

REVIEW

Towards applications of genome-scale metabolic model-based approaches in designing synthetic microbial communities

Huan Du^{1,2}, Meng Li^{1,2}, Yang Liu^{1,2,*}

¹ Archaeal Biology Center, Institute for Advanced Study, Shenzhen University, Shenzhen 518060, China

² Shenzhen Key Laboratory of Marine Microbiome Engineering, Institute for Advanced Study, Shenzhen University, Shenzhen 518060, China

* Correspondence: yangliu@szu.edu.cn

Received July 30, 2022; Revised November 3, 2022; Accepted November 8, 2022

Background: Synthetic microbial communities, with different strains brought together by balancing their nutrition and promoting their interactions, demonstrate great advantages for exploring complex performance of communities and for further biotechnology applications. The potential of such microbial communities has not been explored, due to our limited knowledge of the extremely complex microbial interactions that are involved in designing and controlling effective and stable communities.

Results: Genome-scale metabolic models (GEM) have been demonstrated as an effective tool for predicting and guiding the investigation and design of microbial communities, since they can explicitly and efficiently predict the phenotype of organisms from their genotypic data and can be used to explore the molecular mechanisms of microbe-habitats and microbe-microbe interactions. In this work, we reviewed two main categories of GEM-based approaches and three uses related to design of synthetic microbial communities: predicting multi-species interactions, exploring environmental impacts on microbial phenotypes, and optimizing community-level performance.

Conclusions: Although at the infancy stage, GEM-based approaches exhibit an increasing scope of applications in designing synthetic microbial communities. Compared to other methods, especially the use of laboratory cultures, GEM-based approaches can greatly decrease the trial-and-error cost of various procedures for designing synthetic communities and improving their functionality, such as identifying community members, determining media composition, evaluating microbial interaction potential or selecting the best community configuration. Future efforts should be made to overcome the limitations of the approaches, ranging from quality control of GEM reconstructions to community-level modeling algorithms, so that more applications of GEMs in studying phenotypes of microbial communities can be expected.

Keywords: genome-scale metabolic modeling; microbial community design; interspecies interaction; environmental impact; community-level performance

Author summary: The applications of computational tools have been demonstrated to increase the development of synthetic microbial communities which is an emerging field and can be used in various biotechnology applications. As one effective tool, genome-scale metabolic modeling helps to reconstruct testable metabolic networks from genomic information and can quantitatively simulate entire metabolic fluxes of communities in considering microbe-microbe and microbe-habitat interactions. In-depth study of underlying mechanisms of microbial interactions using metabolic models and of coupling the models with multi-omics data or machine learning can further extend applications in designing synthetic communities.

INTRODUCTION

Synthetic microbial communities are the basis of an emerging research field in synthetic biology that aims at the in-detail study of the properties and functions of microbial communities and development of application of these functions in biotechnology. A synthetic microbial community is artificially built by co-culturing two or more species under controlled conditions [1]. The natural life mode of microorganisms in microbial communities inspires the idea of constructing synthetic communities. Microbes in multi-species consortia can form “metabolic modules” and accomplish complex metabolic processes via cooperation [2,3]. It may be challenging to observe such a phenomenon with a single strain. For example, when a single engineered *Escherichia coli* strain is used to generate the whole glutarate pathway, the accumulation of the intermediate 5-AMV limits the glutarate production. While applying a synthetic consortium composed of two engineered *E. coli* strains, which each contribute a part of the glutarate pathway, the inhibition can be totally removed, leading to a 19.2% of improvement in glutarate production [3]. Moreover, the diversity of metabolic capabilities possessed by the multiple species and the inter-species interactions enhance the stability and robustness of the community against environmental stresses and ecological invasion [4–6]. Multi-species consortia can use more kinds of substrates and reduce the inhibition of intermediate products by regulating the populations. Some experiments have also indicated that during the process of biofuel production, assemblages of algae are better than monocultures at resisting contaminations [7]. In recent years, synthetic microbial communities have revealed strong capabilities in various industrial and biotechnological applications, such as environmental remediation [8], chemical production [9], biofuel production [10,11], drug discovery [12], probiotic-mediated therapies [13,14].

The vast potential of microbial communities is far from harnessed, due to our limited knowledge and ability in rapid design of effective, stable, and robust microbial communities [15]. Among the important challenges for such design are the elusive microbial interactions within the communities [16,17]. Intercellular interactions are one of the key factors for shaping and maintaining community structure [18,19]. Many interactions occur simultaneously to the microbes, including competition for resources and the exchange of metabolites. The tradeoffs led by the interactions make the community composition tend to be stable. In addition, the dynamic variability in interaction patterns can also lead to high metabolic diversity of the

communities and hence make the communities more resistant and responsive to environmental perturbations. However, the microbial interactions are extremely complex and difficult to identify. This is due to the high species diversity and hence metabolic diversity of microbes, due to the enormous species of metabolites to be exchanged, due to the multiple promiscuous interactions among microbes, and due to the dynamic changes in interaction patterns in response to environmental conditions.

In this context, the genome-scale metabolic models (GEMs), which can simulate the metabolic flux distributions of organisms based on their genomic data, offer an effective tool for studying microbial metabolic interactions. Compared to other methods, especially the use of laboratory cultures, GEM-based approaches can explicitly and efficiently predict and study the underlying molecular mechanisms of the multi-species interactions as well as the performance of the whole microbial community and their metabolic network [20–22]. Moreover, the reactions of microbes to different environmental conditions can also be explored, using condition-specific GEMs. With these prediction capacities, the applications of GEM-based approaches can greatly decrease the trial-and-error cost in various procedures for designing synthetic communities and improving their functionality, overcoming challenges such as identifying community members, determining media composition, evaluating microbial interaction potential or selecting the best community configuration. They should thus be considered indispensable for research on synthetic communities. In the following sections, we first present the fundamentals and the reconstruction process of GEMs. Next, several applications of the GEM-based approaches for exploring the microbe-microbe, microbe-habitat interactions, and the community-level performance are presented. Finally, the uncertainties of GEM-based approaches and the future challenges of their applications in the research of synthetic microbial communities are discussed.

GENOME-SCALE METABOLIC-BASED APPROACHES

What are genome-scale metabolic models?

A GEM is a mathematical representation of the metabolic network of an organism, which quantitatively predicts its genotype-phenotype relationship. Using GEMs, a whole optimized set of directional metabolic reactions of a cell can be determined from its genomic data in simulating its entire metabolic flux under preset environmental conditions.

In the process of GEM reconstruction, the first step is

the annotation of the genome sequences using metabolic knowledge bases (Fig. 1A), such as the Kyoto Encyclopedia of Genes and Genomes (KEGG) [23], the MetaCyc [24] or the Biochemical, Genetic and Genomic (BiGG) [25] knowledge bases. The annotated genes are associated with their corresponding reactions with a reaction score calculated through gene-protein-reaction (GPR) rules. It determines if an adequate collection of proteins is present for catalyzing the reactions (Fig. 1A). By deciding the set of biochemical reactions that the organism of interest can carry out, its draft metabolic network can be reconstructed. This draft construction may contain gaps or inaccuracy due to missing or inaccurate gene annotations, which hence should be further revised.

For the model refinement, a biomass reaction [26] is added to this metabolic network to support the cell growth and to connect with the reactions that synthesize precursors for biomass formation (Fig. 1B). The biomass reaction involves major compounds that are

essential for growth of an organism. Ideally, the biomass formulation of a specific organism, such as *E. coli* [27] and *Methanosarcina barkeri* [28], should be directly determined by experimental measurements. Where there is lack of experimental data, then the biomass composition of template models should be used, such as use of *E. coli* for Gram-negative bacteria or use of *Bacillus subtilis* for Gram-positive bacteria [29]. Various GEM construction tools have also built template models for archaea, algae, fungi, plants and human cells [30–32]. In addition to the biomass formulation, a set of exchange reactions is defined to describe the flux of substrates uptake and end-products discharge. The new version of the metabolic network can be further converted into a stoichiometric matrix which compiles the stoichiometry information of all the reactions for mathematically linking the metabolites with their related reactions (Fig. 1C). The stoichiometric matrix offers analysis of the most important constraints on reaction flux to ensure mass and electron

