

# Metabolic network modeling of microbial communities

Matthew B. Biggs,<sup>1</sup> Gregory L. Medlock,<sup>1</sup> Glynis L. Kolling<sup>2</sup> and Jason A. Papin<sup>1\*</sup>

Genome-scale metabolic network reconstructions and constraint-based analyses are powerful methods that have the potential to make functional predictions about microbial communities. Genome-scale metabolic networks are used to characterize the metabolic functions of microbial communities via several techniques including species compartmentalization, separating species-level and community-level objectives, dynamic analysis, the ‘enzyme-soup’ approach, multiscale modeling, and others. There are many challenges in the field, including a need for tools that accurately assign high-level omics signals to individual community members, the need for improved automated network reconstruction methods, and novel algorithms for integrating omics data and engineering communities. As technologies and modeling frameworks improve, we expect that there will be corresponding advances in the fields of ecology, health science, and microbial community engineering. © 2015 Wiley Periodicals, Inc.

## How to cite this article:

*WIREs Syst Biol Med* 2015, 7:317–334. doi: 10.1002/wsbm.1308

## INTRODUCTION

Microbial communities represent a gargantuan force of nature that exerts influence on global geochemical cycles,<sup>1</sup> agriculture,<sup>2</sup> human health,<sup>3</sup> food preparation,<sup>4</sup> and a host of relevant aspects of life on earth.<sup>5,6</sup> Traditional microbiology has made great strides over the last century in describing and categorizing these microscopic neighbors. More recently, advances in sequencing technologies have provided the first glimpses at the composition of natural microbial communities, including insights into the physiology of non-culturable microbes.<sup>7</sup> Databases are filling with mountains of genomic fragments, gene and protein expression data, and other such large-scale ‘-omics’ information, all describing the content of diverse microbial communities.<sup>8,9</sup> Despite the plethora of data, we yet lack true understanding of the mechanisms that cause communities to

function and interact with their environments.<sup>10</sup> Considering the importance of microbial communities to many global ecosystems, health, and various industries, there is a great need to move beyond a descriptive ‘parts list’ approach of the field, and transition to more functional, predictive models of microbial community structure and function.

Predictive community models have the potential to engender many beneficial technologies including: rational probiotic design for restoring a diseased intestinal microbiota,<sup>11</sup> efficient chemical-producing consortia,<sup>12</sup> or optimal bioremediation communities.<sup>13</sup> Furthermore, predictive models will allow novel exploration of basic questions in microbial ecology,<sup>14,15</sup> leading to new insights into the development and evolution of microbial communities<sup>10</sup> (Figure 1). All of these potential applications will require improvements in the mathematical toolbox used to represent biochemical networks and their interactions.

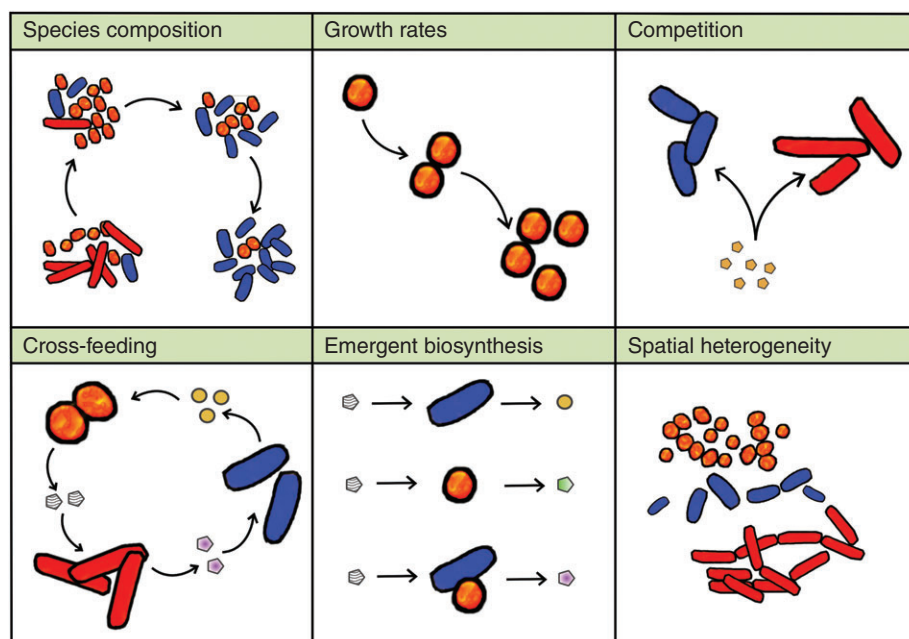
Genome-scale metabolic network reconstructions (GENREs) have been successfully applied to the representation, study, and engineering of single microbes<sup>16</sup> (Figure 2). The last decade has seen extensive tool development for the analysis of models encompassing single strains up to complex microbial

\*Correspondence to: papin@virginia.edu

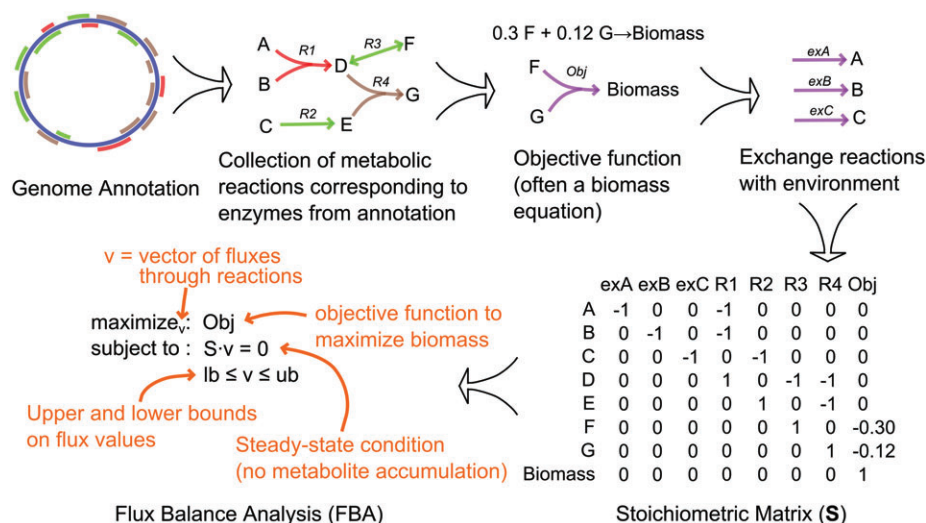
<sup>1</sup>Department of Biomedical Engineering, University of Virginia, Charlottesville, VA, USA

<sup>2</sup>Department of Medicine, Infectious Diseases, University of Virginia, Charlottesville, VA, USA

Conflict of interest: The authors have declared no conflicts of interest for this article.



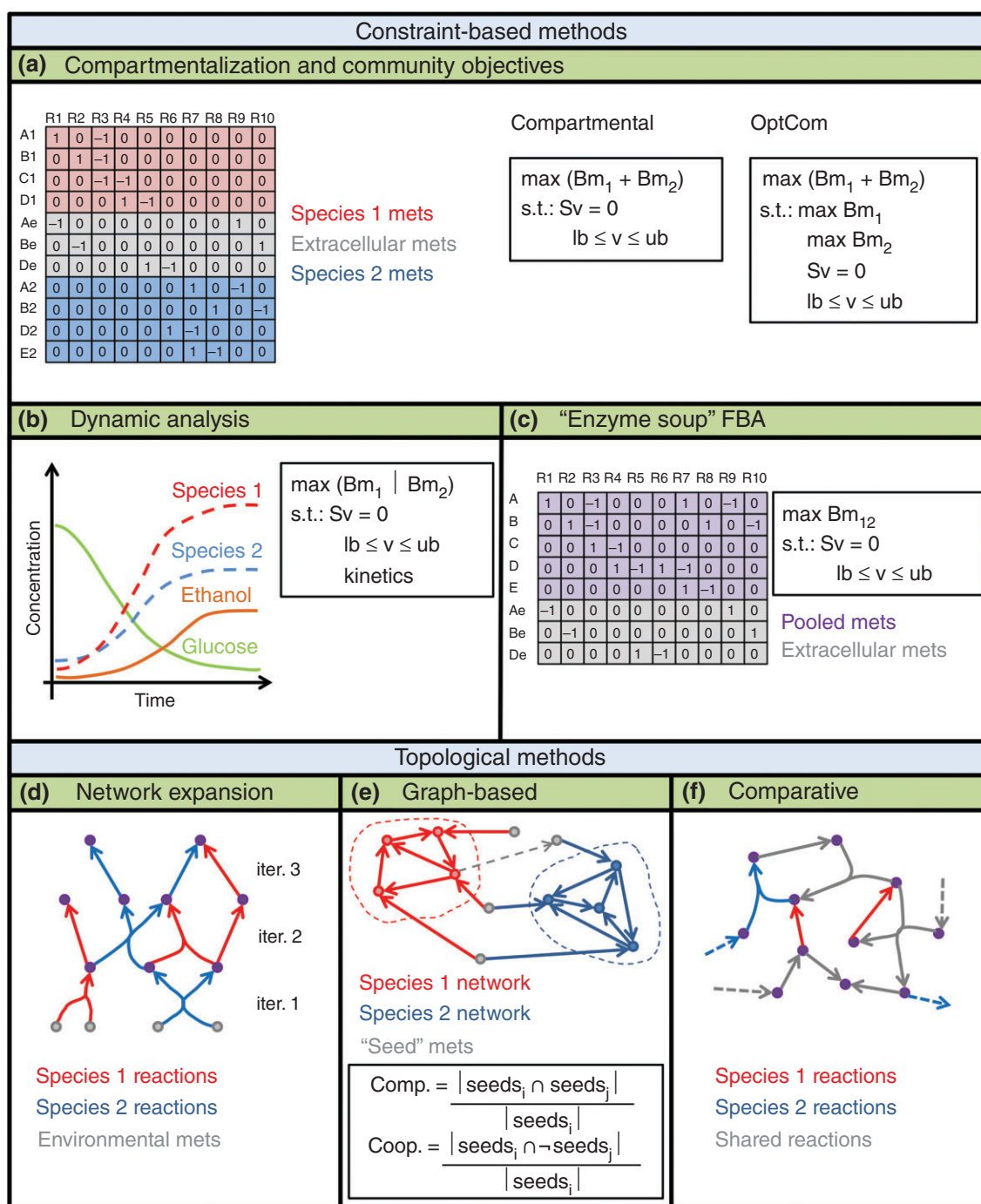
**FIGURE 1** | There are many aspects of life in a microbial community that would be useful to capture using mathematical models. Techniques utilizing constraint-based metabolic models (sometimes in conjunction with other modeling approaches) are capable of capturing all of these scenarios.



