Systems modeling approaches for microbial community studies: from metagenomics to inference of the community structure

Mark Hanemaaijer [1,2], Wilfred F. M. Röling [2], Brett G. Olivier [1], Ruchir A. Khandelwal [1,2], Bas Teusink [1] and Frank J. Bruggeman [1]*

[1] Systems Bioinformatics, Amsterdam Institute for Molecules, Medicines and Systems, VU University Amsterdam, Amsterdam, Netherlands, [2] Molecular Cell Physiology, Amsterdam Institute for Molecules, Medicines and Systems, VU University Amsterdam, Amsterdam, Netherlands

Microbial communities play important roles in health, industrial applications and earth's ecosystems. With current molecular techniques we can characterize these systems in unprecedented detail. However, such methods provide little mechanistic insight into how the genetic properties and the dynamic couplings between individual microorganisms give rise to their dynamic activities. Neither do they give insight into what we call "the community state", that is the fluxes and concentrations of nutrients within the community. This knowledge is a prerequisite for rational control and intervention in microbial communities. Therefore, the inference of the community structure from experimental data is a major current challenge. We will argue that this inference problem requires mathematical models that can integrate heterogeneous experimental data with existing knowledge. We propose that two types of models are needed. Firstly, mathematical models that integrate existing genomic, physiological, and physicochemical information with metagenomics data so as to maximize information content and predictive power. This can be achieved with the use of constraint-based genome-scale stoichiometric modeling of community metabolism which is ideally suited for this purpose. Next, we propose a simpler coarse-grained model, which is tailored to solve the inference problem from the experimental data. This model unambiguously relate to the more detailed genome-scale stoichiometric models which act as heterogeneous data integrators. The simpler inference models are, in our opinion, key to understanding microbial ecosystems, yet until now, have received remarkably little attention. This has led to the situation where the modeling of microbial communities, using only genome-scale models is currently more a computational, theoretical exercise than a method useful to the experimentalist.

Keywords: microbial communities, metagenomic data integration, community modeling, genome-scale stoichiometric modeling, metabolism, flux balance analysis

Introduction

Microbial communities are ubiquitous in nature and play key roles in the ecosystems of our planet. Humans depend on their activities as they play essential roles in element cycling
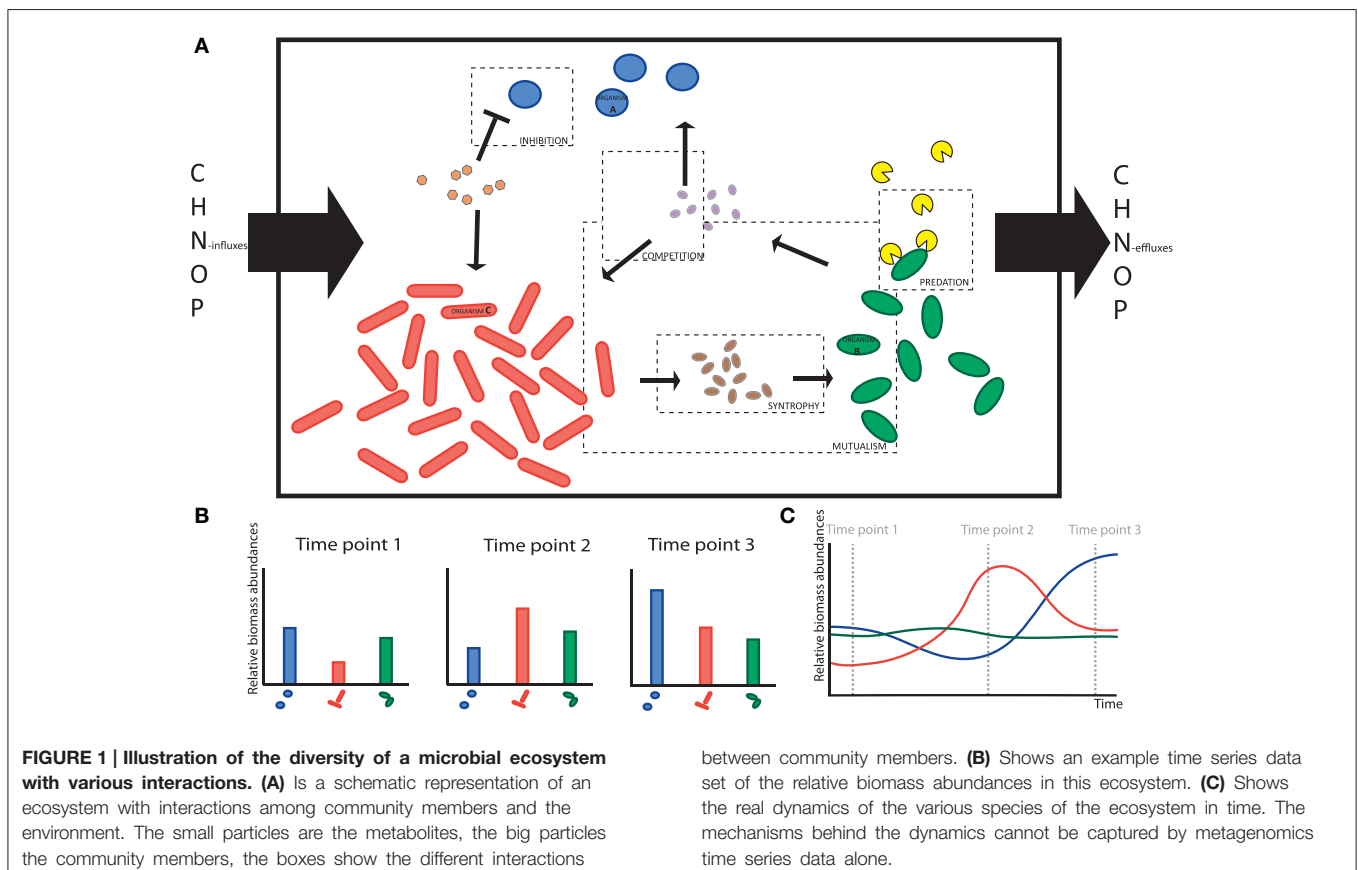
and agriculture, e.g., through interactions between plants on the one hand and mycorrhiza and nitrogen fixing bacteria on the other hand (Fellbaum et al., 2012). Microbial communities are also exploited in food fermentations, e.g., in the production of cheese, yogurt, soy sauce, sauerkraut, and vinegar. Despite this major impact on human society, we still have little understanding of the design principles that determine microbial ecosystem functioning, robustness, evolution, and control. This means that the opportunities to rationally optimize the performance of such communities are currently very limited.