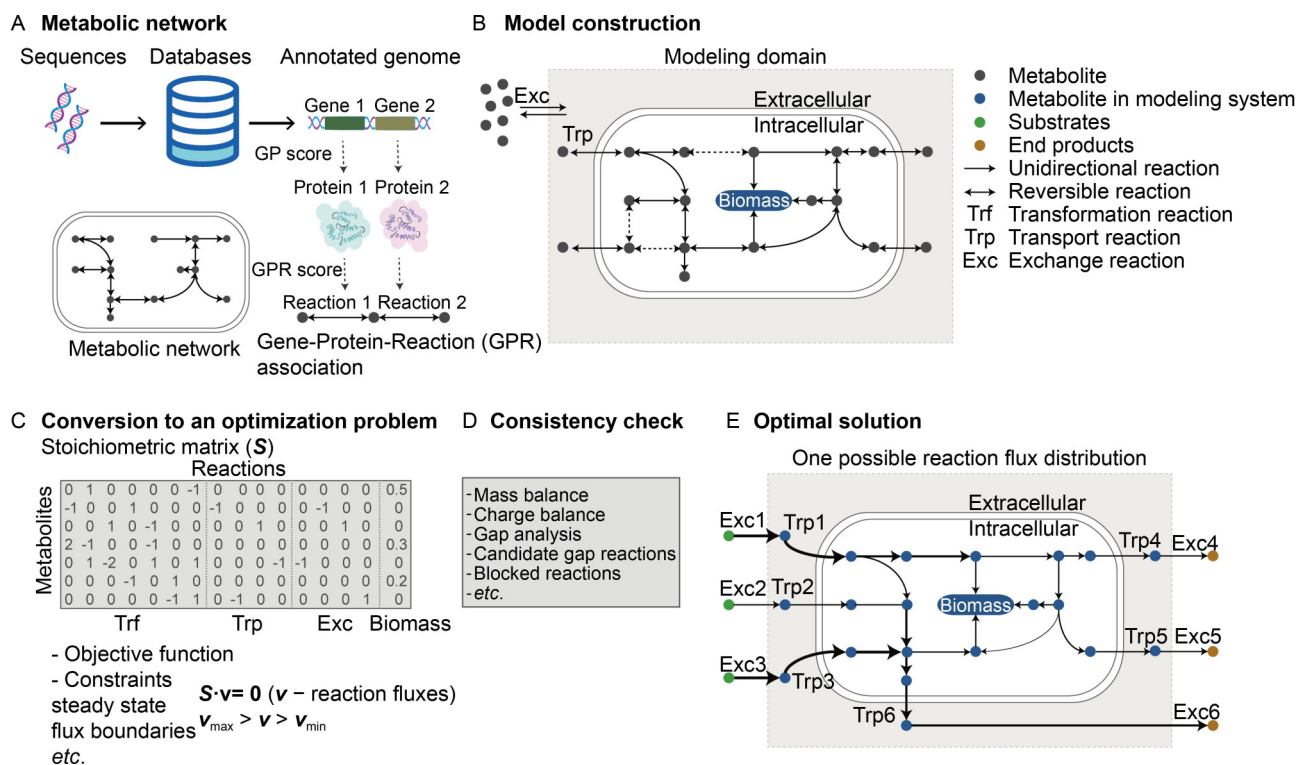


Figure 1. Basic process of GEM reconstructions. (A) Metabolic reactions of the organism of interest are derived from the annotated genome depending on the GPR associations. The metabolic network of this organism (schematic representation) is integrated from the set of metabolic reactions. (B) A biomass reaction is added to the metabolic network, and the boundary of the modeling system and the exchange fluxes with the environment are defined for subsequent model construction. (C) The model deals with an optimization problem with an appropriate objective function and several constraints on the reaction fluxes. The stoichiometric matrix converted from the metabolic network delineates the relation between the reactions and their related metabolites and is used to define the key constraint to ensure the steady state of the organism. (D) Consistency check of the metabolic network. (E) One or multiple possible flux distributions can be found by solving the optimization problem.

conservation. In addition, other constraints can be set according to thermodynamic feasibility [33] or other experimental results [34]. An objective function (*e.g.*, the growth rate, or the production rate of a specific metabolite) can be maximized or minimized using mathematical optimization techniques, such as flux balance analysis (FBA) [35], under these constraints. Subsequently, network evaluation is executed to check the consistency of the network, including the check of mass and charge balance, the check of blocked reactions, and the search of candidate reactions for gap filling (Fig. 1D). After the network evaluation, the generic GEM can be reconstructed, and the optimal solutions can be quantified to provide possible flux distributions for the metabolic network (Fig. 1E). The exchange fluxes describe the uptake and secretion rates of the organism. Generally, the more complex the constraints are, the more precise the reconstructions will be. The environmental conditions, whether as input or as constraints, also greatly affect the flux distributions of reactions in GEMs.

Genome-scale metabolic modeling tools

Since the first GEM of *Haemophilus influenzae* Rd was constructed in 1999 [36], genome-scale metabolic modeling has rapidly developed. Thousands of GEMs have been built for many kinds of cells, such as bacteria [37,38], archaea [39], yeasts [40], plants [41] and even human cells [42]. As the manual process for GEM construction is complex and laborious [26], many computational tools have been developed to make the procedures automatic or semi-automatic, such as the online tools, ModelSEED [31] and Kbase [43], the canonical modeling toolboxes like the COBRA toolbox [44], and other approaches such as Pathway Tools [45], CarveMe [30], AGORA [46], AGREDA [47], RAVEN [48] or Merlin [49]. These tools greatly accelerate the GEM construction process, leading to generation of an increasing number of GEMs [29]. Moreover, some tools can improve the phenotypic predictions of microbes by incorporating specific constraints, such as enzymatic constraints in GECKO [50] or thermodynamic constraints in PSAMM [51]. In addition, the condition-specific (*e.g.*, growth environment, life cycle or specific tissue) metabolisms of organisms can be also simulated based on the generic GEMs and on the experimental observations under specific conditions. These models can help to predict how the microbes allocate nutrients to maximize their growth rate, or their production rate of target chemicals, which further expands the application scope of the GEMs [52–55].

Community-level genome-scale metabolic modeling tools

Community-level GEM-based approaches have been developed by integrating multiple GEMs into one model framework and thereby solving a community-level optimization problem. One category of the approaches is the static modeling approach that connects species via exchange reactions and assumes a steady state for the whole community. This category can be classified into the lumped network-based approaches [56,57] that combine the metabolisms of all community members into one network (Fig. 2A), and the compartment-based approaches like OptCom [58], cFBA [59], SteadyCom [21], DOLMN [60], BioLEGO 2 [61] or SMETANA [62] in which each organism is modeled as a distinct compartment and metabolites exchange between these compartments are explicitly modeled (Fig. 2B). The compartment-based approaches can either solve a single objective for the community performance or integrate the species suboptimization simulations into the community-level optimization. In addition to studying the community-level performance via the interactions among organisms, the multi-level optimization can also describe the trade-offs between individual and community-level fitness criteria.

Alternatively, dynamic approaches like DMMM [63] or dOptCom [64] can explicitly model the temporal variability of microbial communities. This category couples the static compartment-based approaches with the differential equations that capture the dynamic variability of modeling components like biomass or metabolite concentrations. By adding the spatial features, some dynamic approaches, including COMETS [65], BacArena [20], IndiMeSH [66], CODY [67] *etc.*, can also predict the spatial heterogeneity of microbial communities (Fig. 2C). They hence meet the requirements of the spatio-temporal control of engineered communities in a structured environment, which cannot be done by static methods.

APPLICATIONS OF GEM-BASED APPROACHES IN DESIGNING MICROBIAL COMMUNITIES

GEMs have been developed to study, predict and help to engineer the metabolisms of individual microbes and microbial communities, leading to various direct and indirect applications in systems and synthetic biology [68]. Relevant applications of GEMs guiding design of synthetic microbial communities can be divided into three parts (Fig. 3A). Firstly, the microbe-microbe interactions (positive, neutral, or negative types) should be analyzed. Secondly, it is necessary to consider how

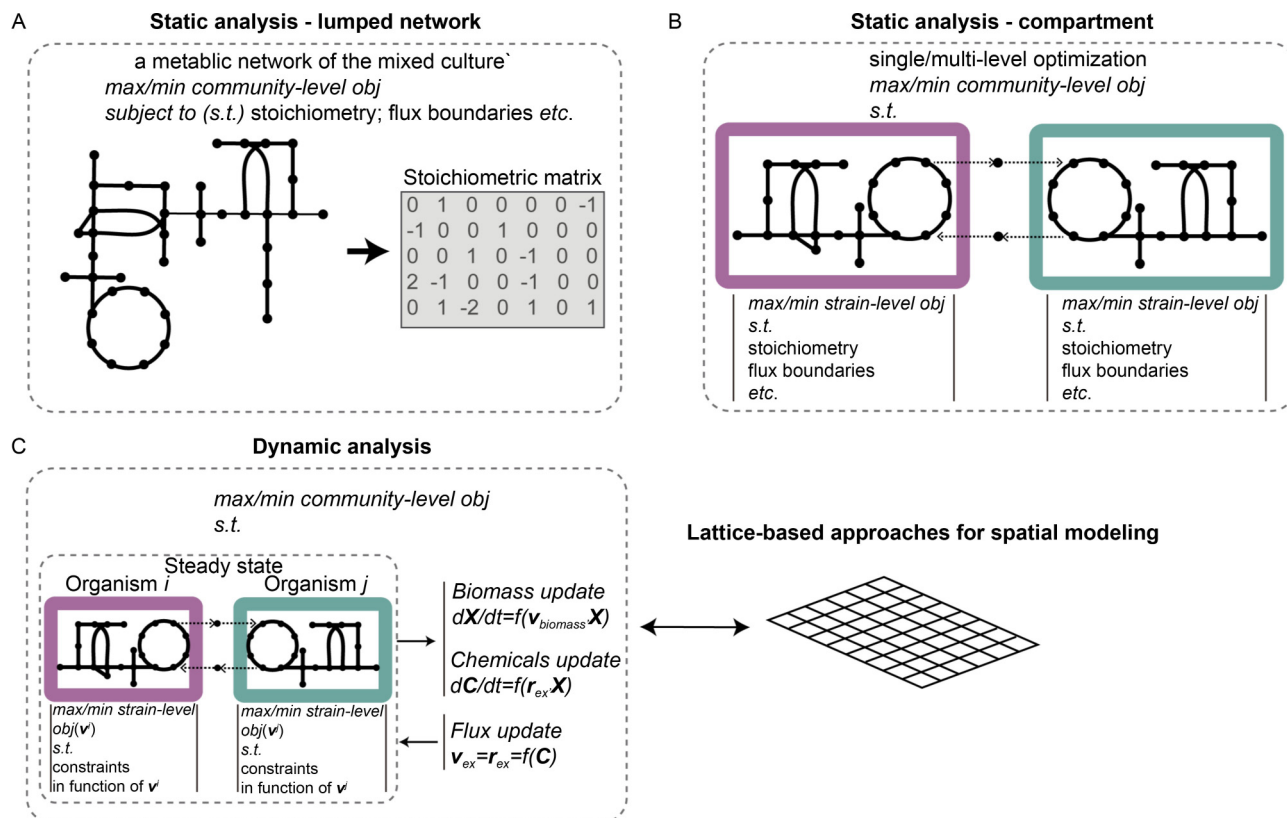


Figure 2. Different categories of GEM-based modeling tools. (A) Modeling of the community as a single entity can be achieved by lumped network-based static analysis that integrates the mixed culture into a unique metabolic network and then converts the network into a whole stoichiometric matrix. (B) Compartment-based static analysis can model each organism as a distinct compartment in considering the metabolites exchange among them. Multi-level optimization strategies are executed by adding the sub-optimization problem for each compartment based on a defined community-level objective. (C) Dynamic analysis couples the steady-state modeling with differential equations that capture the temporal variability of community performance. By integrating the dynamic models into a lattice-based framework, the spatial organization of the community can also be simulated. $\mathbf{X}$, vector denoting the biomass concentration of all the modeled species; $\mathbf{C}$, vector representing all the metabolites concentration; $\mathbf{v}^i$, a subset of all the reaction fluxes of organism i ; $\mathbf{v}_{\text{biomass}}$, vector of biomass reaction fluxes of all the modeled species; $\mathbf{v}_{\text{ex}}$, a subset of exchange reaction fluxes of all the modeled species; $\mathbf{r}_{\text{ex}}$, a subset of substrates uptake rate of all the modeled species; t , time.

environmental fluctuations affect microbial phenotypes and further change microbial interaction patterns. Thirdly, based on the two former studies, the community-level performance can be predicted and optimized. The GEM-based approaches mentioned in the three aspects for designing synthetic microbial communities are summarized in Table 1.