**FIGURE 2** | A simple workflow for genome-scale metabolic network reconstruction and accompanying constraint-based analysis. The process begins with an annotated genome. The metabolic network is derived from this genome annotation by searching databases for homologous proteins with known enzymatic activity. The corresponding metabolic reactions are collected into a draft network reconstruction. This simple procedure can be augmented through gap filling, and often manual curation. A metabolic objective is defined, which for microbes is often assumed to be a biomass equation (i.e., it is assumed that cells are configured to grow as fast as possible). Exchange reactions are defined to allow metabolites to enter and leave the network. All reactions are compiled into a stoichiometric (S) matrix. FBA is a common analytical approach that searches for a flux distribution through the network that optimizes the metabolic objective subject to steady-state constraints and flux bounds.

communities.<sup>10,17–22</sup> Since the first published community model in 2007 of a mutualistic microbial community, the accumulating body of work has highlighted many unique challenges related to microbial community modeling.<sup>23</sup> In this review, we discuss

the existing frameworks that have been developed using GENREs for community analysis (Figure 3 and Table 1), the types of questions that can be addressed, and challenges in the field that present opportunities for progress.



**FIGURE 3** | Community modeling frameworks that feature GENREs. (A) The compartmentalization approach unites all species-level GENREs into a unified stoichiometric matrix with a shared compartment. The objective function is generally assumed to be a linear combination of the individual biomass functions from each species. The community objectives approach (OptCom) is an extension of the simple compartmentalization approach that utilizes a nested, bi-level optimization framework. The bi-level optimization enables the representation of more classes of interactions between species, but comes at the cost of increased computational complexity. (B) Dynamic analysis simulates changes in metabolites and biomass over time, which requires constraints on uptake reaction kinetics. (C) 'Enzyme soup' FBA ignores species boundaries and assumes that all reactions can interact in a community-level meta-GENRE. Other methods include: (D) network expansion, which has been used to identify potential emergent biosynthetic capacity between species by comparing species-specific 'reachable' metabolites to the result of pooling reactions from both species; (E) graph-based methods, which can be used to quantify general characteristics of an interaction between species, such as the level of expected competition or cooperation; (F) comparative analyses, which are used to assess the differences in gene essentiality, biosynthetic capacity, and resource utilization between species. Note that mets signifies metabolites, Bm stands for biomass, s.t. means subject to, ub and lb signify upper and lower bounds, respectively.

**TABLE 1** | A Timeline for Computational Metabolic Systems Biology of Microbial Communities

A Timeline for Computational Metabolic Systems Biology of Microbial Communities		
C	Stolyar et al., 2007 <sup>23</sup>	Recapitulate mutualistic interaction between <i>Desulfovibrio vulgaris</i> and <i>Methanococcus maripaludis</i>
OM	Christian et al., 2007 <sup>24</sup>	Emergent biosynthetic capacity for 99,681 species pairs determined by network expansion
C,ES	Taffs et al., 2009 <sup>25</sup>	Interaction of three microbial guilds. Interrogated using three modeling approaches, including simple compartmentalization, 'enzyme soup', and an approach based on elementary mode analysis (EMA)
C	Bordbar et al., 2010 <sup>26</sup>	<i>Mycobacterium tuberculosis</i> embedded in alveolar macrophage metabolic reconstruction
C	Klitgord and Segrè(2010) <sup>27,28</sup>	Computationally design media to induce commensal and mutualistic interactions between several pairs of species
OM	Sun et al., 2010 <sup>29</sup>	Comparative analysis of <i>Pelobacter carbinolicus</i> and <i>Pelobacter propionicus</i>
C	Freilich et al., 2011 <sup>30</sup>	Prediction of competitive and cooperative potential among 6903 species pairs
CO	Zomorodi and Maranas (2011) <sup>31</sup>	OptCom method introduction and analysis of <i>D. vulgaris</i> and <i>M. maripaludis</i> community
DA	Hanly and Henson (2011) <sup>32</sup>	Optimization of glucose/xylose utilization by mixed cultures of <i>E. coli</i> and <i>Saccharomyces cerevisiae</i>
DA	Zhuang et al., 2011 <sup>33</sup>	Simulation of community responses to nutrient modulation. Community included <i>G. sulfurreducens</i> and <i>R. ferrireducens</i>
DA	Tzamali et al., 2011 <sup>34</sup>	Exploration of interactions between many <i>E. coli</i> gene-knockout strains
C	Heinken et al., 2012 <sup>35</sup>	Interaction of <i>Bacteroides thetaiotamicron</i> and mouse host
C	Khandelwal et al., 2013 <sup>36</sup>	Development of tools to estimate species abundances and yields, applied to co-culture of <i>Escherichia coli</i> auxotrophs
C	Nagarajan et al., 2013 <sup>37</sup>	Electron flow between <i>Geobacter metallireducens</i> and <i>Geobacter sulfurreducens</i> , integrating multi-omics data
C	Shoaie et al., 2013 <sup>38</sup>	Interactions between combinations of 2 and 3 of <i>B. thetaiotamicron</i> , <i>Eubacterium rectale</i> , and <i>Methanobrevibacter smithii</i> . Authors also apply OptCom and compare results
DA	Hanly and Henson (2013) <sup>39</sup>	Optimization of glucose/xylose utilization in community of <i>S. cerevisiae</i> and <i>Scheffersomyces stipitis</i>
OM	Levy and Borenstein (2013) <sup>40</sup>	Analysis of competition and cooperation among all pairs of 154 species using graph-based method
OM	Bartell et al., 2014 <sup>41</sup>	Comparative analysis of <i>Burkholderia cenocepacia</i> and <i>Burkholderia multivorans</i>
OM	Vinay-Lara et al., 2014 <sup>42</sup>	Comparative analysis of two strains of <i>Lactobacillus casei</i>
CO	El-Semman et al., 2014 <sup>43</sup>	Interaction of <i>Bifidobacterium adolescentis</i> and <i>Faecalibacterium prausnitzii</i> . OptCom and classic FBA are used in analysis
CO	Zomorodi et al., 2014 <sup>44</sup>	OptCom is adapted to dynamic simulations. Simulated communities of <i>E. coli</i> auxotrophs, and a uranium-reducing community involving <i>Geobacter sulfurreducens</i> , <i>Rhodoferrax ferrireducens</i> , and <i>Shewanella oneidensis</i>
DA	Chiu et al., 2014 <sup>45</sup>	Screening of 6670 two-species communities for emergent biosynthetic capacity
DA	Harcombe et al., 2014 <sup>46</sup>	Spatial element integrated with dFBA to model interaction of <i>Methylobacterium extorquens</i> , <i>E. coli</i> , and <i>Salmonella enterica</i>
C	Ye et al., 2014 <sup>47</sup>	Analysis of cross-feeding in vitamin-C-producing community composed of <i>Ketogulonicigenium vulgare</i> and <i>Bacillus megaterium</i>
ES	Tobalina et al., 2015 <sup>48</sup>	Analysis of a naphthalene-degrading community

C indicates the compartmentalization approach. CO indicates the community objectives method. DA indicates the dynamic analysis approach. ES indicates the enzyme soup approach. OM indicates other methods including graph-based approaches, network expansion, and the comparative method.



## CURRENT STATE OF THE FIELD

### Genome-Scale Metabolic Reconstructions: What They Are and What They Can Do

GENREs are an organized collection of the metabolic reactions that can occur within a biological system. This collection of reactions is inferred from genome annotations, and the resulting gene-to-protein-to-reaction mapping allows genotype–phenotype predictions (Figure 2). The heart of a GENRE is the stoichiometric (S) matrix, which consists of the stoichiometric coefficients for each reaction represented in the network reconstruction. Mass and charge are balanced for every reaction. This simple representation can be used to explore the space of possible biochemical conversions that can be carried out by the set of reactions in the GENRE. Optimization techniques such as flux balance analysis (FBA) are used to estimate optimal yields given a particular metabolic environment and GENRE.<sup>49</sup> The reconstruction and analysis of GENREs for single organisms have been reviewed extensively.<sup>16,50</sup>

The basic principles for the generation and analysis of GENREs learned from studies of single organisms have been extended in innovative ways to represent the interactions between multiple species within communities. Each of these extensions provides a unique approach to a set of field-specific questions:

1. Structure: How should species models be ‘linked’ together? Should there be an unbiased, wholesale sharing of metabolites, essentially ignoring species boundaries? Or should metabolite sharing be restricted to only those compounds for which there is empirical evidence? Should species models be linked to the environment/host? And if so, how?
2. Analysis: Should optimization be used to estimate optimal species yield or optimal community yield? Is there tension between these two objectives *in silico* or *in vitro*? What optimization strategies lead to the most useful analyses?
3. Refinement: How much curation effort should go into the individual species-level reconstructions?
4. Validation: What kind of experimental data can be used to validate model predictions?
5. Applications: Given the model structure and analysis, how informative are the model predictions? How can the model be used to answer impactful questions or inform engineering design choices?

The answers to these questions depend entirely on the purpose of the study, and the questions being asked. We review the various efforts to understand microbial communities using GENREs, and describe each modeling framework utilized to-date.

### Compartmentalization

The first framework devised for linking GENREs together is an extension of the compartmentalization approach used for eukaryotic GENREs<sup>27</sup> (Figure 3(a)). In eukaryotic models, organelles and other compartments are divided, and reactions specific to each compartment are separated by transport reactions.<sup>27</sup> Along the same lines, multiple species-level GENREs are incorporated into a large ‘meta-stoichiometric matrix’ and transport reactions are explicitly added to enable metabolite flux between species compartments, often with an extracellular compartment inserted as a representation of the local environment.

The first community GENRE was developed to represent the mutualistic interaction between *Desulfovibrio vulgaris* and *Methanococcus maripaludis*.<sup>23</sup> The species GENREs consisted only of reactions in central metabolism and were linked using a compartmentalization approach, with shared byproducts and exchange reactions flowing through a shared compartment. FBA was used to estimate optimal growth rate and metabolite fluxes, where the objective function was chosen to be a linear, weighted combination of the biomass functions for each species (with the weights based on experimentally determined species biomass ratios in active communities). Several results highlight the types of questions that can be addressed using this modeling framework. First, FBA results were closest to experimental results during the active-growth phase, where there were no nutrient limitations, which is consistent with the explicit assumption of pseudo-steady-state growth in FBA. Next, simulated flux patterns of primary metabolites matched experimental measurements, such as the large flux in *D. vulgaris* from lactate (the sole carbon substrate) to acetate, with some CO<sub>2</sub> and formate production, and production of a reduced compound, either hydrogen or formate. Flux through *M. maripaludis* showed consumption of acetate or CO<sub>2</sub> and production of CH<sub>4</sub>. Further, *in silico* simulations offered insight into the amount of non-productive ATP hydrolysis required to match experimentally measured biomass. This study highlights strengths of the compartmentalization and optimization-based approaches. It allowed the exploration of theoretical limits on growth and nutrient fluxes as a function of the metabolic network structure.