In view of the complexity of the systems we are dealing with, it should be clear that experimental data alone will not provide the desired understanding -regardless of how impressive current experimental techniques may be. For example, think about high-throughput DNA-, RNA and protein-sequencing that gives high resolution information on the identities of the occurring species, their (expressed) metabolic potentials and their (relative) abundances. Yet, the information gained from such meta-omics studies (we consider 16S rRNA gene sequencing to fall under meta-omics) is still relatively limited, as they only provide an indirect view into the metabolic activities of such microorganisms and virtually none about their relationship with the environment and each other. These properties must, in turn, be inferred from the data (see **Figure 1**, in which an ecosystem with various interactions between species is depicted). Typically, such data, at best, include metagenomics at different time points, revealing

the differences in relative species abundances in the ecosystem. Methods such as stable isotope probing (SIP) (Dumont and Murrell, 2005), MAR-FISH (Lee et al., 1999) or nano-SIMS (Li et al., 2008) can provide additional and independent information on metabolites that are consumed by the various organisms in an ecosystem (He et al., 2012).

Inferring the metabolic activities of the species in a microbial community from the molecular data is largely an open problem in the field (Myrold et al., 2014). The huge challenge that microbial ecologists currently face is therefore how to infer the community state, i.e., the values of all the metabolite concentrations, species abundances and microbial activities (see also Glossary and below) from experimental data. A key question is therefore to determine what can and cannot be inferred from only metagenomic data and knowing this, what additional experimental measurements and computational methods are required to get a more complete understanding of a microbial ecosystem.

This challenge is to a large extent solved for single species where systems biology has delivered many methods for inference, data integration, and predictive modeling. These methods have become useful tools in more applied fields such as synthetic biology and metabolic engineering. Similarly, experimental methods such as MFA (see Glossary) (Toya and Shimizu, 2013) [e.g., isotopomer based (see Glossary) (Nöh et al., 2006)] and metabolomics (Wisselink et al., 2010) can be used to identify active metabolism of single species. Genome-based and

FIGURE 1 | Illustration of the diversity of a microbial ecosystem with various interactions. (A) Is a schematic representation of an ecosystem with interactions among community members and the environment. The small particles are the metabolites, the big particles the community members, the boxes show the different interactions between community members. (B) Shows an example time series data set of the relative biomass abundances in this ecosystem. (C) Shows the real dynamics of the various species of the ecosystem in time. The mechanisms behind the dynamics cannot be captured by metagenomics time series data alone.

genome-scale metabolic reconstructions can be used to integrate such data and help to understand and predict phenotypes of a microbial species (Bordbar et al., 2014; Long and Antoniewicz, 2014).

The question is to what extent these existing quantitative and computational approaches can be applied to microbial communities; several reviews and perspectives (Röling et al., 2010; Zengler and Palsson, 2012; Röling and Van Bodegom, 2014) pointed out the need for a more quantitative and systems based approach that can be used to understand microbial communities. In this contribution we will focus on the subsequent step: how to deal with quantitative data on ecosystem level? Our approach will be to investigate whether methods developed for monocultures, such as inference and quantitative modeling, are at all applicable for use in microbial ecosystems.

The single species methods we will describe have all been developed for the analysis of, and are therefore focused on, metabolism. One must therefore ask whether the microbial community state is dominantly shaped by metabolism-driven factors or whether it is also significantly dependent on factors unrelated to metabolism? One could distinguish two major classes of interactions between community members and the environment: those dictated by social traits and those driven by metabolism. We postulate that in many relevant cases, the metabolic component of the community is dominant. For instance, in glucose-fed biogas reactors the dominant species are all involved in the process of the metabolic conversion of glucose to methane (Fernandez et al., 2000). Many species-species interactions are metabolically driven, such as cross-feeding, nutrient competition, and predator-prey relations. All such metabolic processes account for mass flow through the ecosystem and the concomitant growth and turnover of microorganisms and metabolite levels.

