



Predicting Microbiome Metabolism and Interactions through Integrating Multidisciplinary Principles

Caleb M. Schmidt,^{a,b} Parsa Ghadermazi,^c Siu Hung Joshua Chan^c

^aAdvanced Pattern Analysis & Countermeasures Group, Boulder, Colorado, USA

^bDepartment of Systems Engineering, Colorado State University, Fort Collins, Colorado, USA

^cDepartment of Chemical and Biological Engineering, Colorado State University, Fort Collins, Colorado, USA

ABSTRACT In this Commentary, we will discuss some of the current trends and challenges in modeling microbiome metabolism. A focus will be the state of the art in the integration of metabolic networks, ecological and evolutionary principles, and spatiotemporal considerations, followed by envisioning integrated frameworks incorporating different principles and data to generate predictive models in the future.

KEYWORDS metabolic modeling, microbial communities, microbiome

The past 2 decades have seen significant increases in the technical ability of metabolic models to more accurately represent the underlying metabolic dynamics in biological systems. What started as models of simple organisms, representing only a subset of their genes (1), has now evolved into multicompartmental human models (2). This positive trend in experimental and modeling capability has also allowed for the construction of multiorganism, microbiome systems with significant nuance. This Commentary describes the state of the art in metabolic modeling (especially as it relates to microbiomes) and entertains a selection of considerations that, we believe, are highly relevant to continued progress in the field (Fig. 1).

GENOME-SCALE METABOLIC NETWORK


A hallmark approach in metabolic modeling is genome-scale metabolic modeling (GEM). The goal of a GEM is to depict the metabolic structure of a cell using the genome sequence as a blueprint, allowing for a complete mapping of reactions in a cell. On top of this blueprint is layered the known metabolic reactions, gene data, and biochemical data to generate a curated reaction network (Fig. 1A). Within this framework, the cell system is framed as a stoichiometric matrix (Fig. 1B). By applying steady-state mass balance across the system and constraining the reaction directionality based on thermodynamics, a solvable system of linear equations can be optimized for a given cellular objective function, e.g., biomass production, to obtain a set of flux distributions across all reactions in the cell. This forms the basis of flux balance analysis (FBA) (3).

The GEM approach described above has shown significant success in human systems (e.g., references 4 and 5). A sex-specific, organ-resolved GEM that incorporates ~80,000 metabolic reactions with relevant anatomical and physiological data was recently published (2). For the microbial counterpart, GEMs have provided a mechanistic framework connecting the chemical environment to microbial metabolism. GEMs can accurately predict metabolic phenotypes and engineering strategies for rewiring microbial metabolism for various biomedical and biotechnological applications, e.g., biochemical production. One important recent advance is the development of automated reconstruction tools (6), such as KBase and CarveMe, that allow for massive rapid model reconstruction in large-scale studies. For example, AGORA2 generates draft models for >7,000 strains using KBase (7). It is currently the largest gut microbial

Citation Schmidt CM, Ghadermazi P, Chan SHJ. 2021. Predicting microbiome metabolism and interactions through integrating multidisciplinary principles. *mSystems* 6: e00768-21. <https://doi.org/10.1128/mSystems.00768-21>.

Copyright © 2021 Schmidt et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Siu Hung Joshua Chan, Joshua.Chan@colostate.edu.

 Commentary on integrating genome-scale metabolic modeling, population and evolutionary dynamics, spatial structures, and omics data for predicting microbiome metabolism

Conflict of Interest Disclosures: C.M.S. has nothing to disclose. P.G. has nothing to disclose. S.H.J.C. has nothing to disclose.

The views expressed in this article do not necessarily reflect the views of the journal or of ASM.

This article is part of a special series sponsored by Floré.