Studying microbe-microbe interactions

Modeling metabolic interactions among microbes is one of the most important applications of GEMs. Using GEM-based approaches, it has been possible to explore underlying mechanisms and predict previously unidentified interactions. The first multispecies GEM was constructed for studying the interaction of *Desulfovibrio*

vulgaris and *Methanococcus maripaludis* via a compartmentalized FBA-based model in which the objective was maximizing a weighted sum of the two-species biomass production fluxes [69]. The model suggested the essentiality of hydrogen transfer for syntrophic growth and accurately predicted the relative cell proportion of the two species during growth. Unlike that approach [69], the FBA framework named OptCom relied on a multi-level optimization formulation [58]. This approach considered a separate FBA problem for each species as sub-optimization problems. It integrated them through the constraints on metabolite exchanges and an outer-level objective of maximizing community biomass production fluxes. Using OptCom, the metabolic interactions between two abundant species in the human gut, *Bifidobacterium adolescentis* and

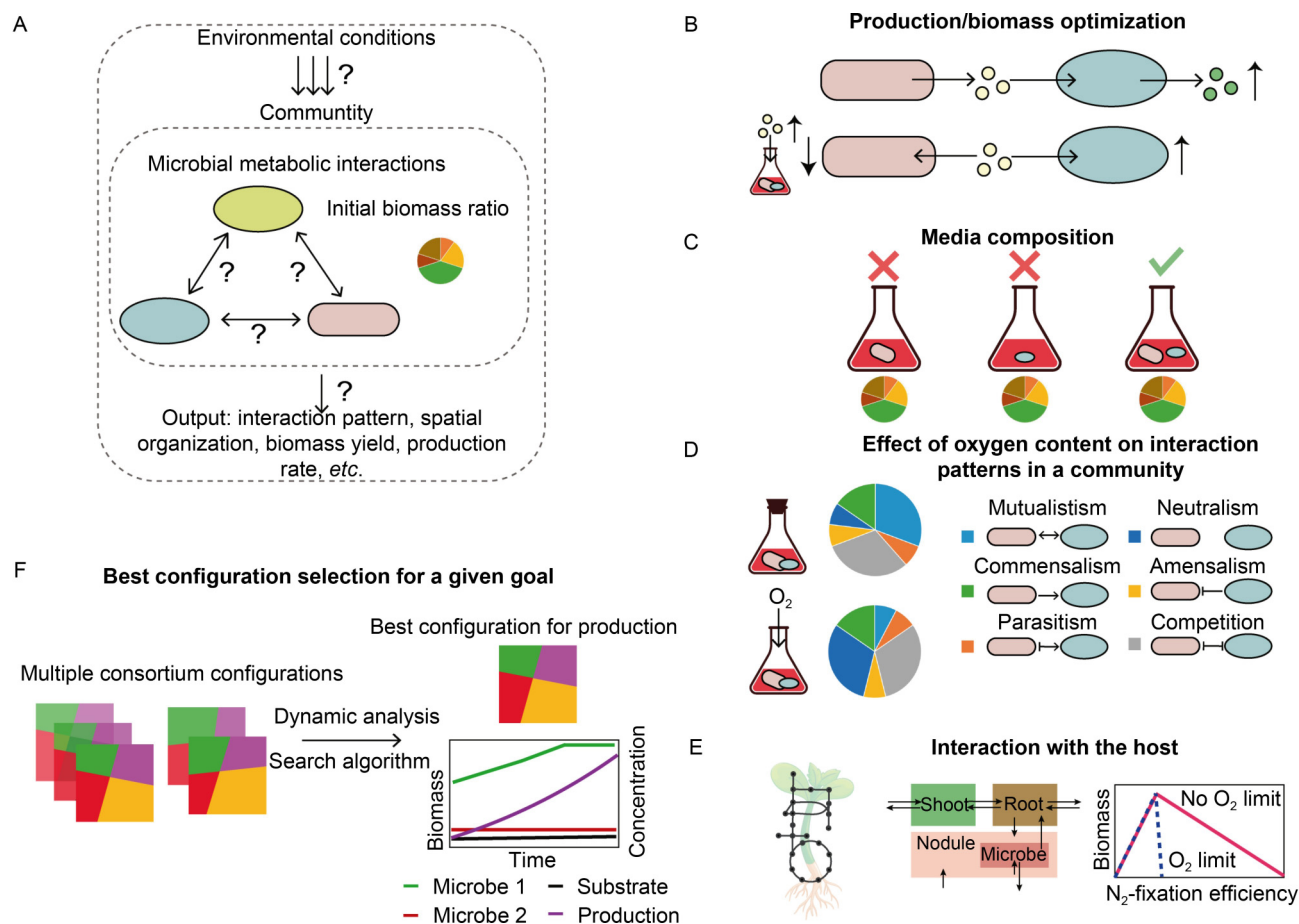


Figure 3. Examples of GEM applications guiding design of synthetic microbial communities. (A) Three parts of the GEM applications in studying synthetic microbial communities: microbe-microbe interactions, environmental impacts, and community-level performance. (B) Biomass/production optimization by microbial interactions. (C) Optimization of media composition for inducing syntrophic interactions of microbes. (D) Prediction of the effect of oxygen content on all pairwise interactions in a community. (E) The effects of the activity of microbes in the nodules on the biomass production rate of the host plant. (F) Best consortium configuration selection among multiple configurations (including media composition, initial strain ratio, environmental conditions, etc.) for a given goal by integrating a dynamic GEM-based approach with a search algorithm.

Faecalibacterium prausnitzii, were explored [70]. The modeling results indicated that the butyrate produced by *F. prausnitzii*, which is essential for colonic homeostasis and cancer prevention, can be promoted by the acetate supplied by *B. adolescentis* (Fig. 2B). In addition to modeling the positive interactions, other interactions (e.g., competition, parasitism, etc.) could be evaluated using GEM-based approaches [58,62,71,72]. For example, an approach (SteadyCom) to identifying the competitive relation among microbes was to simulate the relative abundances of the community members and search for negatively correlated pairs while requiring constant community growth rate [21]. The study used four mutant *E. coli* strains for simulations and succeeded in identifying the competitive pairs in which both the strains relied on lysine and methionine.

In addition to analysis of interaction patterns, GEMs

have been used to explore the underlying mechanisms for generating interactions among species, which are difficult to assess experimentally and can inform the design of synthetic microbial communities. The impact of costless metabolic secretions was evaluated by performing over 2 million pairwise growth simulations of 24 species in different media [73]. The costless metabolic exchange was indicated to be a driver of beneficial interactions contributing to the better growth of microbes in resource-poor environments. Anoxic conditions can provide more opportunities for costless metabolic exchanges and more stable ecological network motifs. Another approach named SMETANA assessed the extent of resource competition and metabolic exchanges among microbes via computing the substrate overlap and the essential exchanged metabolites [62]. The simulations for over 800

Table 1 Summary of GEM-based approaches which can be applied in synthetic community researches

Classification	Method	Short description
Community-level, static, lumped network-based	Rodríguez <i>et al.</i> 2006 [56]	A model to predict product formation from glucose in anaerobic mixed culture fermentation through maximizing a community-level biomass objective.
	Pramanik <i>et al.</i> 1999 [57]	A model to explore biological phosphorus removal metabolism.
Community-level, static, compartment-based	OptCom [58]	An FBA-based framework to describe trade-offs between individual and community-level fitness criteria by optimizing multi-level objectives.
	cFBA [59]	A method to analyze community parameters (maximal growth rate, relative biomass abundance, <i>etc.</i>) at balanced growth.
	SteadyCom [21]	A framework reformulated from cFBA without the limitations on the number of linear programming iterations for predicting the variation in species abundance in response to substrate changes.
	DOLMN [60]	A mixed integer linear programming (MILP) optimization approach to explore possible labor division in communities under constraints (<i>e.g.</i> , limited number of exchange reactions).
	BioLEGO 2 [61]	A Microsoft Azure Cloud-based framework which supports large-scale simulations of biomass serial fermentation processes by two different organisms with single or multiple gene knockouts.
	SMETANA [62]	A tool to estimate pairwise and community-level microbial interaction potential (through SMETANA score) and identify likely exchanged metabolites.
	Stolyar <i>et al.</i> 2007 [69]	The first multi-species GEM to predict community-level fluxes and the ratio of cells.
	The microbiome modeling toolbox [71]	A COBRA-based toolbox to study various types of pairwise microbe-microbe, microbe-host interactions and, to analyze personalized gut microbial communities under different diets.
	MMinte [72]	A methodology to assess pairwise microbial metabolic interactions ends the effect of these interactions on the relative growth rates of microbes from 16S rRNA data.
	Klitgord and Segrè 2010 [76]	A model to identify media that can induce putative symbiotic interactions.
	ViNE [81]	An FBA-based model for analyzing the integrated metabolism of the holobiont consisting of a host plant and its symbiotic bacterium.
	MICOM [86]	A framework for predicting growth rates of diverse bacterial species in human gut and metabolic fluxes of communities by using a heuristic optimization approach based on L2 regularization.
	CASINO [89]	A toolbox for modeling diet-microbiota interactions.
	Zampieri and Sauer 2016 [94]	A mixed-integer bi-level linear programming to infer an optimal combination of nutrients for sustaining pairwise, synergistic growth of microbes with minimum cost of cross-fed metabolites.
Community-level, dynamic, temporal	DMMM [63]	The first method using dFBA at community level to optimize growth rates of each strain within the community.
	dOptCom [64]	A method extended from OptCom for the dynamic metabolic modeling of microbial communities with multi-level objectives.
Community-level, dynamic, spatio-temporal	COMETS [65]	A platform implementing a dFBA algorithm on a lattice to track the spatio-temporal biomass distribution and fluxes of a multi-species community at population level.
	BacArena [20]	An R package integrating dFBA with individual-based approach to generate spatial organization and metabolic phenotype in biofilms over time.
	IndiMeSH [66]	A model combined dFBA with individual-based approach in an angular pore network for spatial modeling of soil aggregates in considering the impact of habitat geometry and hydration conditions.
	CODY [67]	A multi-scale framework to identify and quantify spatiotemporal-specific variations of gut microbiome abundance profiles in the colon as impacted by host physiology.
	FLYCOP [100]	A framework combining COMETS with a local search algorithm to automatically select the best consortium configuration among multiple predefined/random ones for a given goal.
Individual level, integration with macromolecular expression	FoldME [78]	A metabolism and protein expression (ME) model incorporating folding and degradation kinetics to predict the effect of temperature on microbial growth.
	OxidizeME [79]	An ME model to describe the response of microbes to reactive oxygen species stress.
	AcidifyME [80]	An ME model integrating folding and unfolding thermodynamics and kinetics to simulate the response of microbes to pH variations.

communities revealed that competition among microbes was apparent in all communities and indicated a

significant negative correlation between the competition extent and phylogenetic relatedness of the member

species. In addition, no prominent association of co-occurrence with resource competition was observed. But the metabolic interactions were recorded, leading to the conclusion that metabolic interdependency is a major driver of species co-occurrence. In turn, this point of view also provided support for mutualistic system constructions.