However, communities are not solely structured by metabolic interactions. The non-metabolic interactions we exclude are social traits such as chemical warfare, bacteriocin production, quorum sensing, and other cell-to-cell interactions (either direct attachment, or other signaling mechanisms). We cannot exclude that social traits may play an important role in specific cases. Phages, for instance, can shape and alter whole ecosystems (Fuhrman, 1999), yet this important type of interaction cannot be covered by metabolism-based models. On the other hand, interactions such as quorum sensing also play a role in mono-cultures of *Escherichia coli* and *Pseudomonas aeruginosa*, however, model simulations suggests that quorum sensing has a minor effect on the phenotype of those single species (Oberhardt et al., 2008; Orth et al., 2011).

We will argue that limited quantitative experimental data, when combined with coarse-grained metabolic models, will be able to help researchers infer the community structure, i.e., the topology of species interactions (see Glossary for our definition of these often loosely defined terms). What do we mean by coarse-grained models? Basically, reduced complexity of the model. There are several ways to coarse-grain models, but in general it involves either ignoring parts of the system that are deemed less relevant, or lumping of many details of a subsystem into a higher-order description of that subsystem. A whole metabolic pathway can be lumped into one enzymatic reaction; the growth of a cell can be described by one Monod-type equation; a group of organisms can be summarized as one ecotype with an archetypical metabolic profile ("acetate consumer"). Such coarse-grained models can be gradually increased in detail over time. Parallel to these coarse-grained models, genome-scale metabolic models can be used as data repositories where new data and constraints are imposed on the models. These two approaches should in the end converge into the same description of the community. We will provide an example to illustrate how the combination of these approaches could aid in improving mechanistic understanding of microbial communities.

Metabolic fluxes in Single Microorganisms: a Largely Solved Inference Problem

In metabolic networks, metabolites are produced and consumed by enzymes that together form a complete set of chemical processes in the cell. Since enzymes are proteins encoded on the genome, and our biochemical knowledge of metabolic reactions has a long and rich history, it is possible to reconstruct a draft metabolic network largely from genomics data alone. Tools exist that automate this process and generate a complete metabolic network (Henry et al., 2010). In general, such automatic methods can be opaque and additional physiological knowledge considerably aids in curating the resulting metabolic model (Francke et al., 2005; Notebaart et al., 2006).
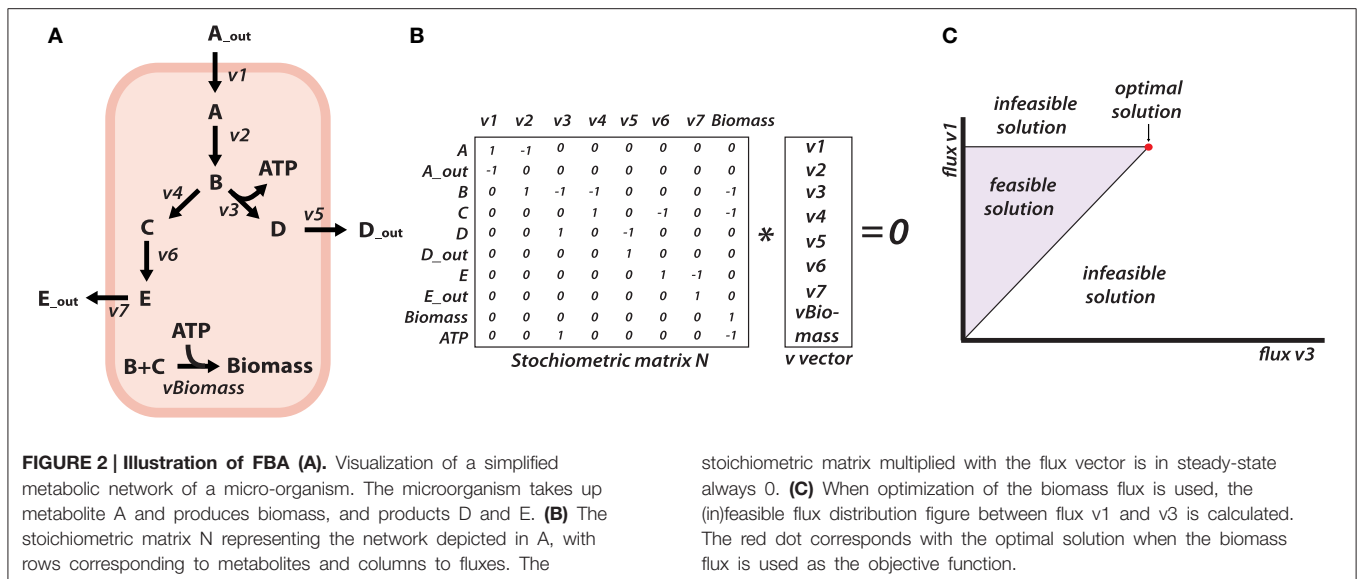
The topology of the metabolic network can be described by the stoichiometry of the constituent reactions. For example, consider the reaction catalyzed by the enzyme hexokinase, in a model description this would be described as: glucose + ATP $\rightleftharpoons$ glucose-6-phosphate + ADP. This relation between reactions and reagents can be summarized for an entire metabolic network. In mathematical form, a stoichiometry matrix is created where rows represent metabolites and the columns the reactions (See **Figure 2B** for an example).