Published 5 October 2021

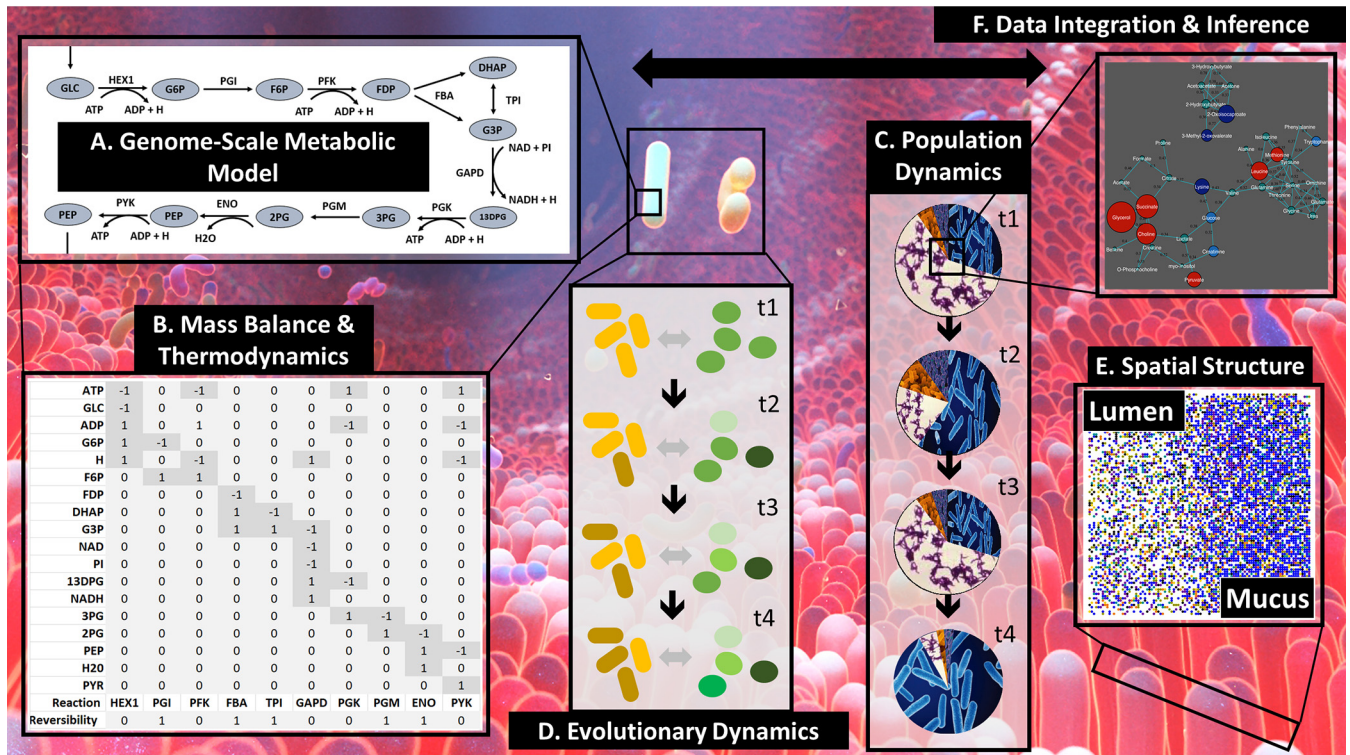


FIG 1 Integrating genome-scale metabolic models (A and B), population (C) and evolutionary (D) dynamics, spatial structures (E) (image adapted from *PLoS Computational Biology* [19]), and omics data (F) for predicting microbiome metabolism.

model collection and can be used to simulate drug metabolism affected by human gut microbes. The history and state of the art are reviewed elsewhere in greater detail (8). When extending to modeling microbiomes, however, there are still significant limitations and challenges, e.g., insufficient taxonomic resolution and biochemical knowledge for nonmodel organisms, computational scalability from single-organism to community models, and more general evolutionary principles governing communities needed. These are all related, and here we will look into this from the perspective of integrating ecological and evolutionary principles.

ECOLOGICAL AND EVOLUTIONARY PRINCIPLES

A decade of constructing microbial community GEMs has made it clear that different model structures and algorithms have different assumptions and answer different questions. In this context, there are three relevant questions that must be addressed. The first is whether to use a computationally efficient supraorganism model (microbiome as a single organism) or a dynamically more accurate community model (microbiome as multiple-assembled organism genomes). A second important consideration is whether static versus dynamic simulations will be used (discussed below) (9). A third relevant question is whether to use a single-level or multilevel objective optimization approach. Each of these considerations involves different ecological and evolutionary principles and assumptions (Fig. 1C and D). We will take a deeper look at the second and third questions.

Static, single-level frameworks. Single-level approaches optimize across the entire community (e.g., by assuming group selection and optimizing the total fitness of the community, instead of individuals). Evolutionary biologists could be critical of this approach because it bypasses selection at the individual level. Indeed, FBA omits certain stable interactions within communities and can be prone to unrealistic altruistic predictions (10). Community FBA was proposed to implement an equal growth rate constraint across the community. This framework allows for a population steady-state

assumption to be followed in stable communities (11). We have proposed a reformulated algorithm, called SteadyCom, that is able to perform flux variability analysis (FVA) and demonstrated its scalability to large community models. Through this approach, we were able to predict the gut microbiota relative abundance profile given an average American diet (10).

