



Inferring composition and function of the human gut microbiome in time and space: A review of genome-scale metabolic modelling tools



Álvaro Altamirano, Pedro A. Saa, Daniel Garrido*

Department of Chemical and Bioprocess Engineering, School of Engineering, Pontificia Universidad Católica de Chile, Vicuña Mackenna, 4860 Santiago, Chile

ARTICLE INFO

Article history:

Received 13 September 2020
Received in revised form 20 November 2020
Accepted 21 November 2020
Available online 1 December 2020

ABSTRACT

The human gut hosts a complex community of microorganisms that directly influences gastrointestinal physiology, playing a central role in human health. Because of its importance, the metabolic interplay between the gut microbiome and host metabolism has gained special interest. While there has been great progress in the field driven by metagenomics and experimental studies, the mechanisms underpinning microbial composition and interactions in the microbiome remain poorly understood. Genome-scale metabolic models are mathematical structures capable of describing the metabolic potential of microbial cells. They are thus suitable tools for probing the metabolic properties of microbial communities. In this review, we discuss the most recent and relevant genome-scale metabolic modelling tools for inferring the composition, interactions, and ultimately, biological function of the constituent species of a microbial community with special emphasis in the gut microbiota. Particular attention is given to constraint-based metabolic modelling methods as well as hybrid agent-based methods for capturing the interactions and behavior of the community in time and space. Finally, we discuss the challenges hindering comprehensive modelling of complex microbial communities and its application for the *in-silico* design of microbial consortia with therapeutic functions.

© 2020 The Author(s). Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Contents

1. Introduction	3897
2. From genome-scale network reconstructions to stoichiometric models of metabolism	3898
2.1. Genome-scale network reconstructions for microbial community modelling	3898
2.2. Genome-scale metabolic models and constrained-based methods	3899
3. Metabolic properties of microbial communities inferred from genome-scale metabolic models	3899
3.1. Prediction of composition in microbial communities	3899
3.2. Probing metabolic interactions within microbial communities	3900
3.3. Inferring biological functions under different environmental conditions	3902
4. Challenges and outlook	3902
CRedit authorship contribution statement	3903
Declaration of Competing Interest	3903
Acknowledgments	3903
References	3903

1. Introduction

Microbiomes could be understood as the collection of microorganisms, their genomes and their surrounding habitat [1]. The human gut hosts a complex and dynamic community of microorganisms that directly modulates gastrointestinal physiology, also

* Corresponding author.

E-mail address: dgarrido@ing.puc.cl (D. Garrido).

playing important roles beyond the digestive system [2]. The gut microbiome is characterized by a large diversity of microbial species belonging to the phyla Firmicutes and Bacteroidetes [3]. Constituent species of these phyla are responsible for the fermentation of dietary substrates into important metabolites including Short Chain Fatty Acids (SCFAs) [4], which serve as energy for the epithelium and modulate important physiological responses [2]. Although gut microbiome profiles are remarkably unique and stable in time for each individual, there are clear transitions in its composition associated with age, diet, antibiotics, or diseases that result in specific signatures [3].

Microbial communities can be studied by the diversity and abundance of their constitutive species, as well as the interactions between them and their environment, e.g., the human host. For instance, composition changes of the intestinal microbiota drive ecological interactions among microbial populations and the host [5], which in the face of external factors or interventions such as diet, drugs, or antibiotics, defines the biological function of the community and its impact on human health [6]. Thus, metabolic interactions between gut microbes and with their host are essential yet they remain largely unknown. Mathematical models of these interactions could help unravel the impact on the microbiome composition upon external interventions, which can be then employed to design microbiome consortia with desired therapeutic functions [7,8].

Experimental elucidation of microbial interactions is challenging as the number of possible relations grows exponentially with the number of species in the community, thereby hindering implementation of conventional experimental designs. On the other hand, computational methods can readily simulate and propose plausible biological interactions within the gut community, which can be later verified experimentally [9]. Depending on the type of computational framework employed, different aspects of the community can be probed. There are four general approaches used for studying microbial communities [10]: modelling accounting for phenotypic traits [11–13], modelling based on sequence abundance [14–16], agent-based modelling [17–19], and constraint-based methods (CBMs) using metabolic reconstructions from annotated genomes [19–22]. The first category, usually applied to small communities, employs non-structured Ordinary Differential Equations (ODEs) models to describe the time-course profile of biomass and metabolite concentration changes in the community, resulting from growth activation/inhibition type interactions. The second group applies statistical methods to infer patterns of co-occurrence and co-exclusion, or to estimate physiological traits (e.g., growth rate) based on metagenomic analyses. Agent-based modelling describes different elements of an autonomous system (e.g., populations, metabolites, or environment), interacting with each other and reacting to perturbations based on defined rules. The last group employs genome-scale stoichiometric models and optimization methods to determine the feasible phenotypic space for a metabolic network, or networks in the case of a community. In particular, CBMs for metabolic modelling have emerged as a popular framework for mechanistically describing the behavior of microbial communities in different contexts and for enabling systematic integration of disparate omics data [10,19,23]. Furthermore, the continuous development of novel methods and tools has prompted CBMs as the framework of choice for studying microbial communities.

In this review, we review the most relevant metabolic modelling tools for inferring the composition, interactions, and ultimately, the biological function of the constituent species in a microbial community with particular emphasis on the human gut microbiota. Following a brief overview of relevant CBM methods for community modelling, we present recent tools and applications for describing microbial community behavior in time and

space, with emphasis but not limited to the human gut. Both time and space dependencies are critical for shaping the collective behavior of the community [22,24,25], yet it has not been until recently that more advanced modelling tools have become available for its use and review. Finally, we discuss the challenges of modelling complex microbial communities such as the gut microbiota in time and space, and the steps required for rationally designing effective microbial consortia for tailored applications.

