Nelson, Joon-Yong Lee, Ronald C. Taylor, Christopher S. Henry, Alexander S. Beliaev, Doraiswami Ramkrishna, and Hans C. Bernstein

Synthetic biology has laid a foundation for integrating fundamental biology and genetics with engineering principles. Significant modifications of metabolic pathways in a single organism, however, often lead to undesirable outcomes such as impaired growth (due to toxicity or metabolic burden) or loss of robustness. The use of microbial communities is drawing attention as an alternative. Microbes are social organisms that exhibit intriguing interspecies dynamics and interactions. Therefore, the key of engineering at the multispecies level lies in the design of interactions. Predictive *in silico* tools such as metabolic network modeling are essential for the rational design of controllable microbial interactions.

45.1
Biological Computer-Aided Design of *Interactions*

Computer-aided design (CAD) practices are principal to all fields of engineering. Metabolic engineers and synthetic biologists have also demonstrated success using various biological CAD (BioCAD) approaches for designing metabolic pathways, genetic devices, and proteins [1]. In the application of these *in silico* tools for engineering microbial communities, we need a nascent concept and workflow such as what we call BioCAD of *interactions* (BioCAD*i*) that particularly focuses on the rational design of microbial interactions (Figure 45.1). While the full potential of mathematical models as design tools is only now beginning to be realized, their future prospects are promising [2].

The BioCAD*i* framework builds upon established computational approaches, which are broadly classified into top-down and bottom-up methods (Figure 45.2). Top-down approaches use data and information obtainable at a community level to infer microbial interactions. Similarity-based approaches have been widely applied to infer microbial interactions over a range of microbial ecosystems including soils, marine environments, and humans [3]. Beyond pair-wise interactions, regression-based methods predict the relationship among multiple

Emerging Areas in Bioengineering, First Edition. Edited by Ho Nam Chang.

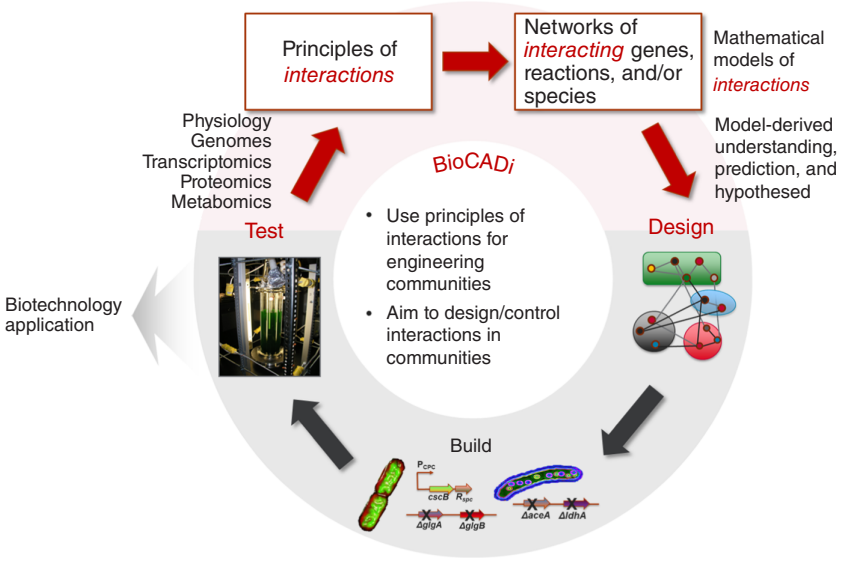

Figure 45.1 A schematic illustration of the concept of BioCAD*i* in the typical engineering cycle of design–build–test.