In many applications we are interested in the behavior of a species at steady-state. At steady-state all metabolite concentrations are time-invariant, while the cells are metabolically active and consume substrate and secrete products. This implies that for each metabolite the rate of production and consumption are balanced.

In MFA, experimental data is used to deduce the true steady-state fluxes. In constraint-based modeling, FBA (see Glossary) in particular, any data that is available is combined with computational optimization techniques to predict a (in this case optimal) flux distribution through the metabolic network. Both methods have their merits and draw-backs, which we will discuss briefly.

Metabolic Flux Analysis: Flux from Data

One of the methods used to extract more information from data is the use of MFA. MFA is a powerful tool in biotechnology and systems biology as it aims to estimate intracellular fluxes in living cells. It generally applies to systems operating at a metabolic steady-state, but is not limited to steady-state conditions. MFA tries to infer intracellular fluxes from measured extracellular fluxes using a coarse-grained metabolic model where steady-state

FIGURE 2 | Illustration of FBA (A). Visualization of a simplified metabolic network of a micro-organism. The microorganism takes up metabolite A and produces biomass, and products D and E. **(B)** The stochiometric matrix N representing the network depicted in A, with rows corresponding to metabolites and columns to fluxes. The stoichiometric matrix multiplied with the flux vector is in steady-state always 0. **(C)** When optimization of the biomass flux is used, the (in)feasible flux distribution figure between flux v1 and v3 is calculated. The red dot corresponds with the optimal solution when the biomass flux is used as the objective function.

mass balances constrain the model. In the early days of MFA, there was no standard description of how the metabolism of a cell could be defined in the stoichiometric model and therefore such descriptions could vary between models. Some defined coarse-grained metabolic blocks in which multi-reaction pathways are reduced to a single reaction, such as catabolism, respiration, product formation, anabolism, and maintenance (De Hollander, 1991). Others limited the description to central metabolism, yet included all individual reactions of central metabolism in this description (Van Gulik and Heijnen, 1995).

The more recent approaches are based on partitioning of stable-isotopic labels over intracellular metabolites (for a historical overview see Szyperski, 1998 and for a more recent review see Zamboni, 2011). ^{13}C-tracer studies result in ^{13}C-enrichment profiles of metabolites in the metabolism of a given strain. The ^{13}C-enrichment in time indicates the amount of flux through the metabolic network. A fast enrichment indicates a high flux through the network, where a slow enrichment indicates a low flux. Exact flux values can be inferred using isotopomer-based flux analysis (see Glossary), using stoichiometric metabolic models. These models are based on biochemical data or genomic information and cover mostly the central metabolism. These approaches enable the estimation of a large fraction of the intracellular fluxes operative in central metabolism, in steady-state or dynamic states (Nöh et al., 2007; Wahl et al., 2008). The dynamic ^{13}C MFA applications are still largely in development, but are perhaps most interesting to microbial ecologists because of the dynamic conditions in many environmental settings. Methods have been developed (Abate et al., 2012) and applied (van Heerden et al., 2014) to infer dynamic changes in intracellular fluxes from ^{13}C tracer-experiments.

Yet, all these methods rely on data with a resolution and quantification that is still very challenging at the community level. Thus, apart from these model-aided methods to turn data into fluxes, other modeling approaches are needed: genome-scale stoichiometric metabolic models to infer fluxes largely from genomics data (Duarte et al., 2004; Teusink et al., 2009; Orth et al., 2011).

Genome-Scale Stoichiometric Models: Flux from Model Exploration

Genome-scale metabolic models contain the coupled metabolic reactions encoded by the genome of an organism (Francke et al., 2005; Feist et al., 2008) (see simplified representation of a metabolic network in **Figure 2A**). These models tend to be much larger than the stoichiometric models used for MFA, as their purpose is to be an inventory of the metabolic potential of an organism based on its genome, not a tool *per se* for inferring fluxes from data (smaller models allow for a better estimation of fewer unknowns).

An average-sized genome-scale model consists of hundreds of reactions (Teusink et al., 2009), whereas a large model can easily contain more than thousand reactions (Herrgård et al., 2008; Thiele et al., 2013). Such large models rely on exploration and optimization for hypothesis generation and flux predictions.

Among the methods used to explore genome-scale metabolic models, FBA is by far the most widely-used method. The idea behind FBA is to find solutions that satisfy some optimal behavior of the metabolic network at steady-state; most often the maximization of biomass yield of a microorganism is used as a proxy for fitness (Orth et al., 2010). To find this optimal behavior, the metabolic network has to be constrained as much as possible (Price et al., 2004). Apart from the mass-balancing constraint that defines the steady-state condition, bounds on flux values of reactions are used as additional constraints. For instance, known irreversible reactions (derived from thermodynamic considerations) are constrained to solely positive or negative flux values. Additionally, measured substrate uptake rates and product secretion rates are also used as bounds on flux values. Other physiological characteristics of the microorganism, such as the non-growth and growth associated maintenances, the biomass composition and P/O ratio help to further constrain the model.