Evaluating microbe-habitat/host interactions

Microbial phenotypes (such as metabolic secretions, growth rate) are significantly related to the environmental conditions, which further affects the interspecies interactions as well as the composition, the stability and even the functions of the whole community. Thus, understanding of the culture conditions, of the potential environmental fluctuations, and of the interactions of microbes with the host are essential for studying and designing synthetic microbial communities. Apart from the interactions among microbes described above, the GEM-based approaches can also be employed for modeling the phenotypic diversity of microbes – ranging from growth rate and substrate uptake rate to gene expression levels – and hence for predicting diverse interspecies interaction patterns under different environmental conditions [74,75].

In a study, the dynamic multispecies metabolic modeling (DMMM) has been employed to investigate the effect of substrate concentrations on the interactions between *Rhodoferrax* and *Geobacter* species, which are both acetate-oxidizing Fe(III)-reducers found in uranium-contaminated groundwater [63]. The model predicted that the high acetate and low ammonium concentrations would increase the ratio of *Geobacter* to *Rhodoferrax*, the former of which can help to reduce the uranium in the environment (Fig. 3B). This result can provide support for designing strategies for bioremediation of uranium-contaminated groundwater. Such effort has also been extended to search the media compositions that sustain a co-culture of two species but do not support the growth of each organism on its own [76]. The results showed that specific media compositions could induce different types of putative symbiotic interactions (Fig. 3C). Environmental fluctuations may be more effective than genetic modifications for inducing symbiotic interactions. It further highlighted the crucial effects of environmental conditions on the generation of symbiotic interactions.

Using a GEM-based approach, oxygen availability was also found to be able to change microbial interaction patterns. Heinken and Thiele [77] used GEMs for 11 representative gut microbes to model pairwise interactions under anoxic and normoxic conditions (Fig. 3D). The mutualistic behaviors of the probiotic

organism *Lactobacillus plantarum* towards six other species under anoxic conditions were found to be entirely abolished under normoxic conditions. Furthermore, by incorporating transcription, translation, and stress response mechanisms into GEMs, the metabolism, the proteomic allocation, and the protein folding rates can be modeled. This enabled modelling of cellular behaviors in more detail and led to investigation of the responses of microbes to other environmental perturbations, such as the thermal [78], oxidative [79], and low-pH stress [80].

The interaction between the microbe and the host is also an important topic that can be studied using GEMs. For example, a model called ViNE integrated the host *Medicago truncatula* (plant) and its symbiotic bacterium *Sinorhizobium meliloti* into a three-tissue (shoot, root, and nodule) framework to study their association patterns [81] (Fig. 3E). The analysis revealed diminishing returns in terms of plant growth when the nitrogen fixation efficiency or the nodulation rate of the bacteria was beyond the optimum, which may have implications for engineering symbiotic nitrogen fixation. Another research direction concerns the metabolic interplay between the host and the gut microbiome, which has been demonstrated to be clearly associated with human health and diseases [82–84]. GEM-based approaches were developed to study the effect of the microbiota on the host and also the impact of diet on the gut microbiome [85–89]. One of the common toolboxes is CASINO, used in a diet-intervention study of 45 obese human individuals [89]. That study estimated the metabolic capabilities of the gut microbes and successfully predicted a significant change in the levels of some short-chain fatty acids and amino acids in response to the dietary intervention. Since abundant GEMs have been systematically constructed specifically for study of the members of the gut microbiome [46,47] and human cells [90–93], we believe that the study on the metabolic exchanges between the microbes, lumen and human cells can be further improved.

Community-level performance: design and optimization

Apart from exploring the underlying mechanisms of microbial interactions, several GEM-based approaches can be employed to model the performance of the whole microbial community and even design or optimize synthetic microbial communities. Here, we provide some examples of applications that use both static and dynamic approaches.

The static approaches can be applied in designing synthetic communities, in various ways. For instance, OptCom can assess the level of sub-optimal growth in

microbial communities [58]; SteadyCom focuses on predicting the variation in species abundance in response to substrate changes [21]; several ad-hoc approaches can optimize medium composition to induce microbial interactions [76,94]; SMETANA is used to evaluate the extent of resource competition and metabolic interaction potential of a whole community [62]. In particular, in a study integrating SMETANA with a network analysis method, a “social” network for a community was constructed based on the pairwise interaction potential of all the community members in mangrove sediments [95]. According to the network analysis and the transcriptomic data, several microbial active functional modules (mAFMs) were extracted from the network as the core modules. The microbes possess relatively high metabolic interactions and can actively realize certain dominant functions in element transformations via cooperation. These mAFMs represent the sub-consortia composed of microbes that are highly associated through their positive interactions, their simultaneously high-level transcriptional activity, and their spatial clustering. They hence could provide clues for synthetic community compositions. In addition, DOLMN has been applied to simulate the trade-off between metabolic self-reliance and mutualistic exchange and to further optimize the strategies for metabolic division of labor in ways that would be difficult to identify manually [60]. The simulations for consortia combined with diverse strains of *E. coli* indicated the nuanced and nonintuitive division of labor, like splitting the tricarboxylic acid (TCA) cycle into two separate halves.

Alternatively, dynamic approaches are more suitable for modeling the impact of the spatial heterogeneity of the media distribution or structured environments on microbial communities. By implementing a dynamic FBA algorithm [96] on a lattice, COMETS realized the simulation of the spatial and temporal diffusion of microbial populations and metabolites [65]. By comparing predictions with the experimental results, this model was verified to be able to precisely predict the impact of a competitor on the growth of a two-species consortium and the spatial distribution of the metabolites’ concentrations. Approaches subsequent to COMETS have made various changes. For example, BacArena, which incorporated GEMs into an agent-based approach, described the individual cells in more details by modeling heterogeneous phenotypic behavior, like cell movement, replication, or cellular lysis [20]; IndiMeSH adapted the model to study microbial dispersion and nutrient diffusion in more complex habitats such as soil, including pore spaces and aqueous phase configurations [66]. Another approach termed CODY [67] was constructed based on elementary flux

mode analysis instead of dynamic FBA. CODY focused on modeling gut microbiota and hence integrated three multiscale modeling components, *i.e.*, species-level microbial dynamics, microbial interactions, and colon physiology. This framework has enabled spatiotemporal predictions of microbial variations in response to diet intervention. In addition, some dynamic analyses have been used to optimize desired community-level functions, such as maximizing ethanol production with *S. cerevisiae* and *E. coli* [97,98] or to improve bioprocessing of cellulose with a clostridial consortium [99].

Going further than the tools for investigating some of the factors involved in constructing synthetic communities, a framework termed FLYCOP was developed to directly engineer microbial communities [100]. This framework combined COMETS with a local search algorithm, rather than tuning each control point for designing synthetic communities one by one, automatically selected the best consortium configuration among multiple predefined or random alternatives for a given goal (Fig. 3F). FLYCOP could realize various applications, like detecting limiting nutrients, optimizing cross-feeding relationships, optimizing strain ratios and pathway fragmentation, identifying aerobic-anaerobic switching time. One example was to optimize a *Synechococcus elongatus*-*Pseudomonas putida* consortium to produce the maximum amount of bioplastic. The simulations proposed the best configuration parameters related to initial low NH_4 concentration and high *S. elongatus* biomass ratio.

PERSPECTIVE AND FUTURE DIRECTIONS

Since the 1990s, the GEMs and the GEM-based modeling approaches have achieved rapid progress and have been applied in various fields [29,101]. Notably, GEM-based approaches have been demonstrated to be indispensable for studying systems and synthetic biology, due to their capacity to predict the genotype-phenotype relationships of organisms [102]. Engineering individual strains, like targets prediction for gene manipulation, has achieved great successes in various applications with the aid of GEM-based approaches [103–106]. Compared with individual cell studies, although in the infancy stage, a growing number of GEM-based approaches belonging to two categories, the community-level steady-state analysis and dynamic framework for spatio-temporal predictions, have been developed to explore the emergent properties of microbial communities. These GEM-based frameworks can contribute to giving mechanistic insights into community-level complexity as well as to improving the design of synthetic microbial communities that are

laborious and even impossible to study by way of laboratory experiments. Such beneficial uses include exploring the core modules of a complex community [95] or optimizing the media composition by testing a large number of component combinations [76].