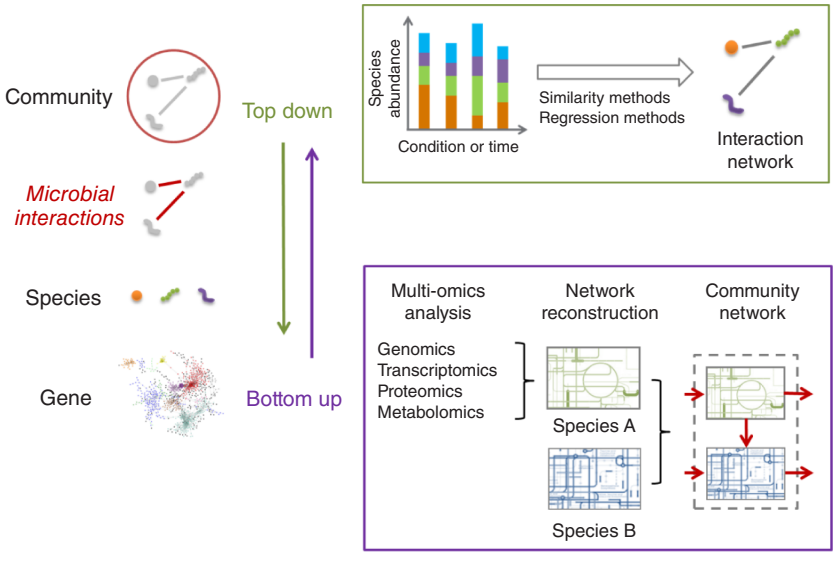

Figure 45.2 Top-down and bottom-up approaches for predicting microbial interactions. The right panel compares the conventional top-down prediction using similarity- and regression-based methods with the bottom-up prediction using the metabolic network simulation.

components, including biotic (i.e., species) and abiotic variables (i.e., environmental parameters) [4]. In contrast, bottom-up approaches such as metabolic network analyses use information on gene sequences and metabolic characteristics in individual species to predict microbial interactions in a community. While our focus is placed on the capabilities of metabolic network analyses in this chapter, the integration of bottom-up and top-down approaches offers an important means for testing and generating new hypotheses for the mechanisms and ecological consequences of microbial interactions [5].

45.2 Community Metabolic Network Reconstruction

Community metabolic network models can be built by combining the networks of individual organisms. For communities formed by combining axenically cultivable species, genome-scale metabolic networks can be reconstructed for individual organisms by the standard procedures shown in Figure 45.3. Pipelines such as