It is also possible to obtain these physiological characteristics by exploiting the genome-scale stoichiometric metabolic models (Feist et al., 2006; Mahadevan et al., 2006). Additionally, any data that can be turned into a constraint, can be integrated into these models and bring the predicted phenotype closer to the real phenotype. Protein and mRNA data, for example, can help to improve the understanding of the fluxes through the metabolic network (Blazier and Papin, 2012; Lee et al., 2012; Reed, 2012), but also metabolomics data in combination with thermodynamic information on Gibbs energies of the reactions (Kümmel et al., 2006). The principle of FBA is shown in **Figure 2C**, where due to the use of a simple network, indeed a single optimal solution is obtained. FBA on metabolic networks of microorganisms generally reveals that even when microorganisms are behaving optimally, many internal flux distributions can enable this optimum. Various algorithms have been developed to calculate the flexibility of the network at its optimal state (Mahadevan and Schilling, 2003; Kelk et al., 2012). This flexibility is indicative for the possible phenotypes an organism could attain.

Using only stoichiometry and no dynamics appears a shortcoming of FBA, especially considering the dynamics in species abundances and activities in many ecosystems. However, dynamics of extracellular nutrients and biomass can be implemented in FBA. In such a case, Monod-type kinetics for uptake processes of substrates are used to constrain the uptake flux in the stoichiometric model, which through FBA will subsequently produce an optimal growth rate under that constraint. Resultant rates of uptake and growth can be integrated into a system of mass balances that constitute a population dynamic model. Microbial dynamics during batch cultivation have been described by using this so-called dynamic FBA (Mahadevan et al., 2002).

Genome-scale stoichiometric modeling of fluxes through metabolic networks have had an huge influence on biotechnology, contributing to the understanding of the physiology of industrial microorganisms and identifying targets for metabolic engineering to improve their performance. This modeling approach has been reviewed in several excellent reviews (Raman and Chandra, 2009; Gianchandani et al., 2010; Santos et al., 2011; McCloskey et al., 2013). We have listed a few relevant achievements of constraint-based modeling in the field of systems biology and biotechnology in **Table 1**.

Modeling in Bacterial Communities

How can we use the modeling methods for monocultures for microbial communities? To be unambiguous about the inference problem at the community scale, let's take a more formal perspective on what we mean by community structure and the community state (see also Glossary). The *community structure* implies the ordering of microorganisms through interactions, ending up in a connected network. It provides the topology of the network, much like the stoichiometry does for a metabolic network. In theory, inference methods can be used to identify such community structure, by measuring the *community state* over time (Gonzalez et al., 2012).

At any given moment in time, this community state is characterized by: (i) the abundances of all microorganisms in the community; (ii) the intracellular and extracellular concentrations of their metabolites, mRNAs, enzymes, and other cellular constituents; and (iii) the process rates (fluxes), i.e., the growth rates of all microorganisms, and the rates of all the intracellular and extracellular biotic metabolic reactions and abiotic processes. Thus, the community state describes the quantitative values of all the concentrations and process rates occurring in the community. For single species cultures, such a data set could be acquired but already constitutes a major task. There are huge experimental complications for microbial communities, and this makes the inference of the community structure also challenging. In theory, if we were to have a kinetic representation of all the processes, the characterization of the community state would suffice to determine the rates of change of all the concentrations in the community. Clearly, we do not know most of the kinetic parameters that would be required. This inference problem is still an enormous challenge in modeling monocultures (Link et al., 2014). As a result, we have to constrain the description of the community system greatly, simplify smartly, and set realistic goals.

In ecosystems, the community-level fluxes that are measured correspond to the sum of cell-specific reaction activities multiplied by their respective abundances. Flux analysis of communities therefore requires the weighing of fluxes by biomass abundances of the community members (Khandelwal et al., 2013). This makes flux analysis for a community more complicated than flux analysis for single microorganisms. The inference method that we envision will, therefore, involve the measurements of community-level fluxes, abundances of species and knowledge of the metabolism of these organisms. This requires the consideration of all the metabolic fluxes in the community, which in complex ecosystems can amount to millions of reactions—many of which can be the same reaction but carried out in different species. The constraints imposed on the network do decrease the amount of solutions in community modeling considerably.

TABLE 1 | Achievements of genome-scale stoichiometric modeling in systems biology.

Achievements	Reference
First genome-scale model	Varma and Palsson, 1994
Models showed consistency with experimental data	Feist et al., 2006; Teusink et al., 2006; Oh et al., 2007
Simulations in correspondence with ^{13}C-tracer studies	Sun et al., 2009
Non-optimal growth of strains	Ibarra et al., 2002; Teusink et al., 2009
Correct gene knock-out strategies for higher product yields	Bro et al., 2006; Pharkya and Maranas, 2006; Park et al., 2007; Lee et al., 2007

Even so, sophisticated computer models will be required to infer knowledge from microbial community data.