2. From genome-scale network reconstructions to stoichiometric models of metabolism

2.1. Genome-scale network reconstructions for microbial community modelling

Genome-Scale Metabolic Models (GSMMs) are mathematical structures derived from GENome-scale Network REconstructions (GENREs; Fig. 1A). The latter represent the complete repertoire of biochemical reactions of a single cell or a microbial community. The collection of reactions, as well as the participating metabolites, are assembled based on the list of enzymes contained in the annotated genome, providing Gene-Protein-Reaction (GPR) relationships that capture the genotype-phenotype relation. Several tools are available for the semi-automatic reconstruction of these networks. Among the most popular are modelSEED [26] and RAVEN [27], among many others [28], which determine the relevant GPRs from the annotated genomic sequence using different databases

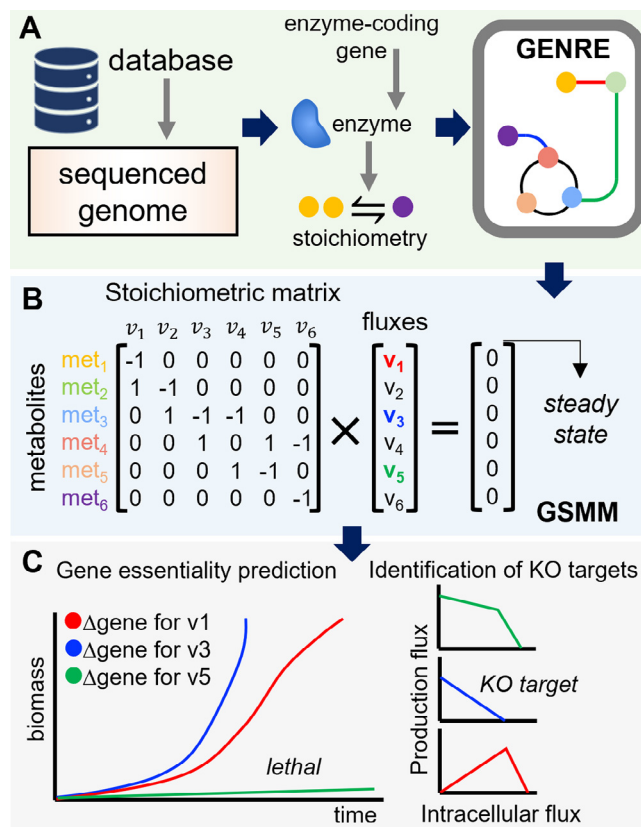


Fig. 1. Illustrative applications and construction of GSMM from GENREs. **A** From an annotated genome, Gene-Protein-Reactions (GPR) relations are defined which link genes, enzymes and reactions. The set of all GPRs defines the metabolic network reconstruction. **B** A GSMM is then constructed using the stoichiometric matrix encoding all the reactions stoichiometries and imposing the steady-state condition among other suitable constraints. **C** The GSMM can be then employed for predicting the effect of gene knockouts on growth or simulating the impact of individual intracellular fluxes on the production of a given metabolite, among other uses.

yielding a draft reconstruction. This draft reconstruction will often require laborious curation, e.g., gap-filling [29,30]. More recently, a different bottom-up approach called CarveMe [31] has been proposed. Starting from a curated universal model containing reactions from well-curated GSMM databases (e.g., BiGG [32]), CarveMe optimizes the reactions to be maintained in the new reconstruction based on sequence similarity with the annotated genome. Notably, this tool has been used to build high-quality metabolic single-species and community models of the human gut microbiota [31].

The availability of genome sequencing data supported by (semi) automatic tools for GENREs construction has led to the emergence of various network reconstruction resources for exploring the behavior of microbial communities. The most comprehensive resource is AGORA, which is an open repository of over 800 semi-automatically reconstructed GENREs of microorganisms from the human gut [33]. This resource has served as a base for multiple studies, including bile acids metabolism in the human gut [34] and the development of a large human gut microbiota metabolic model that integrates metagenomic data [35]. Another key resource is NJC19 [36], which so far represents the largest literature-based network resource that comprises human and mouse microbiome knowledge. NJC19 serves as a reference map and consists of two groups of nodes; the first group is made up of the considered species (883 microbial species and 6 mouse and human cell types), and the second group contains the metabolites related to each organism. Associations between consumed or produced metabolites by species were either identified from empirical data or predicted by bioinformatic algorithms. Importantly, NJC19 has evidenced a limited number of metabolic phenotypes currently supported as measured by the low average number of metabolites produced or consumed [37]. This information can be very valuable for building and curating metabolic networks.

2.2. Genome-scale metabolic models and constrained-based methods

From a defined GENRE, a GSMM can be constructed by mathematically encoding the metabolic reactions in the stoichiometric matrix, which represents the mass balances for each intracellular metabolite located inside the cell. Then, the imposition of the steady-state assumption for the latter enables writing a system of linear equations that represents the metabolic state of the system, here the metabolic network (Fig. 1B). Although the resulting system of equations is typically undetermined, i.e., there are infinite solutions describing the same observations, it is possible to compute particular solutions using CBMs. Briefly, by defining an objective function and applying capacity constraints on the fluxes that simulate specific environmental conditions, the optimal metabolic flux distribution that achieves that goal can be computed. The most popular of such methods is Flux Balance Analysis (FBA) [38], which has been extensively applied for better understanding microbial physiology [39–41], and more lately, to model microbial communities [42–44]. Application of FBA in GSMMs are vast and range from gene essentiality predictions to identification of genetic targets for metabolic engineering (Fig. 1C). The reader is referred elsewhere for a more comprehensive review [28].

There are four types of CBMs that have been employed for modelling microbial communities [45]: 1) lumped approach, 2) compartment per guild [35,46], 3) bi-level optimization simulation [21,47], and 4) dynamic stoichiometric modelling [48,49]. The lumped approach models the union of all reactions and metabolites from all the species of the community as if they were one organism with all common and unique metabolic pathways. In the compartment per guild approach, each microorganism is represented as a different compartment, each with its own set of reactions and metabolites. In both cases, popular CBMs such as FBA

can be employed to perform metabolic flux predictions. In contrast, bi-level optimization frameworks individualize each microorganism but apply two consecutive rounds of optimization: one for each individual species, and a second for the entire community integrating the results from the previous round. Finally, dynamic stoichiometric modelling methods employ dynamic models coupled to GSMMs to describe dynamic changes in metabolite and biomass concentration across time. The most important of such methods is dynamic FBA (d-FBA) [50]. Briefly, d-FBA employs GSMMs equipped with kinetic expressions (see [51] for more details) for describing the rate of consumption/production of external metabolites consistent with intracellular fluxes. By dynamically updating the latter rates using extracellular concentrations and using them as constraints for FBA, metabolic fluxes and biomass production can be calculated in a specific time. The latter fluxes can be then employed to update the concentrations of the external species. By repeating this procedure iteratively, one can solve the ODE system describing the concentration of biomass and external metabolites as a function of time. Importantly, this method can be readily adapted to model space and temporal dependencies by discretizing the modelled domain and solving the Partial Differential Equations (PDEs) that describe the diffusion of species. The methods reviewed hereafter use extensively this or similar approaches for modeling communities in time and space.