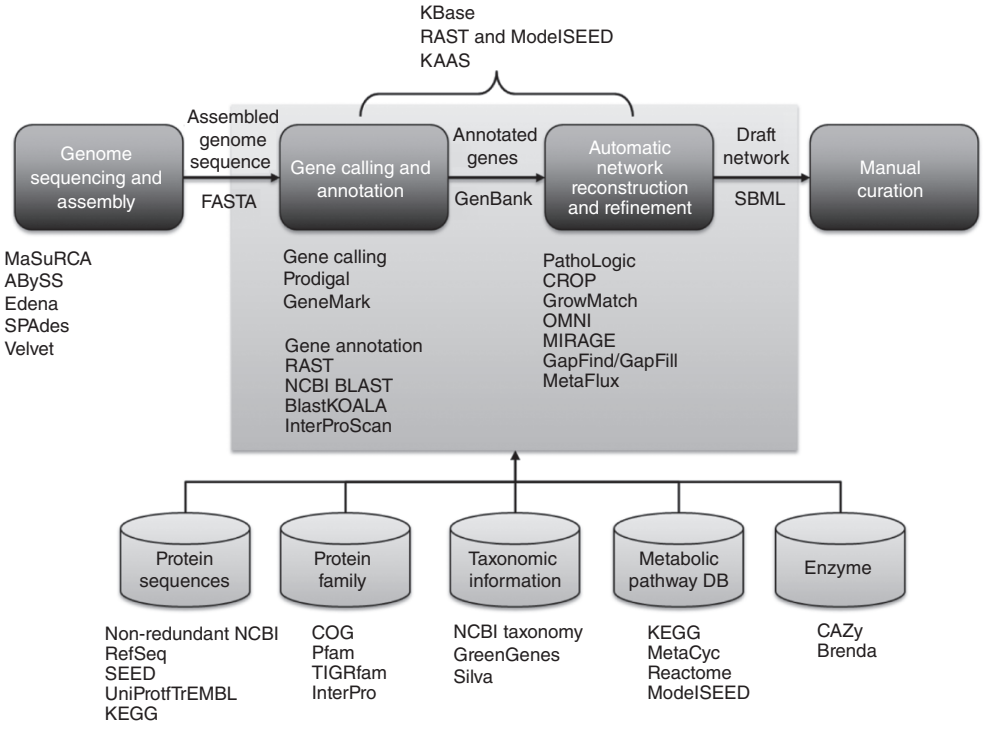

Figure 45.3 Overall procedures of metabolic network reconstruction: gene sequencing and assembly, gene calling and annotation, semiautomatic network reconstruction, and finally manual curation.

Rapid Annotations using Subsystems Technology (RAST) with ModelSEED [6, 7] and Systems Biology Knowledgebase (www.kbase.us) enable a researcher to automatically generate a draft metabolic network of a target organism and computationally improve this network through gap filling and dead-end detection [8]. Manual curation is necessary for improving the model's accuracy and reliability [9].

For environmental communities whose members cannot be well characterized in isolation, metabolic network building becomes more challenging [10]. While it is straightforward to collect community sequence data (metagenome), assembly is frequently inefficient, and there is an additional requirement of segregating contigs into species-specific groups, referred to as binning or genome reconstruction. The grouped contigs can be used as the starting point for building species-specific metabolic models. While several factors can affect the accuracy of genome reconstruction, it is common to capture greater than 90% of the genomic information for abundant community members [11].

45.3 Prediction of Interactions Using Metabolic Networks

45.3.1 Interspecies Interaction Scoring

Metabolic networks can be used to identify and evaluate the type (i.e., competition or cooperation) and extent of metabolic interactions in a community. Various interaction indices have been proposed in the literature. Basic information used for interaction scoring includes (i) the minimal set of essential nutrients required for the growth of individual species (in short minimal nutritional requirement (MNR)) and (ii) the set of metabolites potentially produced in each species (potential metabolic production (PMP)). Competition scores such as effective metabolic overlap and metabolic competition index examine the overlapping of MNR between two organisms [12, 13]. In contrast, cooperation scores, such as Biosynthetic Support Score and Metabolic Complementarity Index, compare MNR (in one species) and PMP (in another species) [14, 15]. MNR represents an interface between metabolism and environment. The analysis of interaction scores is useful in various applications (including designing minimal media, isolating individual organisms from communities, and identifying drug targets for killing pathogens). In a broader sense, the usefulness of these analyses has been described as *reverse ecology* [16].

Zelezniak *et al*. [17] extended the community interaction space to more than two species. For a community of n species, they proposed metabolic resource overlap and metabolic interaction potential as indices for competition and cooperation, respectively. They also proposed the species metabolic interaction analysis (SMETANA) score, which is a more advanced index for assessing the extent

of cooperation for a community. The SMETANA score sums up all interspecies dependencies in the community under a given environment.

Alternatively, metabolic interactions can be assessed by accounting for the impact the presence of one organism has on the growth of another, as determined through flux modeling. In this manner, we can evaluate the positive or negative effect of metabolic interactions by examining the growth rates of organisms when they are cultured in isolation versus together. Freilich *et al*. [18] analyzed all possible binary combinations among 118 metabolic network models. For each of the 6,903 model pairs, they compared the sum of individual growth (SIG) and the co-growth (CG) to evaluate the type and extent of their interactions: competitive (if $SIG > CG$), cooperative (if $SIG < CG$), and neutral (if $SIG \approx CG$).

45.3.2 Steady-State Flux Modeling

Metabolic network simulations can provide details on metabolite exchanges between species and their impact upon community properties. Flux balance analysis (FBA) solves linear programming to estimate intracellular flux distributions, based on the assumption that an organism optimizes metabolism such that a certain metabolic objective is maximized (or minimized) [19]. In the extension of FBA to communities (cFBA) [20], individual metabolic networks can be treated as internal compartments that interact via exchange of metabolites. A typical choice of metabolic objective to maximize in cFBA is the total biomass.