Other studies have successfully utilized the same compartmentalization approach. In one study, this approach was used to computationally design media conditions that induce commensalism or mutualism between microbe pairs.<sup>28</sup> In another, three GENREs (*Bacteroides thetaiotamicron*, *Eubacterium rectale*, and *Methanobrevibacter smithii*) were used to explore the impact of the gut microbiome on host metabolism.<sup>38</sup> To accomplish this, two optimization frameworks were defined which are referred to as the  $\alpha$ -problem and  $\beta$ -problem. The  $\alpha$ -problem is used to predict the uptake and secretion of metabolites when the diet and species abundances are known. The  $\beta$ -problem is the inverse, where the model predicts species abundances when metabolite uptake and secretion rates are known. The results from this novel analytical approach are validated using experimental data from gnotobiotic mice. Finally, a host–pathogen interaction between *Mycobacterium tuberculosis* and an alveolar macrophage was simulated using this compartmentalization strategy.<sup>26</sup> The GENRE for *M. tuberculosis* was included as a compartment within the macrophage model, effectively representing a specific metabolic state that *M. tuberculosis* can inhabit during infection.<sup>26</sup> Several other groups have published models that utilize this compartmentalization approach, summarized in Table 1.

The compartmentalization framework is an intuitive and simple way to represent microbial interactions. This approach has been used more frequently than any other, providing mechanistic insight into community metabolism and good agreement with experiment. However, the compartmentalization strategy may limit the types of analyses that can be performed. First, this representation of a community inherently forces an assumption of balanced growth making it difficult to account for metabolite accumulation in the environment because of steady-state constraints in FBA. Second, single-level optimization-based analyses of compartmentalized models often assume that each species in the community is growing optimally (i.e., the objective function in FBA is often assumed to be a combination of the objective functions for each species). Third, species abundance is assumed to be fixed, rather than allowing for changes in abundance as interactions unfold. These and other limitations are addressed in more recent analytical frameworks.

## Community Objectives

The OptCom approach is an extension of the basic compartmentalization strategy that allows for a

community-level objective function<sup>31</sup> (as opposed to only considering the species-level objective functions as described above<sup>23</sup>) (Figure 3(a)). A nested, bi-level optimization framework enables the simulation of several classes of metabolic interactions, including mutualism, synergism, commensalism, parasitism, or competition. For example, a mutualistic interaction can be represented by setting the outer optimization problem to maximize the biomass of two interacting community members, subject to the inner optimization conditions for each species. The inner optimization conditions can be customized, and may include maximization of biomass production or alternative objective functions and steady-state constraints. However, a parasitic interaction may be better represented by setting the community objective function to only maximize parasite network biomass production.<sup>31</sup> Beyond the flexibility to represent many qualitatively different types of interactions, OptCom offers a powerful way to think about community interactions: as a result of competing objectives between all community members.

Given the bi-level structure of OptCom, a distinct advantage of this framework is the ability to explore trade-offs between individual and community objectives. A hypothetical example could be that two species maintain suboptimal metabolic states allowing them to catabolize disparate carbon sources and share the resulting byproducts. OptCom is an excellent tool for exploring and explaining such trade-offs between individual metabolic states and community-level optimality. In summary, by altering the community objective and the constraints on interspecies fluxes as part of the outer problem, OptCom can be used to explore many types of communities, and the reasons for observed interactions with respect to trade-offs between objectives.

In a representative study, OptCom was used to simulate the interaction between two gut bacteria—*Bifidobacterium adolescentis*, an acetate producer, and *Faecalibacterium prausnitzii*, an acetate consumer and butyrate producer.<sup>43</sup> Flux variability analysis (FVA) was utilized to explore the possible range of flux values for shared metabolites (such as acetate).<sup>43</sup> OptCom is computationally expensive and not appropriate for some optimization solvers due to the nonlinear constraints.<sup>43</sup> In addition, the analysis is sensitive to the user-defined optimization functions and flux constraints. Therefore, OptCom may be less suitable for poorly defined communities where the metabolic interactions are less well known. Studies that have used OptCom are summarized in Table 1.

## Dynamic Analysis

Standard FBA results in a set of fluxes—or metabolite consumption/production rates—across a GENRE during pseudo-steady state conditions. In dynamic FBA (dFBA) these fluxes are integrated over time (using standard numerical integration techniques).<sup>51</sup> With dFBA, it is possible to simulate changes to initial conditions over time, including the consumption and production of metabolites, changes in biomass, and shifts in metabolism in response to environmental changes. dFBA provides an entire time course, as opposed to a single snapshot from standard FBA. Kinetic parameters, particularly relating to uptake rates of limiting metabolites (such as glucose and oxygen),<sup>32,51</sup> are required for the implementation of dFBA. Challenges with the implementation of dFBA include an increased computational load and a paucity of the required kinetic parameters for many systems (Figure 3B).

In the dFBA framework, metabolites are free to accumulate or disappear. Species abundance and metabolic states are free to change in response to interactions and changing environment. Thus, the need for defining a community objective function and to set proper bounds on interspecies fluxes is obviated, given that the proper kinetic parameters are known. Furthermore, it is possible to extend other community modeling methods—including OptCom—and perform dynamic analysis.<sup>44</sup>

As an example of multispecies dFBA the co-culture of *Escherichia coli* and *Saccharomyces cerevisiae* was modeled.<sup>32</sup> Each microbe was capable of consuming a unique sugar (glucose or xylose), and the simulations were used to optimize community ethanol production. A similar approach was taken to model the co-culture of *S. cerevisiae* and *Schereffomyces stipitis*<sup>39</sup> in which the production and degradation of growth-inhibitory compounds such as furfural were represented and growth conditions for ethanol production were optimized.

dFBA has been used to identify emergent biosynthetic capacity in 6670 unique two-species communities.<sup>45</sup> FBA was used to estimate microbial growth at each time point, constrained by kinetic uptake parameters for limiting nutrients.<sup>45</sup> Exchange fluxes can take on a range of possible values, so the lower bound of each was determined by FVA and the sum of all exchange reaction fluxes was minimized. In this way, a reproducible time course was produced by each simulation. *In silico* species could share metabolites through the shared environment, and emergent metabolites were those that could be produced by a co-culture simulation, but not by either species individually.<sup>45</sup> Interestingly, there is a clear window of phylogenetic distances in which two interacting

species are more likely to exhibit emergent metabolic capacity.<sup>45</sup>

Another use of dFBA is to capture spatio-temporal dynamics in a community of organisms. In one case, dFBA was used to model the formation of *Pseudomonas aeruginosa* biofilm, where the ‘community’ is the collection of cells in different metabolic states.<sup>52</sup> In this study, a GENRE corresponding to *P. aeruginosa* was used in a dFBA framework to estimate metabolite secretion, production, and diffusion over time across compartments in an agent-based model of biofilm, recapitulating known features of biofilm formation such as oxygen-limited biofilm growth. In a similar study, dFBA was used to model the spatio-temporal dynamics of three-species microbial communities on a 2D surface.<sup>46</sup> dFBA was used to estimate biomass production, and nutrient concentration changes in local compartments where diffusion allowed changes to impact connected compartments. The authors report successfully predicting the steady-state species composition of an engineered three-member community.<sup>46</sup>

These examples demonstrate the versatility of dFBA for modeling small, well-characterized communities, performing large-scale surveys of many potential communities and the possible emergent properties among them, and accounting for spatial dynamics. dFBA represents an exciting and underexplored area of GENRE analysis. dFBA may present a community modeling option with fewer up-front assumptions. However, use of dFBA may be constrained by computational limitations, given the inherent increase in computation over a time-course<sup>45</sup> or spatial environment.<sup>46</sup> Furthermore, dFBA relies on additional kinetic parameters, which may nullify a primary advantage of FBA-based techniques which avoid extensive parameterization. A summary of dFBA-based community models is presented in Table 1.

It is worthwhile to compare dFBA with other established dynamic models of microbial communities. The Activated Sludge Model (ASM) has a long history in bioreactor control for wastewater treatment.<sup>53</sup> These models are based on ordinary differential equations, and predict the changes in nutrient concentrations and microbial abundances over time. Nutrients of interest are grouped (e.g., carbon, nitrogen, and phosphorus sources), and microbes grouped according to nutrient utilization (e.g., nitrifying bacteria, phosphorus-accumulating bacteria). An ASM can predict the change in abundance of each microbial group as a function of nutrient concentrations, and subsequent changes in nutrient concentrations as a result of microbial growth. The parameters for these

models are chosen to fit experimental data. An ASM does not represent each taxon individually, nor does it account for the metabolic differences between organisms within a group. For example, two bacteria from different taxa may both be classified as ‘nitrifying’. They would likely have different overall metabolic networks, and therefore respond differently to changes in carbon and phosphorus concentrations.<sup>54</sup> An ASM would not account for these taxon-specific metabolic differences, thus resulting predictions may be misleading. In contrast, modeling frameworks that can capitalize on genome-scale metabolic reconstructions, such as dFBA, are capable of accounting for these metabolic differences and can potentially improve prediction accuracy. In addition, dFBA dynamics are typically a function of specific uptake rates while ASM dynamics are a result of fitting the system kinetics to observed data. A recent review describes ASM models as well as other alternative community modeling frameworks that are not based on metabolic network reconstructions.<sup>55</sup>

### ‘Enzyme Soup’ FBA

In contrast to the simple compartmentalization approach, OptCom, and dFBA, the ‘enzyme soup’ approach<sup>18</sup> ignores species boundaries entirely (Figure 3(c)). The emphasis is on exploration of the metabolic potential of an entire community rather than the interactions between species within a community. A community-level ‘enzyme soup’ GENRE is produced by annotating a meta-omic dataset for enzyme presence, and the associated reactions are agglomerated into a single set without an attempt to segregate reactions by species. In this framework, any reaction from any species can potentially connect with any other reaction into a ‘meta-pathway’. Early work on this approach ignored stoichiometric constraints within this network, and examined the topological differences between networks reconstructed from healthy and diseased metagenomic data.<sup>56</sup> More recent work pioneered the analysis of these community-level GENREs using constraint-based methods such as FBA to predict biomass production and substrate utilization.<sup>48</sup> In this work, the authors base their reconstruction on metaproteomic data from a naphthalene-enriched soil microbe dataset. They maintain stoichiometric constraints, and assign metabolic activity to taxa within the community based on the taxonomic annotation of the enzymes in the model. The biomass function is assumed to be a generalized biomass equation borrowed from other organisms, under the assumption that many components of biomass are common to many organisms.<sup>48</sup>