3. Metabolic properties of microbial communities inferred from genome-scale metabolic models

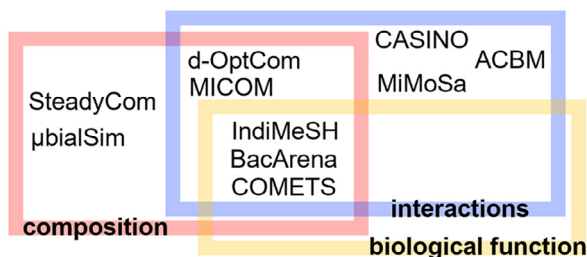
Genome-scale metabolic tools can be categorized according to their mathematical features or biological/ecological scope. Using the latter criterion, we review the most relevant and recent tools for modelling the composition, interaction and biological function of microbial communities (Fig. 2A). The presented tools are mostly based on CBMs, and as such, they provide mechanistic insights into the metabolic interactions for example between microbial species and their impact on the (human) host (Fig. 2B). The main features of the reviewed tools are illustrated in Fig. 3.

3.1. Prediction of composition in microbial communities

Metagenome samples can be sequenced for their 16S rRNA content to determine the species and their abundance – as determined by the number of reads – in the microbial community [52,53]. The reads abundance data can then be integrated into metabolic models for evaluation of the microbial community behavior. Using this strategy, a customizable collection of metabolic models (collectively called MICOM, Fig. 3A) [35] was applied to study the microbial gut composition of Swedish and Danish individuals (healthy or with diabetes) consuming a Western Diet [54]. In the study, 186 metagenome samples were gathered and relative abundance data were used to construct a community model and simulate its behavior (compartment per guild approach). The metagenomic data was relevant for the differential selection of bacterial genera, and the community objective involves a trade-off between the community and individual growth.

A recent CBM for microbial composition prediction is SteadyCom [46]. SteadyCom seeks to determine the optimal microbial composition that supports the maximum community growth rate at steady-state. In particular, SteadyCom predicted the dominance of the bacterial phyla Bacteroides and Firmicutes in the gut microbiota [46]. Notably, the method formulates and solves a finite series of linear optimization problems that do not depend on the number of species considered in the community. A limitation of this approach is that it assumes the same specific growth rate for all members of the community. This assumption is known to hold

A Classification of community modelling tools based on their scope



B Metabolic interaction insights from flux simulations

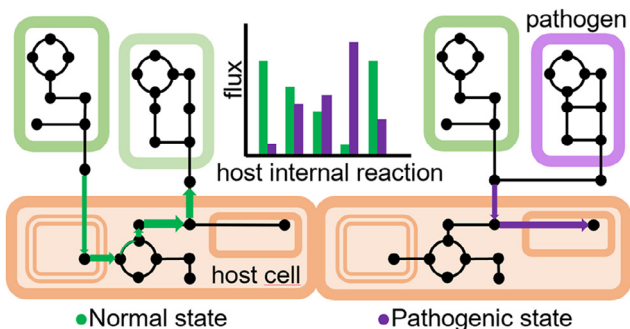


Fig. 2. Scope of different metabolic modelling tools for the study of microbial communities. **A** Metabolic modelling tools can be classified based on their capabilities for probing key ecological aspects of microbial communities, namely: composition, interactions and biological function. **B** Rooted on constraint-based methods such as Flux Balance analysis, metabolic models can provide insights into the metabolic state of a host cell (e.g., enterocyte) upon invasion of a pathogen or in the presence of commensal bacteria by comparing the flux distributions in the two conditions.

over periods of time [42]; however, it cannot be used to model dynamic conditions of microbial growth and communities.

The change of microbial composition over time is of great importance for understanding biological interactions. For this task, d-FBA is the most used CBM for metabolic modelling. µbialSim is a d-FBA-based simulator that describes the dynamic evolution of the biomass composition and metabolite concentrations in a microbial community growing in batch culture (Fig. 3B) [49]. Notably, µbialSim incorporates a numerically robust integration scheme that enables efficient simulation of a large number of species avoiding infeasible trajectories, e.g., reaching negative concentrations. The application of this method enabled satisfactory simulation of a 773-species human gut microbiome that exhibited a complex and dynamic pattern of metabolite exchanges. µbialSim is, however, limited to well-mixed environments, which could miss critical metabolic phenomena arising from spatial interactions, especially when it comes to simulation of the gut microbiota.

Simulation of community growth in time and space and assuming different objectives has also been achieved. For instance, d-OptCom [48], an extension of OptCom [47], uses a multi-level multi-objective optimization approach maximizing the overall community biomass and individual specific growth rates at each time step, which ultimately yields the time-course profile of the consortium composition. As µbialSim, d-OptCom assumes perfect mixing in the community. In contrast, COMETS [22] uses d-FBA to simulate metabolite production/consumption by different species, and diffusion on a discretized lattice to simulate spatial and

temporal gradients. Similarly, BacArena [55] has also described the evolution and spatial distribution of different species and metabolites. However, it employs a hybrid agent-based framework that combines FBA and Michaelis-Menten kinetics for describing the dynamics of the agents. Notably, this approach enabled predicting not only the abundances but also the spatial distribution of members of a representative consortium and niche separation in the human gut [55].