This situation is reminiscent of early-days MFA, where only partial genome information and data on extracellular fluxes were available. For communities the challenges are clearly much larger. FBA-type simulations on the metabolic network models of species in a microbial ecosystem will be needed to fill in inevitable gaps in our data sets. These models can integrate the available meta-omics data of the ecosystem with knowledge about reaction stoichiometries and thermodynamics. Such type of metabolic models can provide detailed insights into the metabolic capacities of single microorganisms in multispecies settings and can be extended to deal with communities (Zengler and Palsson, 2012). They provide a straightforward tool for biologists to integrate data and make realistic predictions, given constraints that derive from basic principles and experimental data. They have been successfully used to understand processes in synthetic ecosystems (Stolyar et al., 2007), but could also explain phenotypic behavior in real ecosystems (Zhuang et al., 2010).

Genome-Scale Modeling Approaches in Bacterial Communities

We therefore believe that the application of stoichiometric models of the coupled metabolisms of microorganisms in communities holds great promise, for several reasons:

1. CSSMs (see Glossary) are very suitable for data integration, as their mathematical description directly maps onto genomic, metabolomic, proteomic, and flux data of the metabolism of individual microorganisms in the community. The mapping can be done on a visualized metabolic map corresponding to the microbial ecosystem according to methods described in Maarleveld et al. (2014).
2. CSSMs allow for calculation of the community state and structure with numerical algorithms.
3. Experimental data can be used as constraints in the associated modeling formalism, to improve predictions when more data has become available (Röling and Van Bodegom, 2014).
4. The systemic consequences at community scale of molecular or physiological perturbations or species augmentation can be explored with CSSMs.
5. CSSMs can be used for experimental or medium design to improve community performance.

Different approaches to extend single-species models to microbial-community metabolic network reconstructions have been proposed, each trying to answer different research questions. The existing applications of FBA to CSSMs can be classified into three different groups: (i) the supra-organism approach (Rodríguez et al., 2006), (ii) the steady-state compartmentalized approach (Stolyar et al., 2007; Khandelwal et al., 2013), and (iii) the dynamic compartmentalized approach, based on dynamic FBA (dFBA) (Salimi et al., 2010; Zhuang et al., 2010; Hanly and Henson, 2011). These methods vary in the complexity of the CSSM description and how they choose to handle individual species.

The supra-organism approach has first been applied by Rodríguez et al. (2006). It combines all metabolic reactions of the various species in the microbial ecosystem to create a single meta-metabolic network to study the metabolic capacities in terms of product and substrate variation of the community. This approach simplifies the complexities of interactions and regulations amongst cohabiting species and makes it easy to predict environmental conditions that can be imposed to optimize a community level objective toward an outcome of interest to biotechnology. By combining all reactions into one network, it ignores the impact of species abundances and the interactions between community members.

The steady-state compartmentalized approach considers the various species as separate compartments where one shared compartment is introduced for the exchange of metabolites between the species. This approach elucidates interactions between species, contributing to an improved insight on host-microbe/pathogen interactions or mutualistic interactions. With this approach, for instance, it is suggested that instead of formate, H_2 is exchanged in a co-culture of *Desulfovibrio vulgaris* and *Methanococcus maripaludis* (Stolyar et al., 2007). Initial compartmentalization approaches neglected biomass concentrations of each species individually, resulting in biased quantitative flux distributions. Implementation of biomass concentrations in the compartmentalization approach has recently been successfully applied and is of particular interest when accurate quantitative transfer rates are required (Khandelwal et al., 2013).

The dynamic compartmentalized approach implements dFBA by using the kinetics of substrate uptake and metabolite exchange between species (Salimi et al., 2010; Zhuang et al., 2010; Hanly and Henson, 2011). Implementation of dynamic behavior is often required to understand ecosystem structure and functionality. Here, biomass concentration of each species is taken into account and can change with time. Therefore, it is possible to simulate competition, predation or other interactions altering the community state. The requirement of quantitative kinetic information is one of the major disadvantages of the dFBA method. Obtaining this for every species in a community will be laborious, if at all possible. However, the kinetic parameters can be inferred by fitting the model with experimental data.

A limitation of all these methods is that they are currently tailored for simulation and not for inference of metabolic activity and abundances of microorganisms from community-level fluxes. Even though such models can actually be used for this purpose, this is rarely done, because of experimental challenges in obtaining the data. In particular resolving fluxes at the level of the single species remains challenging. However, recently species level isotopomer-based flux data were obtained from synthetic consortia (Shaikh et al., 2008; Rühl et al., 2011). When experimental procedures for isotopomer-based analysis of fluxes through microbial communities are developed further, the CSSMs as we describe here, also become relevant to experimentalists. Moreover, recent work (Niebel, Heinemann, personal communication) provides hope that detailed thermodynamic constraints, which are physical constraints that should apply to all organisms, may sufficiently constrain metabolic networks to infer realistic intracellular flux distributions without the need for isotopic labeling data. Thermodynamic constraints can also contribute to the search for novel microorganisms that can carry

out biochemical conversions in nature that appear thermodynamically possible on paper (Raghoebarsing et al., 2006). In such work, initially a coarse-grained perspective is taken to understand overall metabolic potentials in ecosystems, and this approach is supplementary to the genome-scale approaches discussed so far.