The enzyme soup approach has been used successfully to explore the metabolic capacity of complex natural communities. In one study, the interaction of several microbial ‘guilds’ was studied using both the compartmentalization approach and the enzyme soup method (which they refer to as the ‘pooled reactions’ method), and a ‘nested consortium analysis’.<sup>25</sup> The guilds represent community functions attributable to prokaryotic oxygenic phototrophs, filamentous anoxygenic phototrophs, and sulfate reducers found in thermophilic, phototrophic mat communities from Yellowstone National Park (USA). The enzyme soup method is appropriate when there is limited *a priori* knowledge about the community; conversely, the compartmentalization approach is a more accurate representation of the biology if extensive knowledge is available about the various community members.<sup>25</sup> The nested consortium analysis is based on elementary mode analysis (EMA).<sup>25,57</sup> First, elementary modes (i.e., a pool of valid physiological states) are identified for each guild. Second, elementary modes are re-computed for the entire community using the guild-level modes as input.<sup>25</sup> This ‘nested consortium analysis’ also requires significant *a priori* knowledge of the community in order to select useful elementary modes at the guild-level.

The issue of compartmentalization in GENREs has been discussed in the context of eukaryotic organisms, in which extensive compartmentalization is used to represent organelles and compartments within the cell.<sup>27</sup> For example, analysis of the *S. cerevisiae* GENRE shows that compartmentalization significantly impacts basic properties such as network connectivity and the accuracy of analytical results such as flux values.<sup>27</sup> When considering the “enzyme soup” approach, it is important to consider the loss of accuracy associated with dissolving the boundaries between community members. The main advantage of this approach is the low *a priori* knowledge required, such that it is applicable to little-understood systems. In some sense, the “enzyme soup” strategy can be thought of as placing bounds on the potential metabolic capacity of a microbial community. Further compartmentalization will provide more specific solutions within those bounds. Studies utilizing the enzyme soup method are summarized in Table 1.

### Other Methods

Other GENRE-based methods of community analysis have been explored, providing yet more vantage points from which to view the metabolism of complex communities. Network Expansion is an algorithm in which the metabolic potential of a set of reactions is



explored<sup>58</sup> (Figure 3(d)). The algorithm starts with a set of metabolites as input (the environment). An initial set of reactions that can use the input metabolites as substrates are added to the network. This network is expanded in subsequent steps as products of the previously added reactions are made available. New reactions that use some part of the accumulating pool of metabolites are added to the network. This approach was extended to a community-level analysis.<sup>24</sup> Given the reaction sets from two organisms, the algorithm assumes that any intermediate metabolites can be shared, and so performs network expansion by pooling the reaction sets from both organisms.<sup>24</sup> This algorithm has been used to identify emergent properties of pairs of microbes, where the combination can produce metabolites that cannot be produced by either parent species.<sup>24</sup>

In contrast to the network expansion method,<sup>56</sup> approaches based on identifying the ‘seed set’ of a species-level GENRE maintain species-specific boundaries and can be used to estimate metabolic competition or cooperation<sup>59,40</sup> (Figure 3(e)). In this graph-based approach, the stoichiometric structure of the GENRE is decomposed to create a directed graph from all substrate-to-product pairs. The resulting graph represents paths between metabolites, but does not contain stoichiometric information. Metabolites that are consumed, but never produced by any reaction in the network, are assigned as network inputs, or the ‘seed set’.<sup>59</sup> The seed sets for multiple species have been used to estimate the potential for competition or cooperation between species, demonstrating that species tend to co-occur in nature with mutual competitors.<sup>40</sup> These graph-based methods ignore stoichiometry and are therefore not useful for making flux predictions, but rather more generalized statements about network similarity. This approach may be useful when using draft-quality models, for which the accuracy of FBA or similar analyses may be low.

As a final GENRE-based community analysis, we also make note of the comparative approach, which ignores interactions among species and seeks only to identify functional differences between GENREs (Figure 3(f)). For example, a comparative analysis of *Burkholderia cenocepacia* and *Burkholderia multivorans* revealed the unique reactions associated with each, and identified functional outcomes associated with those differences, such as differential virulence factor production capacity.<sup>60</sup> Similarly, a comparative analysis of two strains of *Lactobacillus casei* highlighted functional differences between these industrially-relevant strains.<sup>42</sup> Such comparative analyses can help to identify the functional roles

of species within large communities by identifying both redundancy and differential metabolic capacity between community members.

## Multiscale Models

Outside of the community modeling methods discussed here which are effectively single-scale models (with the possible exception of the spatial models that incorporate dFBA<sup>46,52</sup>), GENREs have also been successfully incorporated into multiscale models. GENREs of soil microbes have been integrated with a reactive transport model based on local geochemical conditions.<sup>61</sup> A GENRE of the human hepatocyte has been connected with a multicompartment pharmacokinetic model of the human body.<sup>62</sup> Opportunities for multiscale modeling abound with the increasing availability of omics data. For example, in one mouse study the presence of several gut microbes was modulated, and the resulting metabolomics data were fit with a compartmental model that represented the host liver, pancreas, kidney, and adipose tissues.<sup>63</sup> It is clear how a GENRE-based compartmental model could be integrated into an analysis of similar data. These studies highlight the potential gains from combining the strengths of multiple modeling frameworks.

## Experimental Data That Meshes with GENRE-Based Community Analysis

GENRE-based community models require data to be constructed, constrained, and validated; for example, quantifying species within a community is important for constraining objective functions (ratios of biomass equations) or validating dFBA simulation results. Many omics technologies are well-suited for this purpose and have been used extensively in GENRE studies of single species. Additional experimental approaches provide insight into microbial community function that may prove useful to GENRE community models in the near future.

Metagenomics answers the question ‘who is there and what might they be able to do?’ The composition of many highly complex microbial communities has been elucidated in recent years through shotgun metagenomic sequencing techniques.<sup>7</sup> This approach can be used to identify the taxa that are present, estimate relative abundances of each taxon, and provide a ‘parts list’ of genes that are present in the community.<sup>64</sup> Metagenomic data can be directly translated into community GENREs using any of the techniques previously discussed; however, difficulties arise during taxonomic assignments within the community. Below, we discuss tools that have been developed to address this issue in metagenomic data.

A parallel technology that may eventually overcome challenges in metagenomic sequencing is single-cell sequencing.<sup>65,66</sup> While throughput is lost, confidence is gained in assignment of function to a distinct organism within the community.

Metatranscriptomics and metaproteomics both help address the question ‘what are they doing?’ Functional data can be used as the basis for constructing a GENRE, or more commonly, to constrain an existing GENRE in order to estimate metabolic pathway usage under specific conditions.<sup>37</sup> Once again, assigning mRNA transcripts or proteins to a specific community member can prove challenging. Single-cell transcriptomics and proteomics can help with this difficulty,<sup>67</sup> as can the generation of reference genomes to which transcripts and peptides can be mapped.<sup>68</sup>

Other techniques allow the quantification of community member abundances, including real-time PCR, flow cytometry, and in some cases, Coulter counters. Real-time PCR can be used to quantify the abundances of specific community members in a mixed culture.<sup>69</sup> There is generally good correlation between real-time PCR and other measures such as optical density and viability data, unless the culture is under stress, in which case real-time PCR tends to overestimate the number of viable cells.<sup>69</sup> Even so, accurate maximum growth rates can be calculated from real-time PCR regardless of culture conditions.<sup>69</sup> Flow cytometry can be an effective technique to quantify community member abundance, but is limited by the availability of species-specific fluorescent markers.<sup>70</sup> Coulter counters have been used in two-member communities where species vary drastically in size.<sup>71</sup>

The burgeoning field of metabolomics offers novel tools to study the anabolic and catabolic capacities of simple and complex communities. Both targeted (measure specific pre-defined metabolites) and non-targeted (measure as many metabolites as possible, without pre-selection) methods have been applied to studies of microbial communities.<sup>72</sup> Both nuclear magnetic resonance (NMR) and mass spectrometry can be applied in targeted and non-targeted ways, and can provide quantification of metabolite abundances.<sup>73–75</sup> Targeted or non-targeted metabolomic data can be used to constrain, and validate functional outcomes of GENRE simulations. A technique of interest, MALDI-TOF imaging mass spectrometry, is a targeted method capable of measuring metabolite concentrations over a physical space.<sup>76,77</sup> This technique has been used to identify compounds that are produced during the co-culture of *Streptomyces coelicolor* with other actinomycetes, and how resulting metabolites localize spatially.<sup>76</sup> This technique may pair particularly well with

GENRE modeling techniques that account for spatial information.<sup>46,52</sup>

Many tools for studying microbial communities are applied to the aggregate, which often makes it necessary to partition the data in order to assign activity to a specific community member. Spent media experiments provide a way to dissect cellular interactions within co-culture systems. Using this technique, spent media is produced by growing the first organism in fresh media. The first organism is removed (i.e., by filtration sterilization) resulting in ‘sterile’ media, and the resulting supernatant is then used to culture a second organism. This technique was used to demonstrate that *Enterobacter cloacae* produced fermentation byproducts that enhanced hydrogen production by *Rhodobacter sphaeroides*.<sup>78</sup> Further, this technique was used to determine the ability of *Lactobacillus* to inhibit growth of *Candida albicans*, *Gardnerella vaginalis*, and *Streptococcus agalactiae*.<sup>79</sup> A disadvantage of this technique is the inability to maintain interspecies signaling, such that the first species (used to produce spent media) cannot respond to the presence of the second species. However, the advantages of this technique are that the direction of interaction and causality can easily be determined, and individual species can be monitored because they are grown separately. In addition, spent-media experiments can be paired with metabolomics and other measures to produce data that can be integrated with GENRE simulations.