However, the uncertain precision of GEMs and the limited capacity for community-level simulations with high-species diversity of current GEM-based approaches hinder the scope and precision of their applications in the field of synthetic microbial communities (Fig. 4B). Basically, the genome sequences, the gene annotations, the biomass formulation and the constraints on metabolic fluxes, which can be obtained from experiments and/or databases, are essential for the construction of GEMs and GEM-based community-level models (Fig. 4A). The uncertainties that emerge from these data, during the different steps of the GEM reconstruction process [107], such as incorrect/missing gene annotations, lack of specific biomass formulation, unknown media uptake rates, or biased flux simulation, fundamentally limit the precision of the GEM reconstructions and hence affect the community-level predictions. For example, after swapping biomass formulations between five different bacterial GEMs, considerable changes (up to 32.8%) have been observed in essentiality predictions of reactions, indicating the great impact of the biomass formulation selection on the prediction capacity of GEMs [108]. The precision of GEMs is one of the reasons why the application examples of the most GEM-based approaches have been executed with communities formed by precisely curated GEMs such as those reconstructed for *E. coli* strains. Thus, the first important task for modeling microbial communities is to certify the quality of GEMs. Recent research has proposed a tool termed MEMOTE for benchmarking GEMs from annotation and basic tests for model components like GPR rules, biomass reaction, and stoichiometry, which significantly contribute to standardized quality control of GEMs [109].

In addition, future efforts should be made to reduce the uncertainties in the GEM reconstruction process and improve GEM precision directly (Fig. 4C). It is firstly essential to utilize efficient methods for providing high-quality genomes to provide the basis for high-quality GEM reconstructions, such as long-read sequencing technologies [110] or droplet-based microfluidics methods [111]. The quality of the assembled genomes can greatly affect the accuracy of the corresponding GEM reconstruction, especially while modeling complex communities as in the gut or soil microbiome. The biomass formulation of a given organism should be estimated more accurately, for instance by lab-culture measurement or by using the software BOFdat [112]. Moreover, some other experimental observations, such

as chemical composition of the microbial habitats, substrate uptake/product secretion rate, multi-level (individual/community) growth rate, species abundance ratio or gene knock-out essays, can either be directly integrated as flux constraints of the related reactions in GEM reconstructions or be used to validate the GEMs [37,113,114]. Integrating multi-omics data, like transcriptome, proteome, or metabolome, into GEMs can also propose more constraints via setting thresholds related to gene/protein expression level or enzymatic activities, to reduce the flux variabilities. Nevertheless, the lack of kinetic information and the high computational demands of this framework result in the challenge for their applications, especially for multi-species modeling. In this context, recent studies have combined machine learning (ML) methods with GEMs to improve the prediction precision of the genotype-phenotype relationship with low computational costs. The ML methods can, on the one hand, be applied to decrease the uncertainties in GEM reconstruction processes. For example, the ML systems can be trained to improve gene annotation precision with the gold standard dataset covering more than 1 million protein sequences and their EC numbers [115]. The important reactions for further manual curation can also be identified by an ML system trained with an ensemble of GEMs generated from a draft GEM by iterative gap-filling [116]. An ML system trained with experimental data consisting of the phenotypic outcomes from single knockout mutants can accurately predict the essentiality of reactions [117]. On the other hand, ML systems can integrate the fluxomic data, generated using GEMs under different growth conditions, with other omics data to reversely improve the prediction power of GEMs, such as in the assessment/improvement of microbial growth and bio-production [118–120], in the exploration of antibiotic efficacy [121], or in the prediction of drug targets [122].

Even if the quality of each single-strain GEM can be accurately reconstructed, particular challenges will still hamper progress in community-level metabolic modeling. Current modeling tools, except the lumped network-based approaches, have not been used to simulate complex communities with high species diversity due to the combinatorial complexity of the multi-level optimizations. For this task, an ML random forest method combined with a dynamic GEM-based approach has classified the interactions and globally predicted the highly interaction-related metabolites for a 100-species gut microbiome [123] (Fig. 4C). Another study trained support vector machine models with over 2 million GEM-generated pairwise simulations to quantify the impact of several conditional variables, such as oxygen availability, species identity and carbon source types, on the secretion of costless metabolites which may promote

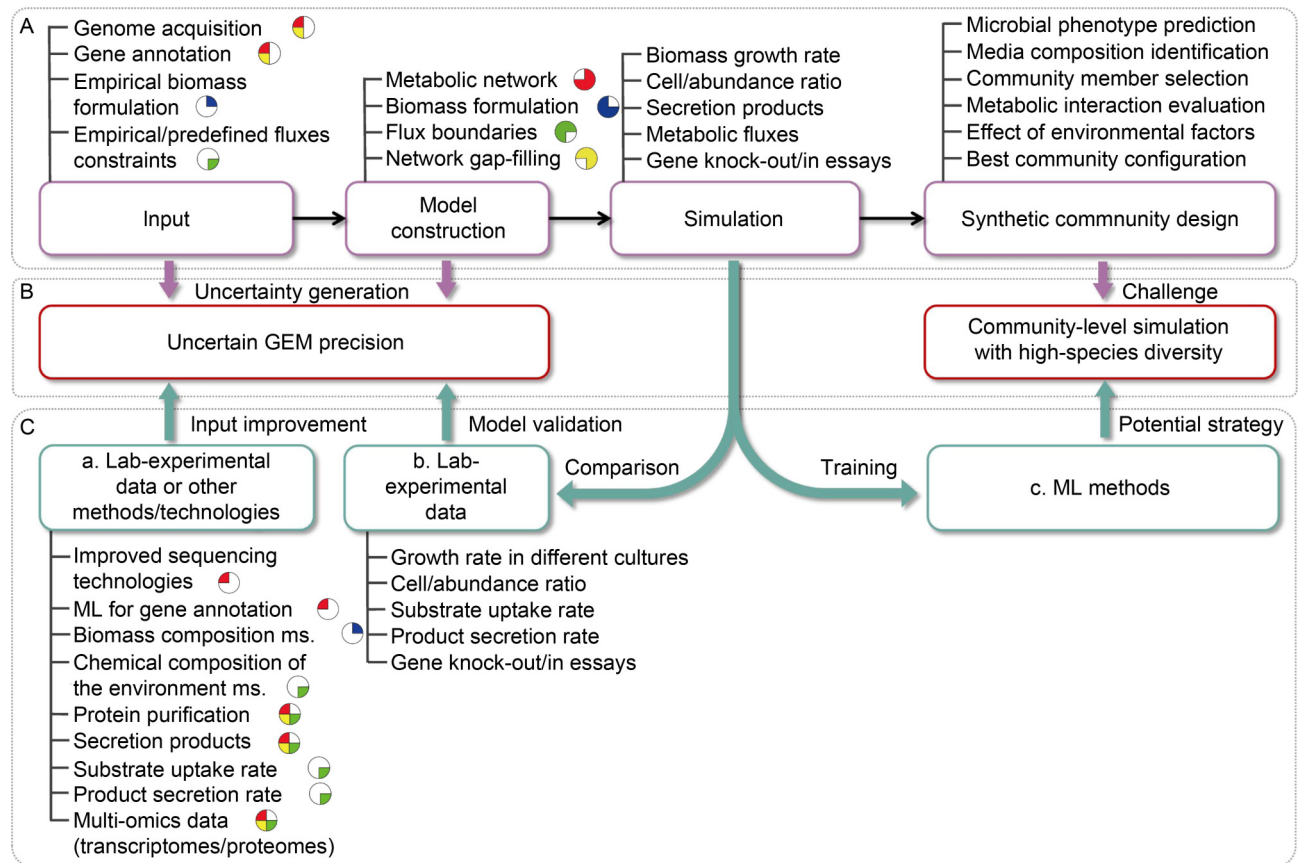


Figure 4. Limitations (B) hindering the application of GEM-based approaches in synthetic community design (A) and the potential improvement strategies (C). (A) An outline summary describing the necessary input data to the corresponding model construction processes (with correspondence represented in the pie chart by the same color), the possible simulated results and the potential applications of GEM-based approaches in synthetic community design. (B) The bottlenecks are mainly caused by the uncertainties generated during GEM constructions and the challenge of community-level simulation with high-species diversity. (C) The potential strategies for improving GEM precision (a, b) and for enabling high-species simulations (c). a. The data obtained from wet lab-experiments, machine learning (ML) systems and other methods/technologies can be directly integrated into different processes of GEM construction (with correspondence of the improvement, based on data, to the model construction processes being represented in the pie chart by the same color); b. The GEMs can be validated by comparing the simulated results to the wet lab-experimental observations; c. Some ML systems trained with the simulated results of GEM-based approaches may explore the underlying interaction mechanisms of the community with high-species diversity for synthetic community design. ML, machine learning; ms., measurement.

inter-microbial interactions [73]. Nevertheless, these frameworks focused on the performance of microbial interactions but cannot directly model the growth and metabolites secretions of a whole community. More work is needed in the future for direct community-level metabolic modeling for large communities. In addition, it would also be interesting to extend the application of GEM-based tools to the large space-scale or even the three-dimensional organization of microbial communities for modeling microbial aggregations like biofilms for chemical production, granules for wastewater treatment, or other microbial growth in structured environments. Ultimately, this paper shows that tools that can systematically engineer microbial communities

are still scarce. Different strategies have distinct focuses regarding environment specification [67], implementation conditions, and outcome results [124]. Thus, while performing synthetic consortia modeling, attention should be paid to selecting an appropriate GEM-based tool depending on the modeling purpose, the assumption consistency, and the available data.

Overall, the GEM-based approaches can guide the design of synthetic microbial communities in various ways, including by optimizing community composition, media composition, culture conditions, microbial interactions, and community-level perturbations under host/habitat condition change. With the advance of omics-data techniques and the emergent strength of

integrating multiple computational methods like GEMs with machine learning, GEM-based approaches exhibit an extending scope of applications. However, future efforts should be made to overcome the limitations so that more applications of GEMs in studying microbial interactions can be expected.

ACKNOWLEDGEMENTS

This work was supported by grants from the National Natural Science Foundation of China (Nos. 92051102, 32200099, 32225003 and 31970105); the Innovation Team Project of Universities in Guangdong Province (No. 2020KCXTD023); and the Shenzhen Science and Technology Program (JCYJ20200109105010363). All sketches used in the Figs. 1 and 3 are obtained from vecteezy website.

COMPLIANCE WITH ETHICS GUIDELINES

The authors Huan Du, Meng Li, Yang Liu declare that they have no conflict of interest or financial conflicts to disclose.

This review does not contain any studies with human or animal subjects performed by any of the authors.