Inferring the Community Structure and State Using Coarse-Grained Models

Even though CSSMs can be very useful, for many applications they will be too detailed and too unparameterized to be of immediate value. The details will result in many degrees of freedom, for example in many potential substrates and products. A more pragmatic approach is therefore often needed. Simplifying and coarse-graining models is also regularly required for models of monocultures. This happens when the data is still insufficient to parameterize a model, or when the model becomes so detailed that it hampers understanding. Even well-curated models might be too detailed and complicated to understand the simulated results, because there are too many variables and

independencies. In such cases there may be predictive power but no understanding. Obviously, it depends on the question what is preferred: ideally, the type of research question determines the required level of detail and initially limits the complexity of the model, at the same time allowing for progressively and gradually increasing the model in size when more experimental data becomes available. The work flow we envision is visualized in **Figure 3**.

The coarse-grained models can be created at different levels of granularity. However, in order to model microbial communities, we envisage that a more detailed (but unparameterized) description, at a lower level of granularity, is always possible. Ultimately, there is the genome of the individual consortium species that allows for a reconstruction of the metabolic potential, which is a solid ground on which to build more detailed reconstructions. When the genome of the consortium species is not present, data from metagenomics is suitable for more detailed metabolic reconstructions. However, the level of detail of the metabolic reconstructions will be largely dependent on the sequencing depth of the metagenome.
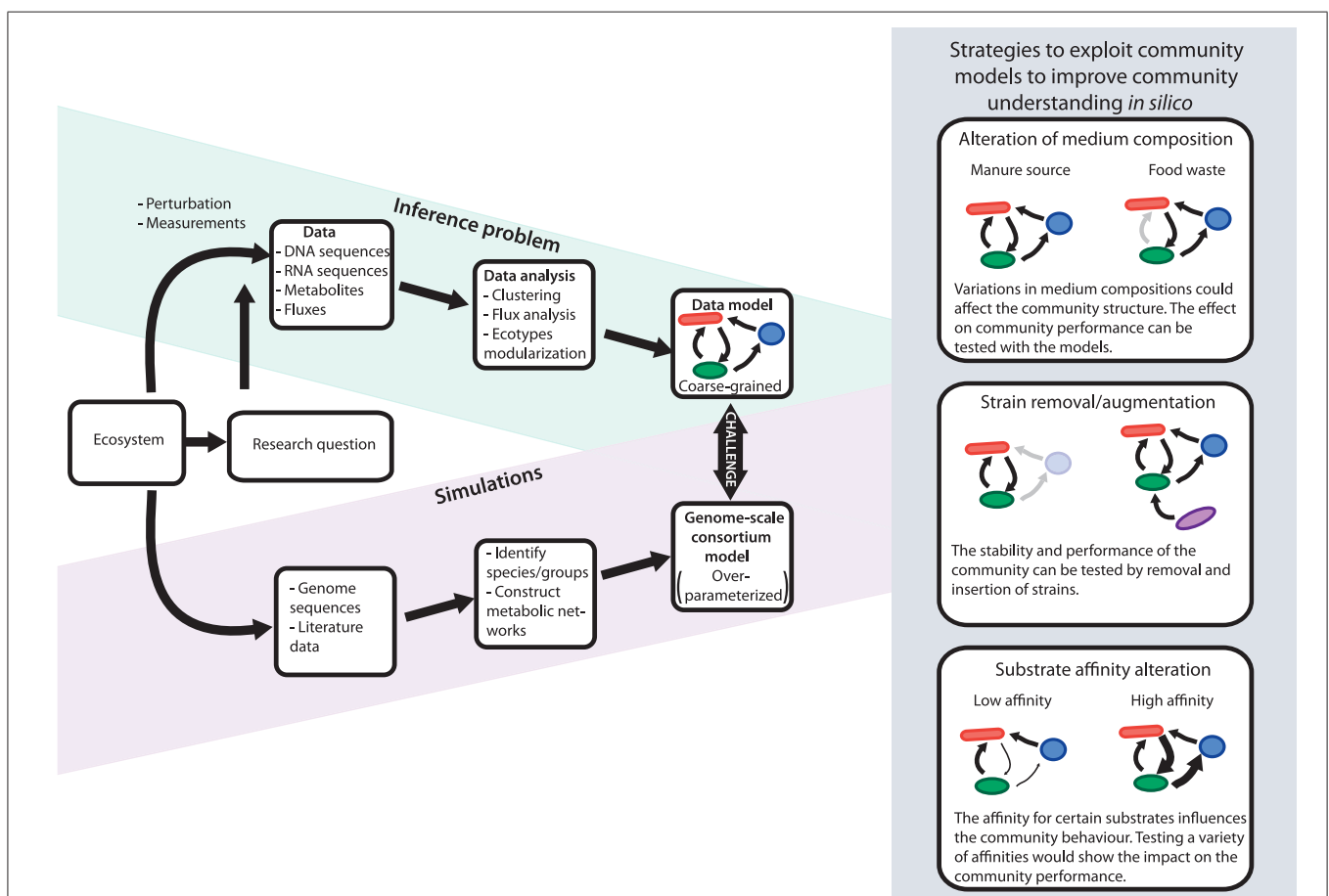
FIGURE 3 | Representation of the work flow to get more information from experimental data. On one hand the inference approach with as result a coarse-grained data-model which fits the measured data and on the other hand the genome-scale models which are highly underdetermined, but can be used as data repositories. The challenge is to create models which fit the research question. This can be done via the represented steps.
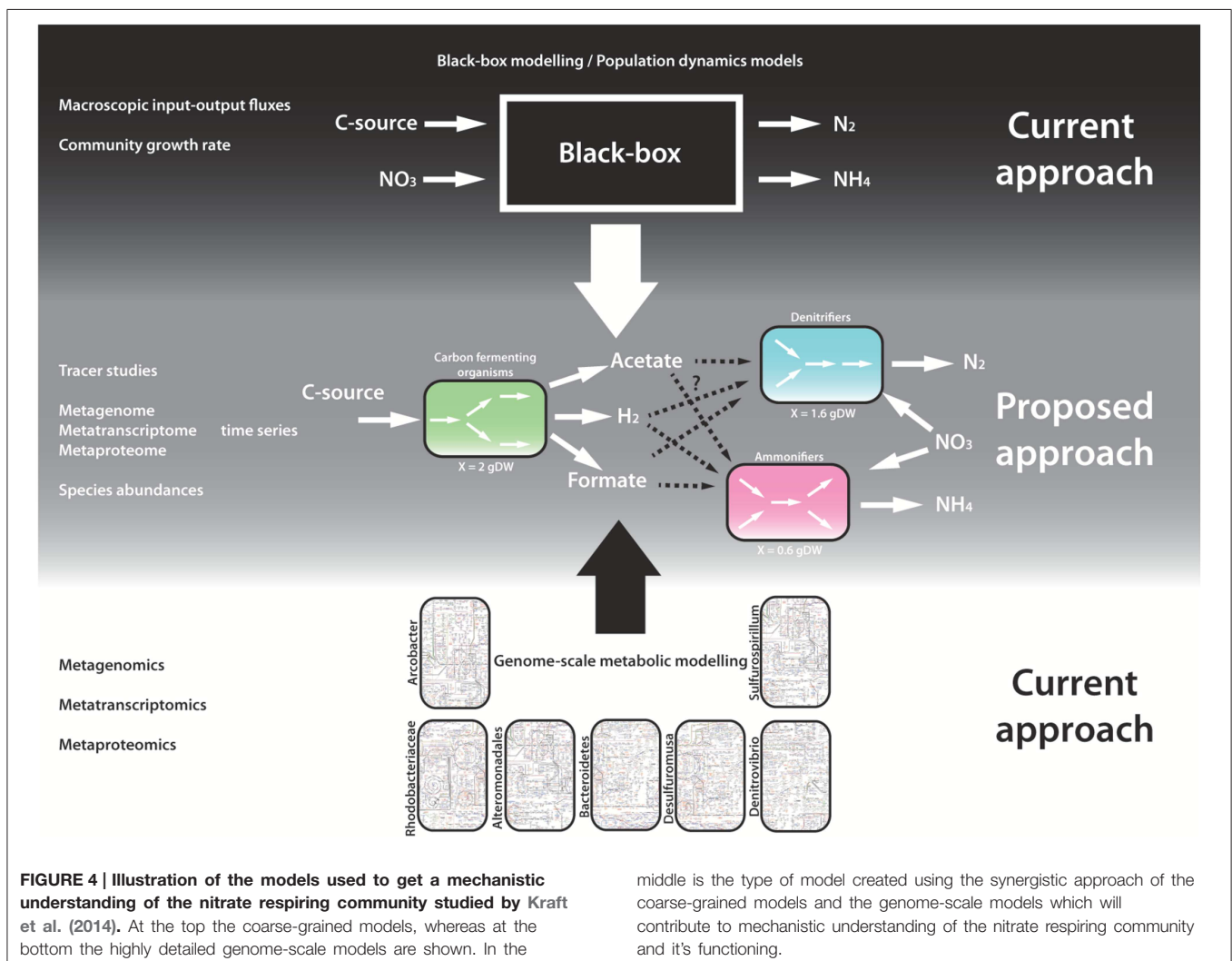
Application to a Case Study