Static, multilevel frameworks. In contrast, multilevel approaches try to find solutions in which each individual member of a community maximizes its fitness function. This approach is closer to the accepted theory of natural selection/adaptation but elevates the computational demand. OptCom was the first bilevel optimization paradigm to consider both individual and community fitness (12). Building on this concept, we have recently proposed a new bilevel algorithm, called NECom, to formally implement the concept of Nash equilibrium (NE) in microbial metabolic models (13). NE is a concept from evolutionary game theory in which no individual microbe can increase its fitness unilaterally and thus each microbe is maximizing its fitness function simultaneously. We showed the superior predictive capability of NECom in a large-scale data set of mutualistic alga-yeast cocultures compared to FBA, with a 70% improvement in prediction of member growth rates under ~500 uptake conditions (13).

Evolutionary dynamic frameworks. Dynamic flux balance analysis (dFBA) juxtaposes the steady-state assumptions of FBA (14). dFBA models the time-dependent extracellular metabolite concentrations by solving FBA-embedded differential equations. In each time step, uptake constraints in FBA are updated to solve for new exchange fluxes. However, we recently showed that dFBA cannot predict the coevolution of community members, e.g., the complementation of amino acid auxotrophic *Escherichia coli* pairs, because each can act only upon metabolites present in their environments (13). That is, a microbe is unable to “know” what other microbes in the system are capable of and act accordingly. Therefore, coevolved, stable, mutualistic interactions favored by natural selection may be missing.

This dFBA deficiency is partially addressed by a recent algorithm based also on NE by Zomorodi and Segrè (15) in which mutants from the same species with alternative metabolic strategies are predefined and simulated until an NE is reached. By solving only linear programs, the algorithm is computationally efficient but requires predefined strategies. In contrast, the NECom algorithm predicts NE using network structures alone but is computationally demanding (solving bilevel programs). An evolutionary dynamic framework with evolving strategies for individual microbes similar to that used in reference 16 could have the advantages of both.

Spatiotemporal dynamics. dFBA has achieved success in ground truthing models to experimental data and allowed for introducing spatiotemporal dynamics to microbial ecosystems. This is possible through the addition of differential equations for spatial dynamics and/or compartmentalization based on spatial structures to produce a realistic picture. We were able to capture the observed differential relative abundance profiles of microbiota radially, from the intestinal lumen to the mucosal layer, and longitudinally, along the intestinal tract, by proposing a dynamic model of the mucosal cell layer shedding into the lumen directing flow into the succeeding intestinal section and combining SteadyCom and dFBA (17). Explicit spatial dynamics further improves the resolution. COMETS incorporates concentration-based diffusion dynamics to model the global spatiotemporal dynamics. It is able to predict nonintuitive competition-mutualism dynamics that match experimental results (18). BacArena (19) (Fig. 1E) and IndiMeSH (20) are also important examples of platforms that allow investigators to spatially resolve microbial populations and molecular signatures in different matrices.

DATA-DRIVEN INTERACTION INFERENCE AND HYPOTHESIS TESTING

Algorithms have been developed to integrate omics data into the metabolic network structure to identify interactions and test mechanistic hypotheses (Fig. 1F). In our previous work, we have tested the validity of mechanistic hypotheses, e.g., bile salt hydrolase activity in the *Clostridia* causing observed gut microbial metabolomic change after drug treatment (21), and tested which microbes are likely to be responsible for

the observed fecal metabolomic changes in infants during initial microbial colonization (22). More sophisticated techniques, such as machine learning (ML) algorithms, have helped in generating, maintaining, and integrating experimental data into models to help infer the phenome and other information about the system. For example, ML has been successful in supporting investigators to better annotate genomes, fill gaps in GEMs, choose constraints for models, and integrate omics data into GEMs (23). Increases in the scale and complexity of data only support the continuation of this trend to better understand and integrate data.

THE FUTURE—INTEGRATION OF PRINCIPLES AND HETEROGENEOUS DATA SET

We have seen how GEMs can capture biological observations with very few parameters and meanwhile how machine learning can improve the accuracy of predictions based on observed data. We believe that seamless integration of the aforementioned (ecological, evolutionary, and spatiotemporal) principles in a systematic modeling framework with parameters learned through rigorous model inference techniques will connect theory and data to enhance the usefulness of GEMs for studying complex microbiomes. This process of integration includes gathering heterogeneous data sets where myriad supervised and unsupervised methods can be used to concatenate, transform, or model to create algorithm-based models (24). For example, one can train uptake kinetic parameters and resource allocation constraints in a community dFBA model that is evolutionarily stable and cross-validated by concentration, relative abundance data, and other omics data. Further, these models can be used as an analytical framework as well as to suggest microbial/genetic/nutritional/spatial perturbations that alter the evolutionary fate of the community.