OPEN ACCESS

This article is licensed by the CC By under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

REFERENCES

- Großkopf, T. and Soyer, O. S. (2014) Synthetic microbial communities. *Curr. Opin. Microbiol.*, 18, 72–77
- Zhou, K., Qiao, K., Edgar, S. and Stephanopoulos, G. (2015) Distributing a metabolic pathway among a microbial consortium enhances production of natural products. *Nat. Biotechnol.*, 33, 377–383
- Wang, X., Su, R., Chen, K., Xu, S., Feng, J. and Ouyang, P. (2019) Engineering a microbial consortium based whole-cell system for efficient production of glutarate from L-lysine. *Front. Microbiol.*, 10, 341
- Hays, S. G., Patrick, W. G., Ziesack, M., Oxman, N. and Silver, P. A. (2015) Better together: engineering and application of microbial symbioses. *Curr. Opin. Biotechnol.*, 36, 40–49
- McCarty, N. S. and Ledesma-Amaro, R. (2019) Synthetic biology tools to engineer microbial communities for biotechnology. *Trends Biotechnol.*, 37, 181–197
- Beyter, D., Tang, P. Z., Becker, S., Hoang, T., Bilgin, D., Lim, Y. W., Peterson, T. C., Mayfield, S., Haerizadeh, F., Shurin, J. B., *et al.* (2016) Diversity, productivity, and stability of an industrial microbial ecosystem. *Appl. Environ. Microbiol.*, 82, 2494–2505
- Shurin, J. B., Abbott, R. L., Deal, M. S., Kwan, G. T., Litchman, E., McBride, R. C., Mandal, S. and Smith, V. H. (2013) Industrial-strength ecology: trade-offs and opportunities in algal biofuel production. *Ecol. Lett.*, 16, 1393–1404
- Senthilvelan, T., Kanagaraj, J., Panda, R. C. and Mandal, A. B. (2014) Biodegradation of phenol by mixed microbial culture: an eco-friendly approach for the pollution reduction. *Clean Technol. Environ. Policy*, 16, 113–126
- Zhang, H., Pereira, B., Li, Z. and Stephanopoulos, G. (2015) Engineering *Escherichia coli* coculture systems for the production of biochemical products. *Proc. Natl. Acad. Sci. U.S.A.*, 112, 8266–8271
- Zuroff, T. R., Xiques, S. B. and Curtis, W. R. (2013) Consortia-mediated bioprocessing of cellulose to ethanol with a symbiotic *Clostridium phytofermentans*/yeast co-culture. *Biotechnol. Biofuels*, 6, 59
- Patle, S. and Lal, B. (2007) Ethanol production from hydrolysed agricultural wastes using mixed culture of *Zymomonas mobilis* and *Candida tropicalis*. *Biotechnol. Lett.*, 29, 1839–1843
- Wang, J. P., Lin, W., Wray, V., Lai, D. and Proksch, P. (2013) Induced production of depsipeptides by co-culturing *Fusarium tricinctum* and *Fusarium begoniae*. *Tetrahedron Lett.*, 54, 2492–2496
- Caballero, S., Kim, S., Carter, R. A., Leiner, I. M., Sušac, B., Miller, L., Kim, G. J., Ling, L. and Pamer, E. G. (2017) Cooperating commensals restore colonization resistance to vancomycin-resistant *Enterococcus faecium*. *Cell Host Microbe*, 21, 592–602.e4
- Brugiroux, S., Beutler, M., Pfann, C., Garzetti, D., Ruscheweyh, H. J., Ring, D., Diehl, M., Herp, S., Lötscher, Y., Hussain, S., *et al.* (2017) Genome-guided design of a defined mouse microbiota that confers colonization resistance against *Salmonella enterica* serovar *Typhimurium*. *Nat. Microbiol.*, 2, 16215
- Che, S. and Men, Y. (2019) Synthetic microbial consortia for biosynthesis and biodegradation: promises and challenges. *J. Ind. Microbiol. Biotechnol.*, 46, 1343–1358
- Johns, N. I., Blazejewski, T., Gomes, A. L. and Wang, H. H. (2016) Principles for designing synthetic microbial communities. *Curr. Opin. Microbiol.*, 31, 146–153
- Mee, M. T., Collins, J. J., Church, G. M. and Wang, H. H. (2014) Syntrophic exchange in synthetic microbial communities. *Proc. Natl. Acad. Sci. U.S.A.*, 111, E2149–E2156
- Embree, M., Liu, J. K., Al-Bassam, M. M. and Zengler, K. (2015) Networks of energetic and metabolic interactions define dynamics in microbial communities. *Proc. Natl. Acad. Sci. U.S.A.*, 112, 15450–15455
- Zengler, K. and Zaramela, L. S. (2018) The social network of

- microorganisms—how auxotrophies shape complex communities. *Nat. Rev. Microbiol.*, 16, 383–390
20. Bauer, E., Zimmermann, J., Baldini, F., Thiele, I. and Kaleta, C. (2017) BacArena: Individual-based metabolic modeling of heterogeneous microbes in complex communities. *PLOS Comput. Biol.*, 13, e1005544
 21. Chan, S. H. J., Simons, M. N. and Maranas, C. D. (2017) SteadyCom: Predicting microbial abundances while ensuring community stability. *PLOS Comput. Biol.*, 13, e1005539
 22. Kumar, M., Ji, B., Zengler, K. and Nielsen, J. (2019) Modelling approaches for studying the microbiome. *Nat. Microbiol.*, 4, 1253–1267
 23. Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M. and Tanabe, M. (2021) KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res.*, 49, D545–D551
 24. Caspi, R., Billington, R., Keseler, I. M., Kothari, A., Krummenacker, M., Midford, P. E., Ong, W. K., Paley, S., Subhraveti, P. and Karp, P. D. (2020) The MetaCyc database of metabolic pathways and enzymes — a 2019 update. *Nucleic Acids Res.*, 48, D445–D453
 25. King, Z. A., Lu, J., Dräger, A., Miller, P., Federowicz, S., Lerman, J. A., Ebrahim, A., Palsson, B. O. and Lewis, N. E. (2016) BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res.*, 44, D515–D522
 26. Thiele, I. and Palsson, B. Ø. (2010) A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.*, 5, 93–121
 27. Pramanik, J. and Keasling, J. D. (1997) Stoichiometric model of *Escherichia coli* metabolism: incorporation of growth-rate dependent biomass composition and mechanistic energy requirements. *Biotechnol. Bioeng.*, 56, 398–421
 28. Feist, A.M., Scholten, J.C.M., Palsson, B., Brockman, F.J., and Ideker, T. (2006) Modeling methanogenesis with a genome-scale metabolic reconstruction of *Methanosarcina barkeri*. *Mol. Syst. Biol.* 2, 2006.0004
 29. Gu, C., Kim, G. B., Kim, W. J., Kim, H. U. and Lee, S. Y. (2019) Current status and applications of genome-scale metabolic models. *Genome Biol.*, 20, 121
 30. Machado, D., Andrejev, S., Tramontano, M. and Patil, K. R. (2018) Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Res.*, 46, 7542–7553
 31. Seaver, S. M. D., Liu, F., Zhang, Q., Jeffries, J., Faria, J. P., Edirisinghe, J. N., Mundy, M., Chia, N., Noor, E., Beber, M. E., *et al.* (2021) The ModelSEED Biochemistry Database for the integration of metabolic annotations and the reconstruction, comparison and analysis of metabolic models for plants, fungi and microbes. *Nucleic Acids Res.*, 49, D575–D588
 32. Aite, M., Chevallier, M., Frioux, C., Trottier, C., Got, J., Cortés, M. P., Mendoza, S. N., Carrier, G., Dameron, O., Guillaudeux, N., *et al.* (2018) Traceability, reproducibility and wiki-exploration for “à-la-carte” reconstructions of genome-scale metabolic models. *PLOS Comput. Biol.*, 14, e1006146
 33. Henry, C. S., Jankowski, M. D., Broadbelt, L. J. and Hatzimanikatis, V. (2006) Genome-scale thermodynamic analysis of *Escherichia coli* metabolism. *Biophys. J.*, 90, 1453–1461
 34. Pereira, B., Miguel, J., Vilaça, P., Soares, S., Rocha, I. and Carneiro, S. (2018) Reconstruction of a genome-scale metabolic model for *Actinobacillus succinogenes* 130Z. *BMC Syst. Biol.*, 12, 61
 35. Orth, J. D., Thiele, I. and Palsson, B. O. (2010) What is flux balance analysis? *Nat. Biotechnol.*, 28, 245–248
 36. Edwards, J. S. and Palsson, B. O. (1999) Systems properties of the *Haemophilus influenzae* Rd metabolic genotype. *J. Biol. Chem.*, 274, 17410–17416
 37. Monk, J. M., Lloyd, C. J., Brunk, E., Mih, N., Sastry, A., King, Z., Takeuchi, R., Nomura, W., Zhang, Z., Mori, H., *et al.* (2017) iML1515, a knowledgebase that computes *Escherichia coli* traits. *Nat. Biotechnol.*, 35, 904–908
 38. Kocabaş, P., Çalık, P., Çalık, G. and Özdamar, T. H. (2017) Analyses of extracellular protein production in *Bacillus subtilis* – I: Genome-scale metabolic model reconstruction based on updated gene-enzyme-reaction data. *Biochem. Eng. J.*, 127, 229–241
 39. Nazem-Bokaei, H., Gopalakrishnan, S., Ferry, J. G., Wood, T. K. and Maranas, C. D. (2016) Assessing methanotrophy and carbon fixation for biofuel production by *Methanosarcina acetivorans*. *Microb. Cell Fact.*, 15, 10
 40. Förster, J., Famili, I., Fu, P., Palsson, B. Ø. and Nielsen, J. (2003) Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res.*, 13, 244–253
 41. Mintz-Oron, S., Meir, S., Malitsky, S., Rupp, E., Aharoni, A. and Shlomi, T. (2012) Reconstruction of *Arabidopsis* metabolic network models accounting for subcellular compartmentalization and tissue-specificity. *Proc. Natl. Acad. Sci. U.S.A.*, 109, 339–344
 42. Mardinoglu, A., Agren, R., Kampf, C., Asplund, A., Nookaew, I., Jacobson, P., Walley, A. J., Froguel, P., Carlsson, L. M., Uhlen, M., *et al.* (2013) Integration of clinical data with a genome-scale metabolic model of the human adipocyte. *Mol. Syst. Biol.*, 9, 649
 43. Arkin, A. P., Cottingham, R. W., Henry, C. S., Harris, N. L., Stevens, R. L., Maslov, S., Dehal, P., Ware, D., Perez, F., Canon, S., *et al.* (2018) KBase: The United States department of energy systems biology knowledgebase. *Nat. Biotechnol.*, 36, 566–569
 44. Heirendt, L., Arreckx, S., Pfau, T., Mendoza, S. N., Richelle, A., Heinken, A., Haraldsdóttir, H. S., Wachowiak, J., Keating, S. M., Vlasov, V., *et al.* (2019) Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v. 3.0. *Nat. Protoc.*, 14, 639–702
 45. Karp, P. D., Midford, P. E., Billington, R., Kothari, A., Krummenacker, M., Latendresse, M., Ong, W. K., Subhraveti, P., Caspi, R., Fulcher, C., *et al.* (2021) Pathway Tools version 23.0 update: software for pathway/genome informatics and systems biology. *Brief. Bioinform.*, 22, 109–126