Co-culture of community members on solid surfaces such as agar plates can also help to overcome difficulties in partitioning community members. ‘Cross streak’ analysis is a technique whereby single cultures are mixed on the surface of a plate.<sup>80</sup> Growth and other phenotypes can be visually assessed or paired with other tools such as imaging mass spectrometry.<sup>76</sup> This approach was elegantly used to identify members of the human microbiota that promote growth of antibiotic resistant *Staphylococcus aureus*.<sup>81</sup> At the most basic level, these screens can be used to qualitatively determine the nature of interactions, which can be used to interpret results of GENRE simulations.

## CHALLENGES AND OPPORTUNITIES

### Partitioning Communities

Assigning activity to particular species is a fundamental difficulty working with mixed communities and is further compounded when complete genomes are not available for all species in the community.<sup>82</sup> This difficulty assigning activity to a particular species motivates the use of ‘enzyme soup’ methods as discussed above. Considerable progress has been made recently

to assemble catalogues of reference genomes, but a great deal remains to be done (and will likely never be ‘complete’).<sup>83</sup> Traditional genome sequencing has relied on culturing individual isolates in order to extract large quantities of purified DNA, but this is not feasible for the vast majority of organisms.<sup>84</sup> Alternative methods rely on computationally partitioning genomes from mixed metagenomic sequencing samples.<sup>82,84–91</sup> Three main types of information are used to bin member genomes from mixed samples: (1) DNA composition-based methods, which rely on an empirically observed trend for genomes to display unique ‘k-mer’ frequencies (patterns of one to five bases)<sup>89,92,93</sup>; (2) Abundance variations across many samples, where contiguous DNA segments with similar abundance profiles across many samples are likely to originate from the same organism<sup>82,84–86</sup>; (3) Taxonomic annotations derived from similarity to known taxa.<sup>91,94–96</sup> There are many active efforts to improve the isolation of individual genetic information from mixed metagenomic samples, and these efforts can translate directly to improved GENRE construction. Species-specific genomes, particularly from non-culturable organisms, will be invaluable resources for understanding the function of complex communities through GENRE-based analysis.

### Automating High-Quality Metabolic Reconstructions

Ideally, automated generation of GENREs from metagenomic or genomic data will result in models that have predictive power with minimal manual curation or experimental validation. In practice, however, even the most well-curated GENREs cannot fully recapitulate experimental phenotypes.<sup>97</sup> Therefore, the validity and usefulness of automatically generated GENREs should be assessed by their utility relative to manual reconstructions. A GENRE that can be used to predict growth conditions and gene essentiality will allow a myriad of applications in community modeling and serves as an attainable short-term goal for the development of algorithms for automatically generating GENREs.

Several attempts at automated and semi-automated creation of GENREs have been made, which have been compared and reviewed previously.<sup>98,99</sup> Many studies report using GENREs created with a combination of automated methods and manual curation,<sup>41,42,43,100</sup> but there are few, if any, reports of automatically generated GENREs used to contextualize experimental data without manual curation to some degree. Greater precision and throughput is necessary when generating GENREs for

uncharacterized, unique communities composed of diverse sets of microbes that vary greatly across environments or hosts, as is often the case in biomedical or ecological applications.

The next generation of algorithms for automated GENRE generation includes a variety of promising approaches in the context of community modeling. When draft reconstructions are created directly from genome annotations, gap-filling is needed to connect dead-end reactions to produce a functional network. Parsimony-based,<sup>101,102</sup> likelihood-based,<sup>103</sup> and phylogeny-based<sup>104</sup> strategies have been developed to fill gaps during automated reconstruction without relying on experimental data. Parsimony-based methods posit that the most parsimonious pathway that fills a gap is the most likely to occur, which results in a smaller GENRE than other gap-filling methods. Likelihood-based methods incorporate multiple gene annotations and use them during the gap-filling stage to present alternative reactions that are each given a likelihood score, greatly expanding the space of possible pathways. While likelihood-based and parsimony-based methods provide similarly accurate results when predicting experimental phenotypes,<sup>103</sup> the former provides a framework for finding low-quality gene annotations, which, when removed or fixed, may improve the quality of GENREs created with other methods. Phylogeny-based methods start with the assumption that reactions tend to be more conserved in closely related species than distantly related species. Phylogenetic relationships have been used in the context of gene annotation by assuming that functionally linked proteins have correlated evolution, thus homologs for functionally linked proteins are likely to be present in the same subset of organisms.<sup>105</sup> The outcome of this assumption is that sets of proteins involved in the same function or metabolic pathway can be more accurately annotated in newly sequenced genomes when corresponding homologs involved in that function are identified in another species.

Evolutionary distances have been shown to have a significant, predictable relationship to gene essentiality and growth phenotypes.<sup>106</sup> A framework called CoReCo has been developed which assumes such relationships *a priori* to enhance the GENRE creation process for multiple species simultaneously.<sup>104</sup> Out of the existing context-based methods, incorporating phylogenetic relationships within a community to guide model creation is particularly interesting because relationships should be obtainable from metagenomic data. Conversely, the assignment of species from metagenomic data could be enhanced by



evaluating the function of GENREs created based on multiple putative phylogenies for putative species.

These gap-filling methods present examples of the types of information that need to be integrated in reconstruction algorithms given the constraints of microbial communities. Assumptions based on evolutionary arguments have proven potential in this regard,<sup>107</sup> and may have exceptional power that needs to be explored in communities containing both closely and distantly related species. Finally, integrating multiple assumptions and sources of information has the potential to increase GENRE validity in an additive manner and should be explored further.

A final step in model generation that may be particularly relevant to community analyses is model reconciliation. Reconciliation removes the differences between GENREs that represent non-biological noise created through the reconstruction process. Such noise can be due to many factors including, but not limited to, gene annotation uncertainties, differences in the naming conventions in reaction databases used, and unspecified or incorrectly specified reaction reversibility.<sup>107</sup> When reconciliation is performed between models for two related species, the result is typically a reduction in the number of reactions that are unique to each model. This could be particularly useful for searching for therapeutic targets in pathogens that are closely related to a non-pathogen in the same community, as is common in the human microbiome.<sup>108</sup> Reconciliation would result in greater certainty that a target is unique to the pathogen, reducing the probability of off-target effects in commensal organisms.

Reconciliation of automatically generated GENREs from a community may be particularly useful because differences in sequencing quality are likely to be small in a community sample and the same model generation algorithm is likely to be used for all species. The resulting GENREs may be very effective at revealing noise introduced from the reconstruction method, since differences in sequencing and model generation algorithms are controlled. However, reconciliation between two models currently requires a significant amount of manual input and user choice, making it difficult to scale-up to large communities.

## Integrating Omics

High-throughput omics technologies such as transcriptomics, proteomics, metabolic flux analysis, and metabolomics all present opportunities for new understanding of microbial communities when integrated with GENRE analysis, but challenges remain with the best approaches for data integration. As with

metagenomic information, partitioning omics datasets and assigning them to a particular community member remains a challenge. Along these lines, difficulties may arise when multiple, highly similar strains exist within the same community interrogated with omics approaches. Transcriptomics and proteomics will both benefit from advances in genome binning and assembly and the resulting species-specific references.<sup>83</sup> Leveraging proteomic techniques, one study used peptide-based <sup>13</sup>C metabolic flux analysis to assign metabolic fluxes to species within a community.<sup>109</sup> Metabolomics is perhaps the most challenging, as it can be very difficult to trace the origins of a metabolite in a shared supernatant. Because of inherent limitations in metabolomic technologies that prevent assignment of metabolites to specific community members, GENRE based analyses offer the most effective way to generate *ab initio* hypotheses about the partitioning of metabolic roles within microbial communities.

Regardless of the method of omics partitioning, it is expected that existing methods for omics integration into single-species models will translate well to community models.<sup>37,38</sup> Many tools for integration of expression data and proteomics data have been developed and validated for individual species<sup>110</sup>; for example, GIMME and MADE represent the trade-off between assumptions and data.<sup>111,112</sup> GIMME constrains a GENRE with expression data by requiring user-supplied thresholds for each gene, and then optimizing the solution based on consistencies in pathway up/down regulation.<sup>111</sup> MADE is a related algorithm that infers gene-specific thresholds based on multiple expression datasets.<sup>112</sup> If expression data can be easily mapped to reference genomes or proteomes for individual species, these and other existing tools will be applicable. However, in a multispecies community, species abundance is convoluted with gene expression, and precautions should account for such effects. The integration of meta-omics data into community GENRE models may follow a similar path to that of genome assembly algorithms. Metagenomic assembly tools are very similar to single genome assembly tools, with minor changes to address the challenges associated with mixed communities.<sup>113,114</sup> Perhaps omics data integration algorithms will follow a similar path.

## Engineering Communities

Perhaps the most exciting aspect of GENRE community models is the opportunity for community design and engineering. As modeling techniques improve, it is hoped that mechanisms of interspecies interaction will become better understood and more predictable. Early computational tools have already proved valuable from a community engineering standpoint, as



demonstrated by the ability to design nutritional environments that modulate the interactions between species in co-culture.<sup>28</sup> FBA-based methods have indicated optimal growth rates of *E. coli* which could be subsequently obtained by adaptive evolution experiments<sup>36,115</sup>; likewise, it may be possible to use FBA-based methods to predict the necessary individual adaptations of synthetic auxotrophs in co-culture.<sup>116</sup> When large community GENRE models are well-validated, they can be used to explore the impact of specific therapeutic interventions such as prebiotics, probiotics, and targeted removal of species in the human gut microbiome.<sup>20,22,117</sup> Analyses such as gene-knockout screens may be extended to species-knockout screens (i.e., sequential removal of each species from the community and analysis of the resulting consequences).<sup>60,118,119</sup> For this type of analysis, it is not currently known what adjustments to flux predictions need to be made to produce accurate simulations. In the case of gene-knockout simulations, techniques such as minimization of metabolic adjustment (MOMA) are used to predict the updated flux distribution.<sup>120</sup> Analogous algorithms will likely be useful for community-level analyses.

It is also unclear how current metabolic engineering tools will be extended to community engineering applications. Algorithms such as OptKnock, OptGene, or the Redirector algorithm have proven useful at a single-organism level.<sup>121–123</sup> Similar optimization frameworks may be devised for community models,

but complexity scales not only with the number of species, but also with the many ways species models can be conjoined. There are great incentives to advance such engineering approaches since examples show that microbial consortia are more efficient and robust than a single engineered species.<sup>12,124</sup>

## CONCLUSION

Metabolic systems biology of microbial communities is an exciting and rapidly developing field with the potential to revolutionize our understanding of microbial communities of societal importance. Existing methods have shown promise, including compartmentalization, separating species-level and community-level objectives, dynamic analysis, the ‘enzyme-soup’ approach, multiscale modeling, and others. The rise of omics technologies has enabled high-level views of microbial community composition and metabolism, but it remains a challenge to partition community function and assign it to individual community members. Future work is also needed to integrate omics data into community-level metabolic models. Moreover, the sheer number of species in many microbial communities demands new automated reconstruction methods that result in GENREs without the need for further manual curation. As technologies and modeling frameworks improve, we expect that there will be corresponding advances in the fields of ecology, health science, and microbial community engineering.