In the first report on cFBA, Stolyar *et al*. [21] analyzed syntrophic interactions between sulfate-reducing bacteria and methanogens to predict the exchange of metabolites, including electron transfer agents such as formate and hydrogen. Klitgord and Segre [22] extended this method to identify a minimal set of metabolites that are potentially exchanged between organisms. Wintermute and Silver [23] developed a computational framework based on the minimization of metabolic adjustment [24], which enables systematic investigation of interactions in synthetic pairs of organisms. Zomorrodi and Maranas [25] proposed an OptCom framework that formulates cFBA as a bi-level optimization problem that sets up objectives at both community and individual species levels. Metabolic interactions are predicted as trade-offs between species- and community-level objectives.

Metabolic pathway analysis such as elementary mode analysis (EMA) can serve as a complementary or alternative approach for cFBA. Community EMA (cEMA) identifies alternative metabolic routes from nutrients to products in each organism and comprehensively explores potential interspecies interactions through metabolite exchanges. Taffs *et al*. [26] applied cEMA to examine potential metabolic interactions among three functional guilds corresponding to photoautotrophic cyanobacteria, anoxygenic photoheterotrophs, and sulfate-reducing bacteria.

45.3.3 Modeling Dynamic Interactions

In varying or even static environmental conditions, microbial interactions in a community are subject to dynamic variation. Hanly and Henson [27] employed dynamic extension of FBA (dFBA) to predict the competition between *Saccharomyces cerevisiae* (consuming glucose) and *Pichia stipitis* (consuming both glucose and xylose). Similarly, Höffner and Barton [28] applied dFBA to investigate the dynamics of the symbiotic relationship between an alga (*Chlamydomonas reinhardtii*) and a yeast (*S. cerevisiae*) in a microalgae farming process. Zhuang *et al*. [29] developed the dynamic multispecies metabolic modeling (DyMMM) framework, a variant of dFBA for community modeling, to predict dynamic competition of two acetate oxidizing Fe(III) reducers in ecological settings. DyMMM has also been applied to simulate dynamic interaction between *Geobacter* and sulfate-reducing bacteria [30]. The dynamic extension of OptCom (d-OptCom) developed by Zomorrodi *et al*. [31] provides a framework for predicting interactive dynamics.

EMA has also been extended for dynamic simulation of microbial communities by combining individual models. EMA-based dynamic metabolic modeling frameworks include the hybrid cybernetic model (HCM) [32, 33], the lumped HCM (L-HCM) [34, 35], and others (e.g., macroscopic bioreaction model [36]). Geng *et al*. [37] applied HCM to predict dynamic competition among three yeast strains (*S. cerevisiae*, *P. stipitis*, and *Kluyveromyces marxianus*) for mixed sugars (glucose, xylose, mannose, and galactose). Using this model, they showed that bioethanol productivity could increase up to 58% by optimal use of multiple species in comparison with monoculture fermentation.

45.4 Conclusions

A shift in engineering focus from the activities of individual species to their interactions with other species is a new paradigm that requires fresh framework. We highlighted the value of mathematical modeling approaches as an *in silico* tool for designing microbial interactions, with a particular focus on multispecies community metabolic networks in the broader context of the proposed BioCAD*i* concept. We envision that the extension of metabolic network models to complex communities will become feasible due to the rapid development of multi-omics analysis and computational tools.