Within the set of experimental techniques to generate these data sets, there are two flavors: bottom-up versus top-down. As an alternative to the more commonly used bottom-up synthetic community approach, top-down approaches to study smaller communities of a complex microbiome are emerging. A recent study deconstructed the soil microbiome into functional modules through targeted enrichment of desired metabolic traits. It allows for investigating a community with significantly less complexity and drawing mechanistic conclusions (25). Another promising approach is to culture the “unculturable” organism in a microbiome using microfluidic droplets as localized, nanoliter-sized bioreactors. By encapsulating organisms from human fecal samples this way, metagenome-assembled genomes of uncharacterized gut commensals with high resolution can be recovered (26). New modeling tools should be ready to integrate these data to gain new insights into the microbiome.

REFERENCES

1. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269:496–512. <https://doi.org/10.1126/science.7542800>.
2. Thiele I, Sahoo S, Heinken A, Hertel J, Heirendt L, Aurich MK, Fleming RM. 2020. Personalized whole-body models integrate metabolism, physiology, and the gut microbiome. *Mol Syst Biol* 16:e8982. <https://doi.org/10.1525/msb.20198982>.
3. Orth JD, Thiele I, Palsson BØ. 2010. What is flux balance analysis? *Nat Biotechnol* 28:245–248. <https://doi.org/10.1038/nbt.1614>.
4. Bordbar A, Lewis NE, Schellenberger J, Palsson BØ, Jamshidi N. 2010. Insight into human alveolar macrophage and *M. tuberculosis* interactions via metabolic reconstructions. *Mol Syst Biol* 6:422. <https://doi.org/10.1038/msb.2010.68>.
5. Rawls KD, Dougherty BV, Vinnakota KC, Pannala VR, Wallqvist A, Kolling GL, Papin JA. 2021. Predicting changes in renal metabolism after compound exposure with a genome-scale metabolic model. *Toxicol Appl Pharmacol* 412:115390. <https://doi.org/10.1016/j.taap.2020.115390>.
6. Mendoza SN, Olivier BG, Molenaar D, Teusink B. 2019. A systematic assessment of current genome-scale metabolic reconstruction tools. *Genome Biol* 20:158. <https://doi.org/10.1186/s13059-019-1769-1>.
7. Heinken A, Acharya G, Ravcheev DA, Hertel J, Nyga M, Okpala OE, Hogan M, Magnúsdóttir S, Martinelli F, Preciat G, Edirisinghe JN, Henry CS, Fleming RMT, Thiele I. 2020. AGORA2: large scale reconstruction of the microbiome highlights wide-spread drug-metabolising capacities. *bioRxiv* <https://www.biorxiv.org/content/10.1101/2020.11.09.375451v1>.
8. Fang X, Lloyd CJ, Palsso BO. 2020. Reconstructing organisms in silico: genome-scale models and their emerging applications. *Nat Rev Microbiol* 18:731–743. <https://doi.org/10.1038/s41579-020-00440-4>.
9. Mahadevan R, Edwards JS, Doyle FJ, III. 2002. Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys J* 83:1331–1340. [https://doi.org/10.1016/S0006-3495\(02\)73903-9](https://doi.org/10.1016/S0006-3495(02)73903-9).
10. Chan SHJ, Simons MN, Maranas CD. 2017. SteadyCom: predicting microbial abundances while ensuring community stability. *PLoS Comput Biol* 13:e1005539. <https://doi.org/10.1371/journal.pcbi.1005539>.
11. Khandelwal RA, Olivier BG, Röling WF, Teusink B, Bruggeman FJ. 2013. Community flux balance analysis for microbial consortia at balanced growth. *PLoS One* 8:e64567. <https://doi.org/10.1371/journal.pone.0064567>.
12. Zomorodi AR, Maranas CD. 2012. OptCom: a multi-level optimization framework for the metabolic modeling and analysis of microbial communities. *PLoS Comput Biol* 8:e1002363. <https://doi.org/10.1371/journal.pcbi.1002363>.
13. Cai J, Tan T, Chan SHJ. 2020. Predicting Nash equilibria for microbial metabolic interactions. *Bioinformatics* <https://doi.org/10.1093/bioinformatics/btaa1014>.