## REFERENCES

1. Rousk J, Bengtson P. Microbial regulation of global biogeochemical cycles. *Front Microbiol* 2014, 5:305–307. doi:10.3389/fmicb.2014.00103.
2. Chaparro JM, Sheflin AM, Manter DK, Vivanco JM. Manipulating the soil microbiome to increase soil health and plant fertility. *Biol Fertil Soils* 2012, 48:489–499. doi:10.1007/s00374-012-0691-4.
3. Kinross JM, Darzi AW, Nicholson JK. Gut microbiome-host interactions in health and disease. *Genome Med* 2011, 3:14. doi:10.1186/gm228.
4. Smid EJ, Erkus O, Spus M, Wolkers-Rooijackers JC, Alexeeva S, Kleerebezem M. Functional implications of the microbial community structure of undefined mesophilic starter cultures. *Microb Cell Fact* 2014, 13(Suppl 1):S2. doi:10.1186/1475-2859-13-S1-S2.
5. Van Der Heijden MGA, Bardgett RD, Van Straalen NM. The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecol Lett* 2008, 11:296–310. doi:10.1111/j.1461-0248.2007.01139.x.
6. Holler T, Widdel F, Knittel K, Amann R, Kellermann MY, Hinrichs K-U, Teske A, Boetius A, Wegener G. Thermophilic anaerobic oxidation of methane by marine microbial consortia. *ISME J* 2011, 5:1946–1956. doi:10.1038/ismej.2011.77.
7. Handelsman J. Metagenomics: application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* 2004, 68:669–685. doi:10.1128/MMBR.68.4.669-685.2004.
8. Meyer F, Paarmann D, D’Souza M, Olson R, Glass E, Kubal M, Paczian T, Rodriguez A, Stevens R, Wilke A, et al. The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 2008, 9:386. doi:10.1186/1471-2105-9-386.

9. Hunter S, Corbett M, Denise H, Fraser M, Gonzalez-Beltran A, Hunter C, Jones P, Leinonen R, McAnulla C, Maguire E, et al. EBI metagenomics – a new resource for the analysis and archiving of metagenomic data. *Nucleic Acids Res* 2014, 42:600–606. doi:10.1093/nar/gkt961.
10. Manor O, Levy R, Borenstein E. Mapping the inner workings of the microbiome: genomic- and metagenomic-based study of metabolism and metabolic interactions in the human microbiome. *Cell Metab* 2014, 20:742–752. doi:10.1016/j.cmet.2014.07.021.
11. Buffie CG, Bucci V, Stein RR, McKenney PT, Ling L, Gobourne A, No D, Liu H, Kinnebrew M, Viale A, et al. Precision microbiome reconstitution restores bile acid mediated resistance to *Clostridium difficile*. *Nature* 2014, 517:205–208. doi:10.1038/nature13828.
12. Brenner K, You L, Arnold FH. Engineering microbial consortia: a new frontier in synthetic biology. *Trends Biotechnol* 2008, 26:483–489. doi:10.1016/j.tibtech.2008.05.004.
13. Brune KD, Bayer TS. Engineering microbial consortia to enhance biomining and bioremediation. *Front Microbiol* 2012, 3:1–6. doi:10.3389/fmicb.2012.00203.
14. Seth EC, Taga ME. Nutrient cross-feeding in the microbial world. *Front Microbiol* 2014, 5:350. doi:10.3389/fmicb.2014.00350.
15. Morris BEL, Henneberger R, Huber H, Moissl-Eichinger C. Microbial syntrophy: interaction for the common good. *FEMS Microbiol Rev* 2013, 37:384–406. doi:10.1111/1574-6976.12019.
16. Oberhardt MA, Palsson BØ, Papin JA. Applications of genome-scale metabolic reconstructions. *Mol Syst Biol* 2009, 5:320. doi:10.1038/msb.2009.77.
17. Røling WFM, Ferrer M, Golyshin PN. Systems approaches to microbial communities and their functioning. *Curr Opin Biotechnol* 2010, 21:532–538. doi:10.1016/j.copbio.2010.06.007.
18. Klitgord N, Segrè D. Ecosystems biology of microbial metabolism. *Curr Opin Biotechnol* 2011, 22:541–546. doi:10.1016/j.copbio.2011.04.018.
19. Karlsson FH, Nookaew I, Petranovic D, Nielsen J. Prospects for systems biology and modeling of the gut microbiome. *Trends Biotechnol* 2011, 29:251–258. doi:10.1016/j.tibtech.2011.01.009.
20. Borenstein E. Computational systems biology and in silico modeling of the human microbiome. *Brief Bioinform* 2012, 13:769–780. doi:10.1093/bib/bbs022.
21. Thiele I, Heinken A, Fleming RMT. A systems biology approach to studying the role of microbes in human health. *Curr Opin Biotechnol* 2013, 24:4–12. doi:10.1016/j.copbio.2012.10.001.
22. Shoaie S, Nielsen J. Elucidating the interactions between the human gut microbiota and its host through metabolic modeling. *Front Genet* 2014, 5:86. doi:10.3389/fgene.2014.00086.
23. Stolyar S, Van Dien S, Hillesland KL, Pinel N, Lie TJ, Leigh J, Stahl D. Metabolic modeling of a mutualistic microbial community. *Mol Syst Biol* 2007, 3:92. doi:10.1038/msb4100131.
24. Christian N, Handorf T, Ebenhöf O. Metabolic synergy: increasing biosynthetic capabilities by network cooperation. *Genome Informatics Int Conf Genome Informatics* 2007, 18:320–329.
25. Taffs R, Aston JE, Brileya K, Jay Z, Klatt CG, McGlynn S, Mallette N, Montross S, Gerlach R, Inskeep WP, et al. In silico approaches to study mass and energy flows in microbial consortia: a syntrophic case study. *BMC Syst Biol* 2009, 3:114. doi:10.1186/1752-0509-3-114.
26. Bordbar A, Lewis NE, Schellenberger J, Palsson BØ, Jamshidi N. Insight into human alveolar macrophage and *M. tuberculosis* interactions via metabolic reconstructions. *Mol Syst Biol* 2010, 6:422. doi:10.1038/msb.2010.68.
27. Klitgord N, Segrè D. The importance of compartmentalization in metabolic flux models: yeast as an ecosystem of organelles. *Genome Informatics Int Conf Genome Informatics* 2010, 22:41–55. doi:10.1142/9781848165786\_0005.
28. Klitgord N, Segrè D. Environments that induce synthetic microbial ecosystems. *PLoS Comput Biol* 2010, 6:e1001002. doi:10.1371/journal.pcbi.1001002.
29. Sun J, Haveman S, Bui O, Fahland T, Lovley D. Constraint-based modeling analysis of the metabolism of two *Pelobacter* species. *BMC Syst Biol* 2010, 4:174. doi:10.1186/1752-0509-4-174.
30. Freilich S, Zarecki R, Eilam O, Segal E, Henry C, Kupiec M, Gophna U, Sharan R, Ruppin E. Competitive and cooperative metabolic interactions in bacterial communities. *Nat Commun* 2011, 2:589. doi:10.1038/ncomms1597.
31. Zomorodi AR, Maranas CD. OptCom: a multi-level optimization framework for the metabolic modeling and analysis of microbial communities. *PLoS Comput Biol* 2012, 8:e1002363. doi:10.1371/journal.pcbi.1002363.
32. Hanly TJ, Henson M. Dynamic flux balance modeling of microbial co-cultures for efficient batch fermentation of glucose and xylose mixtures. *Biotechnol Bioeng* 2011, 108:376–385. doi:10.1002/bit.22954.
33. Zhuang K, Izallalen M, Mouser P, Richter H, Risso C, Mahadevan R, Lovley D. Genome-scale dynamic modeling of the competition between *Rhodospirillum rubrum* and *Geobacter* in anoxic subsurface environments. *ISME J* 2011, 5:305–316. doi:10.1038/ismej.2010.117.
34. Tzamali E, Poirazi P, Tollis I, Reczko M. A computational exploration of bacterial metabolic diversity identifying metabolic interactions and growth-efficient