Acknowledgments

This work was supported by the Genomic Science Program (GSP), Office of Biological and Environmental Research (OBER), U.S. Department of Energy (DOE), and is a contribution of the Pacific Northwest National Laboratory

(PNNL) Foundational Scientific Focus Area. The PNNL is operated for DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830. Author HCB is grateful for the support from the Linus Pauling Distinguished Postdoctoral Fellowship, a laboratory-directed research and development program at PNNL. The authors thank Dr Jim Fredrickson at PNNL for inspiring the discussion that led to this book chapter.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Medema, M.H., van Raaphorst, R., Takano, E., and Breitling, R. (2012) Computational tools for the synthetic design of biochemical pathways. *Nat. Rev. Microbiol.*, **10**, 191–202.
2. Lindemann, S.R., Bernstein, H.C., Song, H.-S., Fredrickson, J.K. *et al.* (2016) Engineering microbial consortia for controllable outputs. *ISME J.*, **10**, 2077–2084.
3. Faust, K. and Raes, J. (2012) Microbial interactions: from networks to models. *Nat. Rev. Microbiol.*, **10**, 538–550.
4. Larsen, P.E., Field, D., and Gilbert, J.A. (2012) Predicting bacterial community assemblages using an artificial neural network approach. *Nat. Methods*, **9**, 621–625.
5. Song, H.-S., Cannon, W.R., Beliaev, A.S., and Konopka, A. (2014) Mathematical modeling of microbial community dynamics: a methodological review. *Processes*, **2**, 711–752.
6. Henry, C.S., DeJongh, M., Best, A.A., Frybarger, P.M. *et al.* (2010) High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.*, **28**, 977–982.
7. Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M. *et al.* (2008) The RAST server: rapid annotations using subsystems technology. *BMC Genomics*, **9**, Article number 75.
8. Faria, J.P., Khazaei, T., Edirisinghe, J., Weisenhorn, P.B. *et al.* (2016) Constructing and analyzing metabolic flux models of microbial communities, in *Hydrocarbon and Lipid Microbiology Protocols* (eds T.J. McGenity, K.N. Timmis, and B. Nogales), Springer-Verlag, Berlin Heidelberg, In press.
9. Thiele, I. and Palsson, B.O. (2010) A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.*, **5**, 93–121.
10. Henry, C.S., Bernstein, H.C., Weisenhorn, P.B., Taylor, R.C. *et al.* (2016) Microbial community metabolic modeling: a community data-driven network reconstruction. *J. Cell. Physiol.*, **231**, 2339–2345.
11. Nelson, W.C., Maezato, Y., Wu, Y.W., Romine, M.F. *et al.* (2016) Identification and resolution of microdiversity through metagenomic sequencing of parallel consortia. *Appl. Environ. Microbiol.*, **82**, 255–267.
12. Kreimer, A., Doron-Faigenboim, A., Borenstein, E., and Freilich, S. (2012) NetCmpt: a network-based tool for calculating the metabolic competition between bacterial species. *Bioinformatics*, **28**, 2195–2197.
13. Levy, R. and Borenstein, E. (2013) Metabolic modeling of species interaction in the human microbiome elucidates community-level assembly rules. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 12804–12809.
14. Borenstein, E. and Feldman, M.W. (2009) Topological signatures of species interactions in metabolic networks. *J. Comput. Biol.*, **16**, 191–200.
15. Levy, R., Carr, R., Kreimer, A., Freilich, S. *et al.* (2015) NetCooperate: a network-based tool for inferring

host-microbe and microbe-microbe cooperation. *BMC Bioinformatics*, **16**, 164.