3.2. Probing metabolic interactions within microbial communities

Pairwise ecological interactions can be broadly considered as detrimental, beneficial, or neutral. Many of these interactions can be captured by GSMMs. In a metabolic modelling context, competition involves the uptake of common substrates (detrimental) from the environment, whereas cross-feeding (syntrophy) describe the exchange of metabolites from one organism to another (beneficial). The previously mentioned MICOM method has been applied to simulate cross-feeding interactions involved in SCFA production [35] (Fig. 3A). Computational results from MICOM were unusual, as they pointed to *Bacteroides* and *Eubacterium* genera to be responsible for acetate and propionate consumption, respectively, while *Shigella* and *Escherichia* genera were responsible for their production, which is not been commonly reported *in vivo*. Additionally, MICOM showed that competition was a common interaction between almost all members of the gut microbiota due to the large set of metabolites that could be consumed. Notably, simulation results also suggested niche partitioning in the gut microbiota (see [56,57] for other examples); there is either consumption of fibers and starches (e.g., inulin, xylan, and pectin) by a group, or consumption of Branched Chain Amino Acids (BCAAs) by another.

By using different metabolic reconstructions from various sources, the Microbiome Modelling Toolbox (MMT) [58] enables studying microbial interactions within a community in a specific context (e.g., diet, fecal compartments) using relative microbial abundances as input (Fig. 3C). Following the construction of a community model contextualized with experimental data [59], this tool indicated a metabolic cross-talk between *Lactobacillus rhamnosus* and *Caco-2* cells. CASINO [21] is another tool that enables similar analyses. By defining a certain diet and microbial abundance (e.g., estimated from 16S rRNA sequencing), CASINO enabled simulation of different metabolite production and consumption profile with the added benefit of predicting the individual contributions of each species, e.g., SCFAs production. The application of this tool is, however, limited to small communities. d-OptCom, already mentioned above, has also been applied for simulating cross-feeding metabolic interactions between *E. coli* auxotrophs [48].

As mentioned in the previous section, agent-based modelling methods are both useful for capturing the dynamic and spatial distribution of species and flexible for integrating results derived from FBA. For instance, Agent and Constraint-Based Modelling Framework (ACBM, Fig. 3D) [18] satisfactorily simulated the cross-feeding between *Faecalibacterium prausnitzii* and *Bifidobacterium adolescentis* in the gut microbiota. In this case, *F. prausnitzii* was predicted to produce four times the amount of butyrate in co-culture with *B. adolescentis*, which was in agreement with experimental data [60]. Similarly, IndiMeSH (Fig. 3E) [24] is a hybrid CBM and ABM tool, which predicted trophic relationships between different populations of a bacterial species, namely *Pseudomonas putida* and *Pseudomonas veronii*. More specifically, simulations predicted the growth of facultative anaerobes with the concomitant production of acetate in the center of a soil aggregate under wet conditions, whereas in the outer region, aerobic growth was sustained by the produced acetate. Another agent-based method called MiMoSa [25] uses a similar approach as the previous