- strain communities. *BMS Syst Biol* 2011, 5:167. doi:10.1186/1752-0509-5-167.
35. Heinken A, Sahoo S, Fleming R, Thiele I. Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. *Gut Microbes* 2013, 4:28–48. doi:10.4161/gmic.22370.
36. Khandelwal R, Olivier BG, Röling WFM, Teusink B, Bruggeman FJ. Community flux balance analysis for microbial consortia at balanced growth. *PLoS One* 2013, 8:e64567. doi:10.1371/journal.pone.0064567.
37. Nagarajan H, Embree M, Rotaru A-E, Shrestha PM, Feist AM, Palsson BØ, Lovley DR, Zengler K. Characterization and modelling of interspecies electron transfer mechanisms and microbial community dynamics of a syntrophic association. *Nat Commun* 2013, 4:2809. doi:10.1038/ncomms3809.
38. Shoaie S, Karlsson F, Mardinoglu A, Nookaew I, Bordel S, Nielsen J. Understanding the interactions between bacteria in the human gut through metabolic modeling. *Sci Rep* 2013, 3:2532. doi:10.1038/srep02532.
39. Hanly TJ, Henson M. Dynamic model-based analysis of furfural and HMF detoxification by pure and mixed batch cultures of *S. cerevisiae* and *S. stipitis*. *Biotechnol Bioeng* 2014, 111:272–284. doi:10.1002/bit.25101.
40. Levy R, Borenstein E. Metabolic modeling of species interaction in the human microbiome elucidates community-level assembly rules. *Proc Natl Acad Sci USA* 2013, 110:12804–12809. doi:10.1073/pnas.1300926110.
41. Bartell J, Yen P, Varga JJ, Goldberg JB, Papin J. Comparative metabolic systems analysis of pathogenic Burkholderia. *J Bacteriol* 2014, 196:210–226. doi:10.1128/JB.00997-13.
42. Vinay-Lara E, Hamilton JJ, Stahl B, Broadbent JR, Reed JL, Steele JL. Genome-scale reconstruction of metabolic networks of *Lactobacillus casei* ATCC 334 and 12A. *PLoS One* 2014, 9:e110785. doi:10.1371/journal.pone.0110785.
43. El-Semman IE, Karlsson FH, Shoaie S, Nookaew I, Soliman TH, Nielsen J. Genome-scale metabolic reconstructions of *Bifidobacterium adolescentis* L2-32 and *Faecalibacterium prausnitzii* A2-165 and their interaction. *BMC Syst Biol* 2014, 8:41. doi:10.1186/1752-0509-8-41.
44. Zomorodi AR, Islam MM, Maranas CD. d-OptCom: dynamic multi-level and multi-objective metabolic modeling of microbial communities. *ACS Synth Biol* 2014, 3:247–257. doi:10.1021/sb4001307.
45. Chiu H-C, Levy R, Borenstein E. Emergent biosynthetic capacity in simple microbial communities. *PLoS Comput Biol* 2014, 10:e1003695. doi:10.1371/journal.pcbi.1003695.
46. Harcombe WR, Riehl WJ, Dukovski I, Granger BR, Betts A, Lang AH, Bonilla G, Kar A, Leiby N, Mehta P, et al. Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics. *Cell Rep* 2014, 7:1104–1115. doi:10.1016/j.celrep.2014.03.070.
47. Ye C, Zou W, Xu N, Liu L. Metabolic model reconstruction and analysis of an artificial microbial ecosystem for vitamin C production. *J Biotechnol* 2014, 182–183:61–67. doi:10.1016/j.jbiotec.2014.04.027.
48. Tobalina L, Bargiela R, Pey J, Herbst F, Lores I, Rojo D, Barbas C, Peláez AI, Sánchez J, von Bergen M, et al. Context-specific metabolic network reconstruction of a naphthalene degrading bacterial community guided by metaproteomic data. *Bioinformatics* 2015, 31:1771–1779. doi:10.1093/bioinformatics/btv036.
49. Gianchandani EP, Chavali AK, Papin JA. The application of flux balance analysis in systems biology. *WIREs Syst Biol Med* 2010, 2:372–382. doi:10.1002/wsbm.60.
50. Thiele I, Palsson BØ. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat Protoc* 2010, 5:93–121. doi:10.1038/nprot.2009.203.
51. Mahadevan R, Edwards JS, Doyle FJ. Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys J* 2002, 83:1331–1340. doi:10.1016/S0006-3495(02)73903-9.
52. Biggs MB, Papin J. Novel multiscale modeling tool applied to *Pseudomonas aeruginosa* biofilm formation. *PLoS One* 2013, 8:1–8. doi:10.1371/journal.pone.0078011.
53. Hauduc H, Rieger L, Oehmen A, van Loosdrecht MCM, Comeau Y, Hédut A, Vanrolleghem PA, Gillot S. Critical review of activated sludge modeling: state of process knowledge, modeling concepts, and limitations. *Biotechnol Bioeng* 2013, 110:24–46. doi:10.1002/bit.24624.
54. Oehmen A, Carvalho G, Lopez-Vazquez CM, van Loosdrecht MCM, Reis MAM. Incorporating microbial ecology into the metabolic modelling of polyphosphate accumulating organisms and glycogen accumulating organisms. *Water Res* 2010, 44:4992–5004. doi:10.1016/j.watres.2010.06.071.
55. Bucci V, Xavier JB. Towards predictive models of the human gut microbiome. *J Mol Biol* 2014, 426:3907–3916. doi:10.1016/j.jmb.2014.03.017.
56. Greenblum S, Turnbaugh PJ, Borenstein E. Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. *Proc Natl Acad Sci USA* 2012, 109:594–599. doi:10.1073/pnas.1116053109.
57. Trinh CT, Wlaschin A, Sreenc F. Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism. *Appl Microbiol Biotechnol* 2009, 81:813–826. doi:10.1007/s00253-008-1770-1.

58. Handorf T, Ebenhöf O, Heinrich R. Expanding metabolic networks: scopes of compounds, robustness, and evolution. *J Mol Evol* 2005, 61:498–512. doi:10.1007/s00239-005-0027-1.
59. Borenstein E, Kupiec M, Feldman MW, Ruppin E. Large-scale reconstruction and phylogenetic analysis of metabolic environments. *Proc Natl Acad Sci USA* 2008, 105:14482–14487. doi:10.1073/pnas.0806162105.
60. Bartell JA, Yen P, Varga JJ, Goldberg JB, Papin JA. Comparative metabolic systems analysis of pathogenic Burkholderia. *J Bacteriol* 2013, 196:210–226. doi:10.1128/JB.00997-13.
61. Fang Y, Scheibe TD, Mahadevan R, Garg S, Long PE, Lovley DR. Direct coupling of a genome-scale microbial in silico model and a groundwater reactive transport model. *J Contam Hydrol* 2011, 122:96–103. doi:10.1016/j.jconhyd.2010.11.007.
62. Krauss M, Schaller S, Borchers S, Findeisen R, Lippert J, Kuepfer L. Integrating cellular metabolism into a multiscale whole-body model. *PLoS Comput Biol* 2012, 8:e1002750. doi:10.1371/journal.pcbi.1002750.
63. Martin FPJ, Sprenger N, Yap IKS, Wang Y, Bibiloni R, Rochat F, Rezzi S, Cherbut C, Kochhar S, Lindon JC, et al. Panorganismal gut microbiome-host metabolic crosstalk. *J Proteome Res* 2009, 8:2090–2105. doi:10.1021/pr801068x.
64. Karlsson FH, Nookaew I, Nielsen J. Metagenomic Data Utilization and Analysis (MEDUSA) and construction of a global gut microbial gene catalogue. *PLoS Comput Biol* 2014, 10:e1003706. doi:10.1371/journal.pcbi.1003706.
65. Lasken RS, McLean JS. Recent advances in genomic DNA sequencing of microbial species from single cells. *Nat Rev Genet* 2014, 15:577–584. doi:10.1038/nrg3785.
66. Woyke T, Xie G, Copeland A, González JM, Han C, Kiss H, Saw JH, Senin P, Yang C, Chatterji S, et al. Assembling the marine metagenome, one cell at a time. *PLoS One* 2009, 4:e5299. doi:10.1371/journal.pone.0005299.
67. Taniguchi Y, Choi PJ, Li G, Chen H, Babu M, Hearn J, Emili A, Xie XS. Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science (New York, NY)* 2010, 329:533–538. doi:10.2142/biophys.51.136.
68. Embree M, Nagarajan H, Movahedi N, Chitsaz H, Zengler K. Single-cell genome and metatranscriptome sequencing reveal metabolic interactions of an alkane-degrading methanogenic community. *ISME J* 2014, 8:757–767. doi:10.1038/ismej.2013.187.
69. Reichert-Schwillinsky F, Pin C, Dzieciol M, Wagner M, Hein I. Stress- and growth rate-related differences between plate count and real-time PCR data during growth of *Listeria monocytogenes*. *Appl Environ Microbiol* 2009, 75:2132–2138. doi:10.1128/AEM.01796-08.
70. Besmer MD, Weissbrodt DG, Kratochvil BE, Sigrist J, Weyland MS, Hammes F. The feasibility of automated online flow cytometry for In-situ monitoring of microbial dynamics in aquatic ecosystems. *Front Microbiol* 2014, 5:1–12. doi:10.3389/fmicb.2014.00265.
71. Hanly TJ, Urello M, Henson M. Dynamic flux balance modeling of *S. cerevisiae* and *E. coli* co-cultures for efficient consumption of glucose/xylose mixtures. *Appl Microbiol Biotechnol* 2012, 93:2529–2541. doi:10.1007/s00253-011-3628-1.
72. Sévin DC, Kuehne A, Zamboni N, Sauer U. Biological insights through nontargeted metabolomics. *Curr Opin Biotechnol* 2015, 34:1–8. doi:10.1016/j.copbio.2014.10.001.
73. Walker A, Pfitzner B, Neschen S, Kahle M, Harir M, Lucio M, Moritz F, Tziotis D, Witting M, Rothballer M, et al. Distinct signatures of host-microbial meta-metabolome and gut microbiome in two C57BL/6 strains under high-fat diet. *ISME J* 2014, 8:2380–2396. doi:10.1038/ismej.2014.79.
74. Kok MGM, Swann JR, Wilson ID, Somssen GW, de Jong GJ. Hydrophilic interaction chromatography-mass spectrometry for anionic metabolic profiling of urine from antibiotic-treated rats. *J Pharm Biomed Anal* 2014, 92:98–104. doi:10.1016/j.jpba.2014.01.008.
75. Gan XT, Ettinger G, Huang CX, Burton JP, Haist JV, Rajapurohitam V, Sidaway JE, Martin G, Gloor GB, Swann JR, et al. Probiotic administration attenuates myocardial hypertrophy and heart failure after myocardial infarction in the rat. *Circ Heart Fail* 2014, 7:491–499. doi:10.1161/CIRCHEARTFAILURE.113.000978.
76. Traxler MF, Watrous JD, Alexandrov T, Dorrestein PC, Kolter R. Interspecies interactions stimulate diversification of the *Streptomyces coelicolor* secreted metabolome. *mBio* 2013, 4:e00459-13. doi:10.1128/mBio.00459-13.
77. Barger SR, Hoefler BC, Cubillos-Ruiz A, Russell WK, Russell DH, Straight PD. Imaging secondary metabolism of *Streptomyces* sp. Mg1 during cellular lysis and colony degradation of competing *Bacillus subtilis*. *Anton van Leeuwenhoek Int J Gen Mol Microbiol* 2012, 102:435–445. doi:10.1007/s10482-012-9769-0.
78. Nath K, Kumar A, Das D. Hydrogen production by *Rhodobacter sphaeroides* strain O.U.001 using spent media of *Enterobacter cloacae* strain DM11. *Appl Microbiol Biotechnol* 2005, 68:533–541. doi:10.1007/s00253-005-1887-4.
79. Seta F, Hunter M, Larsen B. In vitro evaluation of small molecule inhibitors and probiotic byproducts on growth and viability of vaginal microorganisms. *Br J Med Med Res* 2014, 4:5779–5792. doi:10.9734/BJMMR/2014/12327.