46. Magnúsdóttir, S., Heinken, A., Kutt, L., Ravcheev, D. A., Bauer, E., Noronha, A., Greenhalgh, K., Jäger, C., Baginska, J., Wilmes, P., *et al.* (2017) Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat. Biotechnol.*, 35, 81–89
47. Blasco, T., Pérez-Burillo, S., Balzerani, F., Hinojosa-Nogueira, D., Lerma-Aguilera, A., Pastoriza, S., Cendoya, X., Rubio, Á., Gosalbes, M. J., Jiménez-Hernández, N., *et al.* (2021) An extended reconstruction of human gut microbiota metabolism of dietary compounds. *Nat. Commun.*, 12, 4728
48. Agren, R., Liu, L., Shoaie, S., Vongsangnak, W., Nookaew, I. and Nielsen, J. (2013) The RAVEN toolbox and its use for generating a genome-scale metabolic model for *Penicillium chrysogenum*. *PLOS Comput. Biol.*, 9, e1002980
49. Capela, J., Lagoa, D., Rodrigues, R., Cunha, E., Cruz, F., Barbosa, A., Bastos, J., Lima, D., Ferreira, E. C., Rocha, M., *et al.* (2022) merlin, an improved framework for the reconstruction of high-quality genome-scale metabolic models. *Nucleic Acids Res.*, 50, 6052–6066
50. Sánchez, B. J., Zhang, C., Nilsson, A., Lahtvee, P. J., Kerkhoven, E. J. and Nielsen, J. (2017) Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. *Mol. Syst. Biol.*, 13, 935
51. Steffensen, J. L., Dufault-Thompson, K. and Zhang, Y. (2016) PSAMM: A portable system for the analysis of metabolic models. *PLOS Comput. Biol.*, 12, e1004732
52. Björnson, E., Borén, J. and Mardinoglu, A. (2016) Personalized cardiovascular disease prediction and treatment—A review of existing strategies and novel systems medicine tools. *Front. Physiol.*, 7, 2
53. Cho, J. S., Gu, C., Han, T. H., Ryu, J. Y. and Lee, S. Y. (2019) Reconstruction of context-specific genome-scale metabolic models using multiomics data to study metabolic rewiring. *Curr. Opin. Syst. Biol.*, 15, 1–11
54. Esvap, E. and Ulgen, K. O. (2021) Advances in genome-scale metabolic modeling toward microbial community analysis of the human microbiome. *ACS Synth. Biol.*, 10, 2121–2137
55. Lin, L. (2022) Bottom-up synthetic ecology study of microbial consortia to enhance lignocellulose bioconversion. *Biotechnol. Biofuels Bioprod.*, 15, 14
56. Rodríguez, J., Kleerebezem, R., Lema, J. M. and van Loosdrecht, M. C. M. (2006) Modeling product formation in anaerobic mixed culture fermentations. *Biotechnol. Bioeng.*, 93, 592–606
57. Pramanik, J., Trelstad, P. L., Schuler, A. J., Jenkins, D. and Keasling, J. D. (1999) Development and validation of a flux-based stoichiometric model for enhanced biological phosphorus removal metabolism. *Water Res.*, 33, 462–476
58. Zomorodi, A. R. and Maranas, C. D. (2012) OptCom: a multi-level optimization framework for the metabolic modeling and analysis of microbial communities. *PLOS Comput. Biol.*, 8, e1002363
59. Khandelwal, R. A., Olivier, B. G., Röling, W. F. M., Teusink, B. and Bruggeman, F. J. (2013) Community flux balance analysis for microbial consortia at balanced growth. *PLoS One*, 8, e64567
60. Thommes, M., Wang, T., Zhao, Q., Paschalidis, I. C. and Segrè, D. (2019) Designing metabolic division of labor in microbial communities. *mSystems*, 4, e00263–e18
61. Vitkin, E., Gillis, A., Polikovsky, M., Bender, B., Golberg, A. and Yakhini, Z. (2020) Distributed flux balance analysis simulations of serial biomass fermentation by two organisms. *PLoS One*, 15, e0227363
62. Zelezniak, A., Andrejev, S., Ponomarova, O., Mende, D. R., Bork, P. and Patil, K. R. (2015) Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proc. Natl. Acad. Sci. U.S.A.*, 112, 6449–6454
63. Zhuang, K., Izallalen, M., Mouser, P., Richter, H., Risso, C., Mahadevan, R. and Lovley, D. R. (2011) Genome-scale dynamic modeling of the competition between *Rhodospirillum rubrum* and *Geobacter* in anoxic subsurface environments. *ISME J.*, 5, 305–316
64. Zomorodi, A. R., Islam, M. M. and Maranas, C. D. (2014) d-OptCom: Dynamic multi-level and multi-objective metabolic modeling of microbial communities. *ACS Synth. Biol.*, 3, 247–257
65. Harcombe, W. R., Riehl, W. J., Dukovski, I., Granger, B. R., Betts, A., Lang, A. H., Bonilla, G., Kar, A., Leiby, N., Mehta, P., *et al.* (2014) Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics. *Cell Rep.*, 7, 1104–1115
66. Borer, B., Ataman, M., Hatzimanikatis, V. and Or, D. (2019) Modeling metabolic networks of individual bacterial agents in heterogeneous and dynamic soil habitats (IndiMeSH). *PLOS Comput. Biol.*, 15, e1007127
67. Geng, J., Ji, B., Li, G., López-Isunza, F. and Nielsen, J. (2021) CODY enables quantitatively spatiotemporal predictions on in vivo gut microbial variability induced by diet intervention. *Proc. Natl. Acad. Sci. U.S.A.*, 118, e2019336118
68. O'Brien, E. J., Monk, J. M. and Palsson, B. O. (2015) Using genome-scale models to predict biological capabilities. *Cell*, 161, 971–987
69. Stolyar, S., Van Dien, S., Hillesland, K. L., Pinel, N., Lie, T. J., Leigh, J. A. and Stahl, D. A. (2007) Metabolic modeling of a mutualistic microbial community. *Mol. Syst. Biol.*, 3, 92
70. El-Semman, I. E., Karlsson, F. H., Shoaie, S., Nookaew, I., Soliman, T. H. and Nielsen, J. (2014) Genome-scale metabolic reconstructions of *Bifidobacterium adolescentis* L2-32 and *Faecalibacterium prausnitzii* A2-165 and their interaction. *BMC Syst. Biol.*, 8, 41
71. Baldini, F., Heinken, A., Heirendt, L., Magnúsdóttir, S., Fleming, R. M. T. and Thiele, I. (2019) The Microbiome Modeling Toolbox: from microbial interactions to personalized microbial communities. *Bioinformatics*, 35, 2332–2334
72. Mendes-Soares, H., Mundy, M., Soares, L. M. and Chia, N. (2016) MMinte: an application for predicting metabolic interactions among the microbial species in a community. *BMC*