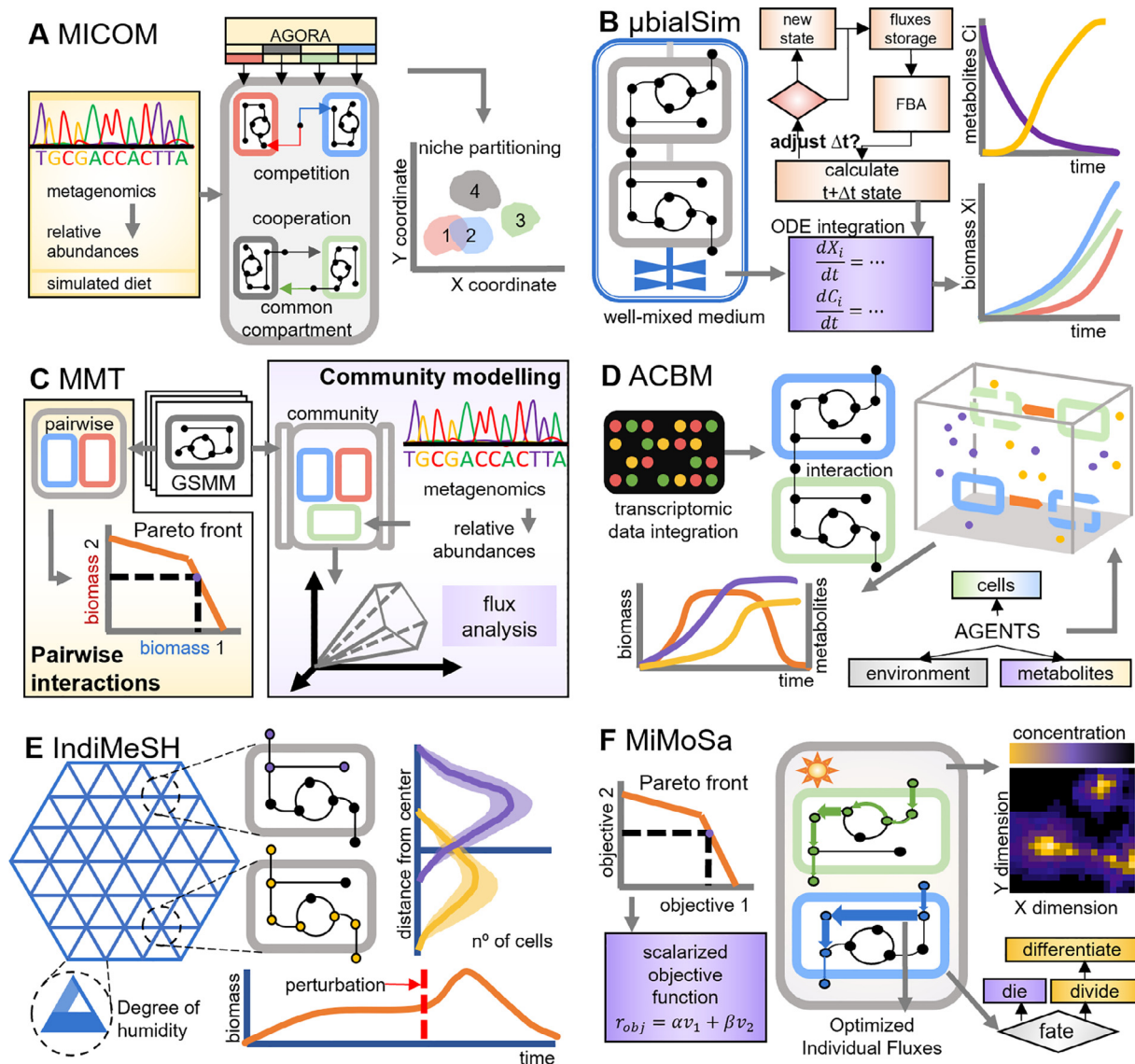


Fig. 3. Schematic representations of different genome-scale metabolic modelling tools for probing microbial communities in space and time. **A** MICOM computes microbial interactions in the human gut by integrating abundance data from metagenomic samples and GSMMs obtained from AGORA under defined conditions (e.g., western diet). **B** µbialSim describes the time-course growth and metabolic consumption/production of multiple microbial species in a well-mixed culture. **C** MMT employs a simple steady-state approach for studying pairwise interactions within the community and building the associated Pareto front. Community models generated with MMT differ from others in that non-traditional compartments can be considered. **D** ACBM uses a 3D cubic space (agents) to simulate bacterial chemotactic behavior in a microbial community. This method also enables the incorporation of transcriptomic data using constraint-based methods to characterize different phenotypes within the community. **E** IndiMeSH uses a 2D lattice to simulate the dynamics and movement of bacteria and diffusion of chemical compounds within a pore-network. IndiMeSH can simulate changes in abundance of cells upon external perturbations as well as the differentiation of the microbial population according to the local environment. **F** MiMoSa employs a dynamic multi-objective hybrid constraint-based approach for simulating concentration changes in time and space for different species. Individual optimized fluxes for each cell type (blue and green weighted arrows) are used to update the chemical concentrations of the 2D environment and also to decide whether cells divide or differentiate based on their size. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

method, but it employs multi-objective optimization instead (Fig. 3F). This tool was employed to model two different cell types – photoautotrophic and diazotrophic – of *Trichodesmium erythraeum*, and accurately predicted the differential activation of biochemical pathways by varying the weights of the objective function according to cell size. This function prioritized biomass production at small cell sizes and production of either glycogen or cyanophycin for larger cell sizes below the threshold of cell division. Although these tools have been deployed for simulating environmental habitats, they can be readily applied to represent other situations such as proliferation of infectious bacteria in

human tissues like the lungs or aerated epithelial tissues colonized with microbes like the mouth [24].

Finally, the study of pairwise interactions of mutualistic nature, or between two closely interrelated cell types, might be complemented with the quantification of the metabolic support that each species brings to the other. To this task, the Metabolic Support Index (MSI) quantifies the number of reactions of a metabolic network that are *enabled* (i.e., increased metabolic capacities) when the metabolic network of another organism interacts in a simulated co-culture [61]. Although MSI is not a method *per se*, it can be readily implemented within a CBM framework, for example,

for better understanding how mutualistic interactions may be affected by media composition.

3.3. Inferring biological functions under different environmental conditions

The biological role or function of a community depends inevitably on environmental factors such as nutrients and oxygen availability, light intensity, shear forces, water activity, or temperature. To appropriately capture this interdependency, the time evolution and spatial distribution of the cells and molecules need to be described. These are particularly the cases of ACBM, InSiMeSH, and MiMoSa.

IndiMeSH predicted the spatial distribution and metabolic behavior of *Pseudomonas stutzeri* A1501 under wet and dry conditions of soil. Under dry conditions, the entire population was in fully aerobic metabolism thanks to oxygen penetration, whereas under wet conditions, nitrate consumption was predicted (Fig. 3E). Notably, the latter resulted in overflow metabolism and the production of acetate by a denitrifying subpopulation at the center of the lattice [24]. As opposed to IndiMeSH, ACBM models single agents in a 3D cube across time. By incorporating transcriptomic data from *Escherichia coli* MG1655 in the form of constraints for metabolic simulation, this tool predicted the emergence of two distinct phenotypes in different locations determined by the glucose concentration at the end of the log phase (Fig. 3D). One of the drawbacks of this approach is that it assumed no cost for motility and momentaneous starvation, i.e., no maintenance energy requirement considered. In turn, MiMoSa enabled spatial and time-course simulation of how changes in the microenvironment – particularly light intensity – affect the metabolic flux distribution of single cells (Fig. 3F). In the *T. erythraeum* case study, MiMoSa predicted the excretion of 20% of the fixed nitrogen into the environment, improving previous estimates in terms of the energetic costs [62].

In the context of the effect of the gut microbiota on human health, there have been two recent studies that showcased how CBMs can probe biological functions [34,58]. In one study, MMT was used to generate personalized community models based on metagenomic data simulating an average European diet (Fig. 3C). FBA calculations yielded the maximum potential capability of conjugation and biotransformation of bile acids for each individual model, which is relevant for inflammatory bowel disease patients. Similarly, the metabolic potential was analyzed in microbiomes obtained from Parkinson's disease patients. A relationship between the potential of maximal production of certain metabolites and constipation – among other traits of Parkinson's disease – was found, which suggests that alterations in the gut microbiota composition may play a role in metabolic potential and related Parkinson's disease symptoms.

4. Challenges and outlook

GSMs are excellent tools for exploring the properties of microbial communities as they are suited for the application of a variety of computational methods [23]. From predicting the composition of environmental communities or host-associated consortia (e.g., human gut) [46], to understanding the biological function of individual species in the context of their community and environment [24,34,58], there is a growing family of methods that have been developed to resolve these fundamental questions, and to grasp even more complex features related to dynamics and spatial distribution microbial communities. Despite the many virtues of GSMs and the associated tools, there are still many challenges hindering further progress.

Construction of predictive GSMs from GENRES remains a laborious and challenging task. Although there are different pipelines and workflows available for streamlining the reconstruction process [26,27], detailed manual curation is still required for arriving at high-quality networks. For instance, recent comprehensive GENRES resources (e.g., AGORA [33]) contain several hundreds of reconstructions, which are mostly developed in a semi-automated fashion. Careful manual curation and revision are thus needed to ensure the quality of the network. This difficulty is only exacerbated when modelling the gut microbiota, as many of its members lack detailed biochemical information of their metabolism and functional annotation of their genes [63]. Furthermore, even if single-species GSMs can be satisfactorily constructed, formulation of a community model is not a trivial task as different strategies can be employed. The selection of the most appropriate strategy will mostly depend on the underlying assumptions, research question, and available data, which may largely vary for different communities [45]. For instance, what are the most relevant interactions within the community remains an open question for most relevant microbial systems [64].