80. Michelsen CF, Christensen A-MJ, Bojer MS, Hoiby N, Ingmer H, Jelsbak L. *Staphylococcus aureus* alters growth activity, autolysis, and antibiotic tolerance in a human host-adapted *Pseudomonas aeruginosa* lineage. *J Bacteriol* 2014, 196:3903–3911. doi:10.1128/JB.02006-14.
81. Hammer ND, Cassat JE, Noto MJ, Lojek LJ, Chadha AD, Schmitz JE, Creech CB, Skaar EP. Inter- and intraspecies metabolite exchange promotes virulence of antibiotic-resistant *Staphylococcus aureus*. *Cell Host Microbe* 2014, 16:531–537. doi:10.1016/j.chom.2014.09.002.
82. Nielsen HB, Almeida M, Juncker AS, Rasmussen S, Li J, Sunagawa S, Plichta DR, Gautier L, Pedersen AG, Le Chatelier E, et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nat Biotechnol* 2014, 32:822–828. doi:10.1038/nbt.2939.
83. Human Microbiome Jumpstart Reference Strains Consortium. A catalog of reference genomes from the human microbiome. *Science (New York, NY)* 2010, 328:994–999. doi:10.1126/science.1183605.
84. Albertsen M, Hugenholtz P, Skarshewski A, Nielsen KL, Tyson GW, Nielsen PH. Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nat Biotechnol* 2013, 31:533–538. doi:10.1038/nbt.2579.
85. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C. Binning metagenomic contigs by coverage and composition. *Nat Methods* 2014, 11:1144–1146. doi:10.1038/nmeth.3103.
86. Sharon I, Morowitz MJ, Thomas BC, Costello EK, Relman D, Banfield JF. Time series community genomics analysis reveals rapid shifts in bacterial species, strains, and phage during infant gut colonization. *Genome Res* 2012, 23:111–120. doi:10.1101/gr.142315.112.
87. Carr R, Shen-Orr SS, Borenstein E. Reconstructing the genomic content of microbiome taxa through shotgun metagenomic deconvolution. *PLoS Comput Biol* 2013, 9:e1003292. doi:10.1371/journal.pcbi.1003292.
88. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Loman NJ, Andersson AF, Quince C. CONCOCT: Clustering cONTigs on COverage and ComposiTiOn. *Arxiv preprint arXiv:13124038v1*; 2013;28.
89. Gori F, Mavroedis D, Jetten MSM, Marchiori E. Genomic signatures for metagenomic data analysis: Exploiting the reverse complementarity of tetranucleotides. In: *2011 IEEE International Conference on Systems Biology*, 2011:149–154. doi:10.1109/ISB.2011.6033147
90. Wu Y-W, Tang Y-H, Tringe SG, Simmons B, Singer SW. MaxBin: an automated binning method to recover individual genomes from metagenomes using an expectation-maximization algorithm. *Microbiome* 2014, 2:26. doi:10.1186/2049-2618-2-26.
91. Brady A, Salzberg SL. Phymm and PhymmBL: metagenomic phylogenetic classification with interpolated Markov models. *Nat Methods* 2009, 6:673–676. doi:10.1038/nmeth.1358.
92. Teeling H, Meyerdierks A, Bauer M, Amann R, Glöckner FO. Application of tetranucleotide frequencies for the assignment of genomic fragments. *Environ Microbiol* 2004, 6:938–947. doi:10.1111/j.1462-2920.2004.00624.x.
93. Kelley DR, Salzberg SL. Clustering metagenomic sequences with interpolated Markov models. *BMC Bioinformatics* 2010, 11:544. doi:10.1186/1471-2105-11-544.
94. Patil KR, Haider P, Pope PB, Turnbaugh PJ, Morrison M, Scheffer T, McHardy AC. Taxonomic metagenome sequence assignment with structured output models. *Nat Methods* 2011, 8:191–192. doi:10.1038/nmeth0311-191.
95. MacDonald NJ, Parks DH, Beiko RG. Rapid identification of high-confidence taxonomic assignments for metagenomic data. *Nucleic Acids Res* 2012, 40:1–13. doi:10.1093/nar/gks335.
96. Jiang H, An L, Lin SM, Feng G, Qiu Y. A statistical framework for accurate taxonomic assignment of metagenomic sequencing reads. *PLoS One* 2012, 7:e46450 doi:10.1371/journal.pone.0046450.
97. Orth JD, Conrad TM, Na J, Lerman JA, Nam H, Feist AM, Palsson BØ. A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism—2011. *Mol Syst Biol* 2011, 7:535. doi:10.1038/msb.2011.65.
98. Hamilton JJ, Reed JL. Software platforms to facilitate reconstructing genome-scale metabolic networks. *Environ Microbiol* 2014, 16:49–59. doi:10.1111/1462-2920.12312.
99. Saha R, Chowdhury A, Maranas CD. Recent advances in the reconstruction of metabolic models and integration of omics data. *Curr Opin Biotechnol* 2014, 29C:39–45. doi:10.1016/j.copbio.2014.02.011.
100. Alam MT, Medema MH, Takano E, Breitling R. Comparative genome-scale metabolic modeling of actinomycetes: the topology of essential core metabolism. *FEBS Lett* 2011, 585:2389–2394. doi:10.1016/j.febslet.2011.06.014.
101. Satish Kumar V, Dasika MS, Maranas CD. Optimization based automated curation of metabolic reconstructions. *BMC Bioinformatics* 2007, 8:212. doi:10.1186/1471-2105-8-212.
102. Kumar VS, Maranas CD. GrowMatch: An automated method for reconciling in silico/in vivo growth predictions. *PLoS Comput Biol* 2009, 5:e1000308. doi:10.1371/journal.pcbi.1000308.
103. Benedict MN, Mundy MB, Henry CS, Chia N, Price ND. Likelihood-based gene annotations for gap filling and quality assessment in genome-scale metabolic

- models. *PLoS Comput Biol* 2014, 10:e1003882. doi:10.1371/journal.pcbi.1003882.
104. Pitkänen E, Jouhten P, Hou J, Syed MF, Blomberg P, Kludas J, Oja M, Holm L, Penttilä M, Rousu J, et al. Comparative genome-scale reconstruction of gapless metabolic networks for present and ancestral species. *PLoS Comput Biol* 2014, 10:e1003465. doi:10.1371/journal.pcbi.1003465.
105. Pellegrini M, Marcotte EM, Thompson MJ, Eisenberg D, Yeates TO. Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. *Proc Natl Acad Sci USA* 1999, 96:4285–4288. doi:10.1073/pnas.96.8.4285.
106. Plata G, Henry CS, Vitkup D. Long-term phenotypic evolution of bacteria. *Nature* 2014, 517:369–372. doi:10.1038/nature13827.
107. Oberhardt MA, Puchalka J, dos Santos VAPM, Papin JA. Reconciliation of genome-scale metabolic reconstructions for comparative systems analysis. *PLoS Comput Biol* 2011, 7:e1001116. doi:10.1371/journal.pcbi.1001116.
108. Stecher B, Chaffron S, Käppeli R, Hapfelmeier S, Friedrich S, Weber TC, Kirundi J, Suar M, McCoy KD, Von Mering C, et al. Like will to like: abundances of closely related species can predict susceptibility to intestinal colonization by pathogenic and commensal bacteria. *PLoS Pathog* 2010, 6:e1000711. doi:10.1371/journal.ppat.1000711.
109. Ghosh A, Nilmeier J, Weaver D, Adams PD, Keasling JD, Mukhopadhyay A, Petzold CJ, Martín HG. A peptide-based method for <sup>13</sup>C Metabolic Flux Analysis in microbial communities. *PLoS Comput Biol* 2014, 10:e1003827. doi:10.1371/journal.pcbi.1003827.
110. Blazier AS, Papin JA. Integration of expression data in genome-scale metabolic network reconstructions. *Front Physiol* 2012, 3:299. doi:10.3389/fphys.2012.00299.
111. Becker S, Palsson BO. Context-specific metabolic networks are consistent with experiments. *PLoS Comput Biol* 2008, 4:e1000082. doi:10.1371/journal.pcbi.1000082.
112. Jensen P, Papin J. Functional integration of a metabolic network model and expression data without arbitrary thresholding. *Bioinformatics* 2011, 27:541–547. doi:10.1093/bioinformatics/btq702.
113. Namiki T, Hachiya T, Tanaka H, Sakakibara Y. MetaVelvet: an extension of velvet assembler to de novo metagenome assembly from short sequence reads. *Nucleic Acids Res* 2012, 40:1–12. doi:10.1093/nar/gks678.
114. Afiahayati, Sato K, Sakakibara Y. MetaVelvet-SL : an extension of Velvet assembler to de novo metagenomic assembler utilizing supervised learning. *DNA Res* 2015, 22:69–77. doi:10.1093/dnares/dsu041.
115. Ibarra RU, Edwards JS, Palsson BO. *Escherichia coli* K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* 2002, 420:186–189. doi:10.1038/nature01149.
116. Wintermute EH, Silver P. Emergent cooperation in microbial metabolism. *Mol Syst Biol* 2010, 6:407. doi:10.1038/msb.2010.66.
117. Waldor MK, Tyson G, Borenstein E, Ochman H, Moeller A, Finlay BB, Kong HH, Gordon JL, Nelson KE, Dabbagh K, et al. Where next for microbiome research? *PLoS Biol* 2015, 13:e1002050. doi:10.1371/journal.pbio.1002050.
118. Becker S, Feist AM, Mo ML, Hannum G, Palsson BO, Herrgard MJ. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA toolbox. *Nat Protoc* 2007, 2:727–738. doi:10.1038/nprot.2007.99.
119. Oberhardt M, Puchalka J, Fryer KE, Martins dos Santos VP, Papin J. Genome-scale metabolic network analysis of the opportunistic pathogen *Pseudomonas aeruginosa* PAO1. *J Bacteriol* 2008, 190:2790–2803. doi:10.1128/JB.01583-07.
120. Segrè D, Vitkup D, Church GM. Analysis of optimality in natural and perturbed metabolic networks. *Proc Natl Acad Sci USA* 2002, 99:15112–15117. doi:10.1073/pnas.232349399.
121. Burgard AP, Pharkya P, Maranas CD. Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol Bioeng* 2003, 84:647–657. doi:10.1002/bit.10803.
122. Patil KR, Rocha I, Förster J, Nielsen J. Evolutionary programming as a platform for in silico metabolic engineering. *BMC Bioinformatics* 2005, 6:308. doi:10.1186/1471-2105-6-308.
123. Rockwell G, Guido NJ, Church GM. Redirector: designing cell factories by reconstructing the metabolic objective. *PLoS Comput Biol* 2013, 9:e1002882. doi:10.1371/journal.pcbi.1002882.
124. Shong J, Jimenez Diaz MR, Collins CH. Towards synthetic microbial consortia for bioprocessing. *Curr Opin Biotechnol* 2012, 23:798–802. doi:10.1016/j.copbio.2012.02.001.