16. Levy, R. and Borenstein, E. (2012) Reverse ecology: from systems to environments and back. *Adv. Exp. Med. Biol.*, **751**, 329–345.
17. Zelezniak, A., Andrejev, S., Ponomarova, O., Mende, D.R. *et al.* (2015) Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 6449–6454.
18. Freilich, S., Kreimer, A., Meilijson, I., Gophna, U. *et al.* (2010) The large-scale organization of the bacterial network of ecological co-occurrence interactions. *Nucleic Acids Res.*, **38**, 3857–3868.
19. Feist, A.M. and Palsson, B.O. (2010) The biomass objective function. *Curr. Opin. Microbiol.*, **13**, 344–349.
20. Khandelwal, R.A., Olivier, B.G., Roling, W.F.M., Teusink, B. *et al.* (2013) Community flux balance analysis for microbial consortia at balanced growth. *Plos One*, **8**, e64567.
21. Stolyar, S., Van Dien, S., Hillesland, K.L., Pinel, N. *et al.* (2007) Metabolic modeling of a mutualistic microbial community. *Mol. Syst. Biol.*, **3**, Article number 92.
22. Klitgord, N. and Segre, D. (2010) Environments that induce synthetic microbial ecosystems. *PLoS Comput. Biol.*, **6**, e1001002.
23. Wintermute, E.H. and Silver, P.A. (2010) Emergent cooperation in microbial metabolism. *Mol. Syst. Biol.*, **6**, 407.
24. Segre, D., Vitkup, D., and Church, G.M. (2002) Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 15112–15117.
25. Zomorrodi, A.R. and Maranas, C.D. (2012) OptCom: a multi-level optimization framework for the metabolic modeling and analysis of microbial communities. *PLoS Comput. Biol.*, **8**, e1002363.
26. Taffs, R., Aston, J.E., Brileya, K., Jay, Z. *et al.* (2009) In silico approaches to study mass and energy flows in microbial consortia: a syntrophic case study. *BMC Syst. Biol.*, **3**, 114.
27. Hanly, T.J. and Henson, M.A. (2013) Dynamic metabolic modeling of a microaerobic yeast co-culture: predicting and optimizing ethanol production from glucose/xylose mixtures. *Biotechnol. Biofuels*, **6**, 44.
28. Höffner K, Barton PI. 2014. *Design of Microbial Consortia for Industrial Biotechnology*. Proceedings of the 8th International Conference on Foundations of Computer-Aided Process Design, Elsevier. Cle Elum, Washington, USA, July 13–17, pp. 65–74.
29. Zhuang, K., Izallalen, M., Mouser, P., Richter, H. *et al.* (2011) Genome-scale dynamic modeling of the competition between Rhodoferax and Geobacter in anoxic subsurface environments. *ISME J.*, **5**, 305–316.
30. Zhuang, K., Ma, E., Lovley, D.R., and Mahadevan, R. (2012) The design of long-term effective uranium bioremediation strategy using a community metabolic model. *Biotechnol. Bioeng.*, **109**, 2475–2483.
31. Zomorrodi, A.R., Islam, M.M., and Maranas, C.D. (2014) d-OptCom: dynamic multi-level and multi-objective metabolic modeling of microbial communities. *ACS Synth. Biol.*, **3**, 247–257.
32. Kim, J.I., Varner, J.D., and Ramkrishna, D. (2008) A hybrid model of anaerobic *E. coli* GJT001: combination of elementary flux modes and cybernetic variables. *Biotechnol. Progr.*, **24**, 993–1006.
33. Song, H.S., Morgan, J.A., and Ramkrishna, D. (2009) Systematic development of hybrid cybernetic models: application to recombinant yeast co-consuming glucose and xylose. *Biotechnol. Bioeng.*, **103**, 984–1002.
34. Song, H.S. and Ramkrishna, D. (2010) Prediction of metabolic function from limited data: lumped hybrid cybernetic modeling (L-HCM). *Biotechnol. Bioeng.*, **106**, 271–284.
35. Song, H.S. and Ramkrishna, D. (2011) Cybernetic models based on lumped elementary modes accurately predict strain-specific metabolic function. *Biotechnol. Bioeng.*, **108**, 127–140.

36. Provost, A., Bastin, G., Agathos, S.N., and Schneider, Y.J. (2006) Metabolic design of macroscopic bioreaction models: application to Chinese hamster ovary cells. *Bioproc. Biosyst. Eng.*, **29**, 349–366.

37. Geng, J., Song, H.S., Yuan, J.Q., and Ramkrishna, D. (2012) On enhancing productivity of bioethanol with multiple species. *Biotechnol. Bioeng.*, **109**, 1508–1517.