A critical limitation of current genome-scale metabolic modelling tools is related to their temporal and spatial resolution capabilities for modelling microbial communities. This is particularly relevant for the human gut microbiota. The gut microbiota is a highly heterogeneous group of microbes spatially distributed along various sections of the human intestines. It is known that the adopted spatial organization has an important effect on human health, as it can promote different types of Inflammatory Bowel Diseases (IBDs) [65]. A recent study using a combination of CBMs modelled representative bacteria (5) of the human gut and yielded valuable insights about the role of oxygen as a key contributing factor in the longitudinal and radial spatial organization (composition) of the simulated community [66]. Some of the tools reviewed here display similar capabilities for modelling spatial distributions with even greater detail [18,22,25], however, they are all only suitable for describing small communities. Furthermore, in all these cases, there is no consideration as to how the environment – or host in the case of the human gut – will respond. In the latter case, multi-scale models of the human body [67] could be constructed and adapted for capturing the dynamics and interplay between host metabolism, microbiota, and possibly diet. In this way, more holistic therapeutic strategies may be designed for more personalized site-specific perturbations of the gut microbiota.

A more detailed understanding of the composition, metabolic interactions, and ultimately, biological functions of microbial communities within a given environment will enable the design of synthetic consortia with desired functions. Synthetic communities have already been used in bioremediation [68], and there is growing interest in designing consortia with therapeutic activities to treat relevant pathologies [7,8]. Experimental efforts can benefit from computational methods as they enable narrowing the vast design space to a more practical set of consortia. Framework based on dynamic modelling of GSMs could be extremely helpful in predicting emergent properties of microbial communities. However, they have not been applied to the design of large consortia due to the combinatorial nature of the problem. To this task, other tools – typically graph-based – have been proposed as they can cope with the massive scale at the cost of yielding only qualitative predictions. For instance, Miscoto [69] has been shown to exhaustively enumerate millions of possible minimal consortia (i.e., with the least number of members) capable of performing a defined metabolic task, i.e., producing a metabolite from a set of substrates. This method is suited for screening large sets of compounds and networks and enables redundancy analysis of metabolic tasks. The identification of keystone species is also critical for properly modelling in microbial communities [19]. These species are essen-

tial for the development of the community and, when removed, they drastically disrupt the community and environment. Eventually, experimental studies will need to be conducted to determine the presence/absence of such microbes in the synthetic community [70]. Importantly, by leveraging the power of computational methods, the biological function of keystone species will be identified and preserved – or even engineered using *in silico* tools [71] – in the synthetic consortia.

CRediT authorship contribution statement

Álvaro Altamirano: Conceptualization, Writing - original draft. **Pedro A. Saa:** Conceptualization, Writing - original draft, Writing - review & editing, Funding acquisition. **Daniel Garrido:** Conceptualization, Writing - review & editing, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded in part by ANID grants Fondecyt Regular 1190074 and Fondecyt EQM190070 (DG), Fondecyt Iniciación 11190871 (PS) and ANID Doctoral Scholarship (AA). PS acknowledges the support by FONDAP, Grant 15090007, of the Center for Genome Regulation (CGR).

References

[1] Marchesi JR, Ravel J. The vocabulary of microbiome research: a proposal. *Microbiome* 2015;3(1). <https://doi.org/10.1186/s40168-015-0094-5>.

[2] Dalile B, Van Oudenhove L, Vervliet B, Verbeke K. The role of short-chain fatty acids in microbiota–gut–brain communication. *Nat Rev Gastroenterol Hepatol* 2019;16(8):461–78. <https://doi.org/10.1038/s41575-019-0157-3>.

[3] Zmora N, Suez J, Elinav E. You are what you eat: diet, health and the gut microbiota. *Nat Rev Gastroenterol Hepatol* 2019;16(1):35–56. <https://doi.org/10.1038/s41575-018-0061-2>.

[4] Koh A, De Vadder F, Kovatcheva-Datchary P, Bäckhed F. From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. *Cell* 2016;165(6):1332–45. <https://doi.org/10.1016/j.cell.2016.05.041>.

[5] Coyte KZ, Rakoff-Nahoum S. Understanding competition and cooperation within the mammalian gut microbiome. *Curr Biol* 2019;29(11):R538–44. <https://doi.org/10.1016/j.cub.2019.04.017>.

[6] Rowland I, Gibson G, Heinken A, Scott K, Swann J, Thiele I, Tuohy K. Gut microbiota functions: metabolism of nutrients and other food components. *Eur J Nutr* 2018;57(1):1–24. <https://doi.org/10.1007/s00394-017-1445-8>.

[7] Vázquez-Castellanos JF, Bicló A, Vrancken G, Huys GRB, Raes J. Design of synthetic microbial consortia for gut microbiota modulation. *Curr Opin Pharmacol* 2019;49:52–9. <https://doi.org/10.1016/j.coph.2019.07.005>.

[8] Pascale A, Marchesi N, Govoni S, Coppola A, Gazzaruso C. The role of gut microbiota in obesity, diabetes mellitus, and effect of metformin: new insights into old diseases. *Curr Opin Pharmacol* 2019;49:1–5. <https://doi.org/10.1016/j.coph.2019.03.011>.

[9] Frioux C, Singh D, Korcsmaros T, Hildebrand F. From bag-of-genes to bag-of-genomes: metabolic modelling of communities in the era of metagenome-assembled genomes. *Comput Struct Biotechnol J* 2020;18:1722–34. <https://doi.org/10.1016/j.csbj.2020.06.028>.

[10] Kumar M, Ji B, Zengler K, Nielsen J. Modelling approaches for studying the microbiome. *Nat Microbiol* 2019;4(8):1253–67. <https://doi.org/10.1038/s41564-019-0491-9>.

[11] Pinto F, Medina DA, Pérez-Correa JR, Garrido D. Modeling metabolic interactions in a consortium of the infant gut microbiome. *Front Microbiol* 2017;8(2507).

[12] Muñoz-Tamayo R, Laroche B, Walter É, Doré J, Duncan SH, Flint HJ, et al. Kinetic modelling of lactate utilization and butyrate production by key human colonic bacterial species. *Fems Microbiol Ecol* 2011;76(3):615–24.