14. Zhuang K, Izallalen M, Mouser P, Richter H, Risso C, Mahadevan R, Lovley DR. 2011. Genome-scale dynamic modeling of the competition between *Rhodospirillum rubrum* and *Geobacter* in anoxic subsurface environments. *ISME J* 5: 305–316. <https://doi.org/10.1038/ismej.2010.117>.
15. Zomorodi AR, Segrè D. 2017. Genome-driven evolutionary game theory helps understand the rise of metabolic interdependencies in microbial communities. *Nat Commun* 8:1563. <https://doi.org/10.1038/s41467-017-01407-5>.
16. van Hoek MJA, Merks RMH. 2017. Emergence of microbial diversity due to cross-feeding interactions in a spatial model of gut microbial metabolism. *BMC Syst Biol* 11:56. <https://doi.org/10.1186/s12918-017-0430-4>.
17. Chan SHJ, Friedman ES, Wu GD, Maranas CD. 2019. Predicting the longitudinally and radially varying gut microbiota composition using multi-scale microbial metabolic modeling. *Processes* 7:394. <https://doi.org/10.3390/pr7070394>.
18. Lee C, Jung YJ, Lee SJ, Im CH. 2017. COMETS2: an advanced MATLAB toolbox for the numerical analysis of electric fields generated by transcranial direct current stimulation. *J Neurosci Methods* 277:56–62. <https://doi.org/10.1016/j.jneumeth.2016.12.008>.
19. Bauer E, Zimmermann J, Baldini F, Thiele I, Kaleta C. 2017. BacArena: individual-based metabolic modeling of heterogeneous microbes in complex communities. *PLoS Comput Biol* 13:e1005544. <https://doi.org/10.1371/journal.pcbi.1005544>.
20. Borer B, Ataman M, Hatzimanikatis V, Or D. 2019. Modeling metabolic networks of individual bacterial agents in heterogeneous and dynamic soil habitats (IndiMeSH). *PLoS Comput Biol* 15:e1007127. <https://doi.org/10.1371/journal.pcbi.1007127>.
21. Zhang L, Xie C, Nichols RG, Chan SHJ, Jiang C, Hao R, Smith PB, Cai J, Simons MN, Hatzakis E, Maranas CD, Gonzalez FJ, Patterson AD. 2016. Farnesoid X receptor signaling shapes the gut microbiota and controls hepatic lipid metabolism. *mSystems* 1:e00070-16. <https://doi.org/10.1128/mSystems.00070-16>.
22. Bittinger K, Zhao C, Li Y, Ford E, Friedman ES, Ni J, Kulkarni CV, Cai J, Tian Y, Liu Q, Patterson AD, Sarkar D, Chan SHJ, Maranas C, Saha-Shah A, Lund P, Garcia BA, Mattei LM, Gerber JS, Elovitz MA, Kelly A, DeRusso P, Kim D, Hofstaedter CE, Goulian M, Li H, Bushman FD, Zemel BS, Wu GD. 2020. Bacterial colonization reprograms the neonatal gut metabolome. *Nat Microbiol* 5:838–847. <https://doi.org/10.1038/s41564-020-0694-0>.
23. Rana P, Berry C, Ghosh P, Fong SS. 2020. Recent advances on constraint-based models by integrating machine learning. *Curr Opin Biotechnol* 64: 85–91. <https://doi.org/10.1016/j.copbio.2019.11.007>.
24. Zampieri G, Vijayakumar S, Yaneske E, Angione C. 2019. Machine and deep learning meet genome-scale metabolic modeling. *PLoS Comput Biol* 15:e1007084. <https://doi.org/10.1371/journal.pcbi.1007084>.
25. Naylor D, Fansler S, Brislawn C, Nelson WC, Hofmockel KS, Jansson JK, McClure R. 2020. Deconstructing the soil microbiome into reduced-complexity functional modules. *mBio* 11:e01349-20. <https://doi.org/10.1128/mBio.01349-20>.
26. Tan JY, Wang S, Dick GJ, Young VB, Sherman DH, Burns MA, Lin XN. 2020. Co-cultivation of microbial sub-communities in microfluidic droplets facilitates high-resolution genomic dissection of microbial ‘dark matter’. *Integr Biol (Camb)* 12:263–274. <https://doi.org/10.1093/intbio/zyaa021>.