- Bioinformatics, 17, 343
73. Pacheco, A. R., Moel, M. and Segrè, D. (2019) Costless metabolic secretions as drivers of interspecies interactions in microbial ecosystems. *Nat. Commun.*, 10, 103
 74. O'Brien, E. J., Lerman, J. A., Chang, R. L., Hyduke, D. R. and Palsson, B. Ø. (2013) Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Mol. Syst. Biol.*, 9, 693
 75. Khodayari, A. and Maranas, C. D. (2016) A genome-scale *Escherichia coli* kinetic metabolic model k-ecoli457 satisfying flux data for multiple mutant strains. *Nat. Commun.*, 7, 13806
 76. Klitgord, N. and Segrè, D. (2010) Environments that induce synthetic microbial ecosystems. *PLOS Comput. Biol.*, 6, e1001002
 77. Heinken, A. and Thiele, I. (2015) Anoxic conditions promote species-specific mutualism between gut microbes In *Silico*. *Appl. Environ. Microbiol.*, 81, 4049–4061
 78. Chen, K., Gao, Y., Mih, N., O'Brien, E. J., Yang, L. and Palsson, B. O. (2017) Thermosensitivity of growth is determined by chaperone-mediated proteome reallocation. *Proc. Natl. Acad. Sci. U.S.A.*, 114, 11548–11553
 79. Yang, L., Mih, N., Anand, A., Park, J. H., Tan, J., Yurkovich, J. T., Monk, J. M., Lloyd, C. J., Sandberg, T. E., Seo, S. W., *et al.* (2019) Cellular responses to reactive oxygen species are predicted from molecular mechanisms. *Proc. Natl. Acad. Sci. U.S.A.*, 116, 14368–14373
 80. Du, B., Yang, L., Lloyd, C. J., Fang, X. and Palsson, B. O. (2019) Genome-scale model of metabolism and gene expression provides a multi-scale description of acid stress responses in *Escherichia coli*. *PLOS Comput. Biol.*, 15, e1007525
 81. diCenzo, G. C., Tesi, M., Pfau, T., Mengoni, A. and Fondi, M. (2020) Genome-scale metabolic reconstruction of the symbiosis between a leguminous plant and a nitrogen-fixing bacterium. *Nat. Commun.*, 11, 2574
 82. Shreiner, A. B., Kao, J. Y. and Young, V. B. (2015) The gut microbiome in health and in disease. *Curr. Opin. Gastroenterol.*, 31, 69–75
 83. Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., Liang, S., Zhang, W., Guan, Y., Shen, D., *et al.* (2012) A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature*, 490, 55–60
 84. Turnbaugh, P. J., Hamady, M., Yatsunenko, T., Cantarel, B. L., Duncan, A., Ley, R. E., Sogin, M. L., Jones, W. J., Roe, B. A., Affourtit, J. P., *et al.* (2009) A core gut microbiome in obese and lean twins. *Nature*, 457, 480–484
 85. Heinken, A., Ravcheev, D. A., Baldini, F., Heirendt, L., Fleming, R. M. T. and Thiele, I. (2019) Systematic assessment of secondary bile acid metabolism in gut microbes reveals distinct metabolic capabilities in inflammatory bowel disease. *Microbiome*, 7, 75
 86. Diener, C., Gibbons, S. M. and Resendis-Antonio, O. (2020) MICOM: Metagenome-scale modeling to infer metabolic interactions in the gut microbiota. *mSystems*, 5, e00606–e00619
 87. Kumar, M., Ji, B., Babaei, P., Das, P., Lappa, D., Ramakrishnan, G., Fox, T. E., Haque, R., Petri, W. A., Bäckhed, F., *et al.* (2018) Gut microbiota dysbiosis is associated with malnutrition and reduced plasma amino acid levels: Lessons from genome-scale metabolic modeling. *Metab. Eng.*, 49, 128–142
 88. Shoaie, S. and Nielsen, J. (2014) Elucidating the interactions between the human gut microbiota and its host through metabolic modeling. *Front. Genet.*, 5, 86
 89. Shoaie, S., Ghaffari, P., Kovatcheva-Datchary, P., Mardinoglu, A., Sen, P., Pujos-Guillot, E., de Wouters, T., Juste, C., Rizkalla, S., Chilloux, J., *et al.* (2015) Quantifying diet-induced metabolic changes of the human gut microbiome. *Cell Metab.*, 22, 320–331
 90. Swainston, N., Smallbone, K., Hefzi, H., Dobson, P. D., Brewer, J., Hanscho, M., Zielinski, D. C., Ang, K. S., Gardiner, N. J., Gutierrez, J. M., *et al.* (2016) Recon 2. 2: from reconstruction to model of human metabolism. *Metabolomics*, 12, 109
 91. Brunk, E., Sahoo, S., Zielinski, D. C., Altunkaya, A., Dräger, A., Mih, N., Gatto, F., Nilsson, A., Preciat Gonzalez, G. A., Aurich, M. K., *et al.* (2018) Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nat. Biotechnol.*, 36, 272–281
 92. Duarte, N. C., Becker, S. A., Jamshidi, N., Thiele, I., Mo, M. L., Vo, T. D., Srivas, R. and Palsson, B. Ø. (2007) Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc. Natl. Acad. Sci. U.S.A.*, 104, 1777–1782
 93. Sahoo, S. and Thiele, I. (2013) Predicting the impact of diet and enzymopathies on human small intestinal epithelial cells. *Hum. Mol. Genet.*, 22, 2705–2722
 94. Zampieri, M. and Sauer, U. (2016) Model-based media selection to minimize the cost of metabolic cooperation in microbial ecosystems. *Bioinformatics*, 32, 1733–1739
 95. Du, H., Pan, J., Zou, D., Huang, Y., Liu, Y., and Li, M. (2022) Microbial active functional modules derived from network analysis and metabolic interactions decipher the complex microbiome assembly in mangrove sediments. *Microbiome*, 10, 224
 96. Mahadevan, R., Edwards, J. S. and Doyle, F. J. 3rd. (2002) Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys. J.*, 83, 1331–1340
 97. Hanly, T. J. and Henson, M. A. (2013) Dynamic metabolic modeling of a microaerobic yeast co-culture: predicting and optimizing ethanol production from glucose/xylose mixtures. *Biotechnol. Biofuels*, 6, 44
 98. Hanly, T. J. and Henson, M. A. (2011) Dynamic flux balance modeling of microbial co-cultures for efficient batch fermentation of glucose and xylose mixtures. *Biotechnol. Bioeng.*, 108, 376–385
 99. Salimi, F., Zhuang, K. and Mahadevan, R. (2010) Genome-scale metabolic modeling of a clostridial co-culture for consolidated bioprocessing. *Biotechnol. J.*, 5, 726–738
 100. García-Jiménez, B., García, J. L. and Nogales, J. (2018)

- FLYCOP: metabolic modeling-based analysis and engineering microbial communities. *Bioinformatics*, 34, i954–i963
101. Colarusso, A. V., Goodchild-Michelman, I., Rayle, M. and Zomorodi, A. R. (2021) Computational modeling of metabolism in microbial communities on a genome-scale. *Curr. Opin. Syst. Biol.*, 26, 46–57
 102. García-Jiménez, B., Torres-Bacete, J. and Nogales, J. (2020) Metabolic modelling approaches for describing and engineering microbial communities. *Comput. Struct. Biotechnol. J.*, 19, 226–246
 103. Burgard, A. P., Pharkya, P. and Maranas, C. D. (2003) Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol. Bioeng.*, 84, 647–657
 104. Segrè, D., Vitkup, D. and Church, G. M. (2002) Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl. Acad. Sci. U.S.A.*, 99, 15112–15117
 105. Tepper, N. and Shlomi, T. (2010) Predicting metabolic engineering knockout strategies for chemical production: accounting for competing pathways. *Bioinformatics*, 26, 536–543
 106. Choi, H. S., Lee, S. Y., Kim, T. Y. and Woo, H. M. (2010) In silico identification of gene amplification targets for improvement of lycopene production. *Appl. Environ. Microbiol.*, 76, 3097–3105
 107. Bernstein, D. B., Sulheim, S., Almaas, E. and Segrè, D. (2021) Addressing uncertainty in genome-scale metabolic model reconstruction and analysis. *Genome Biol.*, 22, 64
 108. Xavier, J. C., Patil, K. R. and Rocha, I. (2017) Integration of biomass formulations of genome-scale metabolic models with experimental data reveals universally essential cofactors in prokaryotes. *Metab. Eng.*, 39, 200–208
 109. Lieven, C., Beber, M. E., Olivier, B. G., Bergmann, F. T., Ataman, M., Babaei, P., Bartell, J. A., Blank, L. M., Chauhan, S., Correia, K., *et al.* (2020) MEMOTE for standardized genome-scale metabolic model testing. *Nat. Biotechnol.*, 38, 272–276
 110. Maghini, D. G., Moss, E. L., Vance, S. E. and Bhatt, A. S. (2021) Improved high-molecular-weight DNA extraction, nanopore sequencing and metagenomic assembly from the human gut microbiome. *Nat. Protoc.*, 16, 458–471
 111. Zheng, W., Zhao, S., Yin, Y., Zhang, H., Needham, D. M., Evans, E. D., Dai, C. L., Lu, P. J., Alm, E. J. and Weitz, D. A. (2022) High-throughput, single-microbe genomics with strain resolution, applied to a human gut microbiome. *Science*, 376, eabm1483
 112. Lachance, J. C., Lloyd, C. J., Monk, J. M., Yang, L., Sastry, A. V., Seif, Y., Palsson, B. O., Rodrigue, S., Feist, A. M., King, Z. A., *et al.* (2019) BOFdat: Generating biomass objective functions for genome-scale metabolic models from experimental data. *PLOS Comput. Biol.*, 15, e1006971
 113. Leng, H., Wang, Y., Zhao, W., Sievert, S. M. and Xiao, X. (2022) An expanded deep-branching thermophilic bacterial clade sheds light on the early evolution of bacteria. *BioRxiv*
 114. Dreyfuss, J. M., Zucker, J. D., Hood, H. M., Ocasio, L. R., Sachs, M. S. and Galagan, J. E. (2013) Reconstruction and validation of a genome-scale metabolic model for the filamentous fungus *Neurospora crassa* using FARM. *PLOS Comput. Biol.*, 9, e1003126
 115. Ryu, J. Y., Kim, H. U. and Lee, S. Y. (2019) Deep learning enables high-quality and high-throughput prediction of enzyme commission numbers. *Proc. Natl. Acad. Sci. U.S.A.*, 116, 13996–14001
 116. Medlock, G. L. and Papin, J. A. (2020) Guiding the refinement of biochemical knowledgebases with ensembles of metabolic networks and machine learning. *Cell Syst.*, 10, 109–119.e3
 117. Plaimas, K., Mallm, J. P., Oswald, M., Svara, F., Sourjik, V., Eils, R. and König, R. (2008) Machine learning based analyses on metabolic networks supports high-throughput knockout screens. *BMC Syst. Biol.*, 2, 67
 118. Zhang, J., Petersen, S. D., Radivojevic, T., Ramirez, A., Pérez-Manríquez, A., Abeliuk, E., Sánchez, B. J., Costello, Z., Chen, Y., Fero, M. J., *et al.* (2020) Combining mechanistic and machine learning models for predictive engineering and optimization of tryptophan metabolism. *Nat. Commun.*, 11, 4880
 119. Oyetunde, T., Liu, D., Martin, H. G. and Tang, Y. J. (2019) Machine learning framework for assessment of microbial factory performance. *PLoS One*, 14, e0210558
 120. Culley, C., Vijayakumar, S., Zampieri, G. and Angione, C. (2020) A mechanism-aware and multiomic machine-learning pipeline characterizes yeast cell growth. *Proc. Natl. Acad. Sci. U.S.A.*, 117, 18869–18879
 121. Yang, J. H., Wright, S. N., Hamblin, M., McCloskey, D., Alcantar, M. A., Schrübbers, L., Lopatkin, A. J., Satish, S., Nili, A., Palsson, B. O., *et al.* (2019) A white-box machine learning approach for revealing antibiotic mechanisms of action. *Cell*, 177, 1649–1661.e9
 122. Li, L., Zhou, X., Ching, W. K. and Wang, P. (2010) Predicting enzyme targets for cancer drugs by profiling human metabolic reactions in NCI-60 cell lines. *BMC Bioinformatics*, 11, 501
 123. DiMucci, D., Kon, M. and Segrè, D. (2018) Machine learning reveals missing edges and putative interaction mechanisms in microbial ecosystem network. *mSystems*, 3, e00181–e18
 124. Perez-Garcia, O., Lear, G. and Singhal, N. (2016) Metabolic network modeling of microbial interactions in natural and engineered environmental systems. *Front. Microbiol.*, 7, 673