[13] Van Wey AS, Lovatt SJ, Roy NC, Shorten PR. Determination of potential metabolic pathways of human intestinal bacteria by modeling growth kinetics from cross-feeding dynamics. *Food Res Int* 2016;88:207–16.

[14] Mandakovic D, Rojas C, Maldonado J, Latorre M, Travisany D, Delage E, et al. Structure and co-occurrence patterns in microbial communities under acute environmental stress reveal ecological factors fostering resilience. *Sci Rep-Uk* 2018;8(1):5875.

[15] Vandeputte D, Falony G, Vieira-Silva S, Tito RY, Joossens M, Raes J. Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. *Gut* 2016;65(1):57.

[16] Korem T, Zeevi D, Suez J, Weinberger A, Avnit-Sagi T, Pompan-Lotan M, et al. Growth dynamics of gut microbiota in health and disease inferred from single metagenomic samples. *Science* 2015;349(6252):1101.

[17] Lin C, Culver J, Weston B, Underhill E, Gorky J, Dhurjati P. GutLogo: Agent-based modeling framework to investigate spatial and temporal dynamics in the gut microbiome. *PLoS ONE* 2018;13(11):e0207072.

[18] Karimian E, Motamedian E. ACBM: an integrated agent and constraint based modeling framework for simulation of microbial communities. *Sci Rep-Uk* 2020;10(1):8695.

[19] Muller EEL, Faust K, Widder S, Herold M, Martínez Arbas S, Wilmes P. Using metabolic networks to resolve ecological properties of microbiomes. *Curr Opin Syst Biol* 2018;8:73–80.

[20] Shashkova T, Popenko A, Tyakht A, Peskov K, Kosinsky Y, Bogolubsky L, et al. Agent based modeling of human gut microbiome interactions and perturbations. *PLoS ONE* 2016;11(2):e0148386.

[21] Shoaie S, Ghaffari P, Kovatcheva-Datchary P, Mardinoglu A, Sen P, Pujos-Guillot E, et al. Quantifying diet-induced metabolic changes of the human gut microbiome. *Cell Metab* 2015;22(2):320–31.

[22] Harcombe WR, Riehl WJ, Dukovski I, Granger BR, Betts A, Lang AH, et al. Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics. *Cell Rep* 2014;7(4):1104–15.

[23] Sen P, Orešič M. Metabolic modeling of human gut microbiota on a genome scale: an overview. *Metabolites* 2019;9(2).

[24] Borer B, Ataman M, Hatzimanikatis V, Or D. Modeling metabolic networks of individual bacterial agents in heterogeneous and dynamic soil habitats (IndiMeSH). *Plos Comput Biol* 2019;15(6):e1007127.

[25] Gardner JJ, Hodge B-MS, Boyle NR. Multiscale Multiobjective Systems Analysis (MiMoSA): an advanced metabolic modeling framework for complex systems. *Sci Rep-Uk* 2019;9(1):16948.

[26] Henry CS, DeJongh M, Best AA, Frybarger PM, Linsay B, Stevens RL. High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat Biotechnol* 2010;28(9):977–82.

[27] Wang H, Marcišauskas S, Sánchez BJ, Domenzain I, Hermansson D, Agren R, et al. RAVEN 2.0: A versatile toolbox for metabolic network reconstruction and a case study on *Streptomyces coelicolor*. *Plos Comput Biol* 2018;14(10):e1006541.

[28] Gu C, Kim GB, Kim WJ, Kim HU, Lee SY. Current status and applications of genome-scale metabolic models. *Genome Biol* 2019;20(1):121.

[29] Satish Kumar V, Dasika MS, Maranas CD. Optimization based automated curation of metabolic reconstructions. *BMC Bioinf* 2007;8(1):212.

[30] Schroeder WL, Saha R. OptFill, a tool for infeasible cycle-free gapfilling of stoichiometric metabolic models. *iScience* 2020;23(1).

[31] Machado D, Andrejev S, Tramontano M, Patil KR. Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Res* 2018;46(15):7542–53.

[32] King ZA, Lu J, Dräger A, Miller P, Federowicz S, Lerman JA, et al. BiGG Models: a platform for integrating, standardizing and sharing genome-scale models. *Nucl Acids Res* 2016;44(D1):D515–22.

[33] Magnúsdóttir S, Heinken A, Kutt L, Ravcheev DA, Bauer E, Noronha A, et al. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat Biotechnol* 2017;35(1):81–9.

[34] Heinken A, Ravcheev DA, Baldini F, Heirendt L, Fleming RMT, Thiele I. Systematic assessment of secondary bile acid metabolism in gut microbes reveals distinct metabolic capabilities in inflammatory bowel disease. *Microbiome* 2019;7(1):75.

[35] Diener C, Gibbons SM, Resendis-Antonio O. MICOM: metagenome-scale modeling to infer metabolic interactions in the gut microbiota. *mSystems* 2020, 5(1):e00606-00619.

[36] Lim R, Cabatbat JJT, Martin TLP, Kim H, Kim S, Sung J, et al. Large-scale metabolic interaction network of the mouse and human gut microbiota. *Sci Data* 2020;7(1):204.

[37] Tweeddale H, Ntley-McRobb L, Ferenci T. Effect of slow growth on metabolism of *Escherichia coli*, as revealed by global metabolite pool (“metabolome”) analysis. *J Bacteriol* 1998;180(19):5109–16.

[38] Varma A, Palsson BO. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Appl Environ Microb* 1994;60(10):3724–31.

[39] Torres P, Saa PA, Albiol J, Ferrer P, Agosin E. Contextualized genome-scale model unveils high-order metabolic effects of the specific growth rate and oxygenation level in recombinant *Pichia pastoris*. *Metab Eng Commun* 2019;9:e00103.

[40] Boyle NR, Morgan JA. Flux balance analysis of primary metabolism in *Chlamydomonas reinhardtii*. *Bmc Syst Biol* 2009;3(1):4.

[41] Orth JD, Conrad TM, Na J, Lerman JA, Nam H, Feist AM, et al. A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism–2011. *Mol Syst Biol* 2011;7:535.

- [42] Khandelwal RA, Olivier BG, Røling WFM, Teusink B, Bruggeman FJ. Community flux balance analysis for microbial consortia at balanced growth. *PLoS ONE* 2013;8(5):e64567.
- [43] Louca S, Doebeli M. Calibration and analysis of genome-based models for microbial ecology. *Elife* 2015;4:e08208.
- [44] Zhuang K, Izallalen M, Mouser P, Richter H, Risso C, Mahadevan R, et al. Genome-scale dynamic modeling of the competition between *Rhodospirillum rubrum* and *Geobacter* in anoxic subsurface environments. *ISME J* 2011;5(2):305–16.
- [45] Perez-Garcia O, Lear G, Singhal N. Metabolic network modeling of microbial interactions in natural and engineered environmental systems. *Front Microbiol* 2016;7(673).
- [46] Chan SHJ, Simons MN, Maranas CD. SteadyCom: predicting microbial abundances while ensuring community stability. *Plos Comput Biol* 2017;13(5):e1005539.
- [47] Zomorodi AR, Maranas CD. OptCom: a multi-level optimization framework for the metabolic modeling and analysis of microbial communities. *Plos Comput Biol* 2012;8(2):e1002363.
- [48] Zomorodi AR, Islam MM, Maranas CD. d-OptCom: dynamic multi-level and multi-objective metabolic modeling of microbial communities. *ACS Synth Biol* 2014;3(4):247–57.
- [49] Popp D, Centler F. μ BialSim: constraint-based dynamic simulation of complex microbiomes. *Front Bioeng Biotechnol* 2020;8(574).
- [50] Mahadevan R, Edwards JS, Doyle 3rd FJ. Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys J* 2002;83(3):1331–40.
- [51] Saa PA, Nielsen LK. Formulation, construction and analysis of kinetic models of metabolism: a review of modelling frameworks. *Biotechnol Adv* 2017;35(8):981–1003.
- [52] Campanaro S, Treu L, Kougias PG, De Francisci D, Valle G, Angelidaki I. Metagenomic analysis and functional characterization of the biogas microbiome using high throughput shotgun sequencing and a novel binning strategy. *Biotechnol Biofuels* 2016;9(1):26.
- [53] Medina DA, Pedreros JP, Turiel D, Quezada N, Pimentel F, Escalona A, et al. Distinct patterns in the gut microbiota after surgical or medical therapy in obese patients. *PeerJ* 2017;5:e3443.
- [54] Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin Lipidol* 2002;13(1):3–9.
- [55] Bauer E, Zimmermann J, Baldini F, Thiele I, Kaleta C. BacArena: Individual-based metabolic modeling of heterogeneous microbes in complex communities. *Plos Comput Biol* 2017;13(5):e1005544.
- [56] Albenberg L, Espipova TV, Judge CP, Bittinger K, Chen J, Laughlin A, et al. Correlation between intraluminal oxygen gradient and radial partitioning of intestinal microbiota. *Gastroenterology* 2014;147(5):1055–1063.e1058.
- [57] Tannock GW, Wilson CM, Loach D, Cook GM, Eason J, O'Toole PW, et al. Resource partitioning in relation to cohabitation of *Lactobacillus* species in the mouse forestomach. *ISME J* 2012;6(5):927–38.
- [58] Baldini F, Heinken A, Heirendt L, Magnusdottir S, Fleming RMT, Thiele I. The Microbiome Modeling Toolbox: from microbial interactions to personalized microbial communities. *Bioinformatics* 2018;35(13):2332–4.
- [59] Greenhalgh K, Ramiro-Garcia J, Heinken A, Ullmann P, Bintener T, Pacheco MP, et al. Integrated in vitro and in silico modeling delineates the molecular effects of a symbiotic regimen on colorectal-cancer-derived cells. *Cell Rep* 2019;27(5):1621–1632.e1629.
- [60] Rios-Covian D, Gueimonde M, Duncan SH, Flint HJ, de los Reyes-Gavilan CG. Enhanced butyrate formation by cross-feeding between *Faecalibacterium prausnitzii* and *Bifidobacterium adolescentis*. *Fems Microbiol Lett* 2015;362(21).
- [61] Ravikrishnan A, Blank LM, Srivastava S, Raman K. Investigating metabolic interactions in a microbial co-culture through integrated modelling and experiments. *Comput Struct Biotechnol J* 2020;18:1249–58.
- [62] Glibert PM, Bronk DA. Release of dissolved organic nitrogen by marine diazotrophic cyanobacteria, *trichodesmium* spp. *Appl Environ Microb* 1994;60(11):3996.
- [63] Lugli GA, Duranti S, Albert K, Mancabelli L, Napoli S, Viappiani A, et al. Unveiling genomic diversity among members of the species *Bifidobacterium pseudolongum*, a widely distributed gut commensal of the animal kingdom. *Appl Environ Microb* 2019;85(8):e03065–03018.
- [64] Faust K, Raes J. Microbial interactions: from networks to models. *Nat Rev Microbiol* 2012;10(8):538–50.
- [65] Honda K, Littman DR. The microbiota in adaptive immune homeostasis and disease. *Nature* 2016;535(7610):75–84.
- [66] Chan SHJ, Friedman ES, Wu GD, Maranas CD. Predicting the longitudinally and radially varying gut microbiota composition using multi-scale microbial metabolic modeling. *Processes* 2019;7(7).
- [67] Toroghi MK, Cluett WR, Mahadevan R. A multi-scale model of the whole human body based on dynamic parsimonious flux balance analysis. *IFAC-PapersOnLine* 2016;49(7):937–42.
- [68] Brune K, Bayer T. Engineering microbial consortia to enhance biomineralization and bioremediation. *Front Microbiol* 2012;3(203).
- [69] Frioux C, Fremy E, Trotter C, Siegel A. Scalable and exhaustive screening of metabolic functions carried out by microbial consortia. *Bioinformatics* 2018;34(17):i934–43.
- [70] Gutiérrez N, Garrido D. Species deletions from microbiome consortia reveal key metabolic interactions between gut microbes. *mSystems* 2019;4(4):e00185–00119.
- [71] Saa PA, Cortés MP, López J, Bustos D, Maass A, Agosin E. Expanding metabolic capabilities using novel pathway designs: computational tools and case studies. *Biotechnol J* 2019;1800734.