We will make use of an elegant study performed on a nitrate respiring community (Kraft et al., 2014) as an example to illustrate and explain how the proposed work flow could be applied in practice.

Kraft et al. (2014) incubated a nitrate respiring community, originating from sediment in the German Wadden Sea, in reactors with continuous substrate supply, and altered environmental factors to investigate if denitrification or ammonification evolved as the major nitrate respiring process. The researchers measured community level fluxes, performed metagenomics, metatranscriptomics, and metaproteomics, and used fluorescence *in situ* hybridization to unravel the effect of specific environmental conditions on community behavior. A schematic overview of the metabolic interactions was created, based on the meta-omics data, with fermenting microorganisms generating fermentation products that were consumed by the microorganisms respiring nitrate to either ammonium (ammonification) or nitrogen gas (denitrification) (see **Figure 4**). In total 7 dominant populations related to ammonification and denitrification were identified.

Denitrification became dominant at short generation times and with nitrite supply, while nitrate supply and long generation times with nitrite as supply favored ammonification. The same enzymes are responsible for the first steps in both processes, which branch after nitrite reductase. Therefore, they hypothesized that differences in apparent affinity of this enzyme for nitrite determines the dominating respiration process, with the nitrite reductase of denitrifying microorganisms having higher affinities.

The study by Kraft et al. (2014) covers the "Data" and the "Data analysis" part of the "Inference problem" in **Figure 3**, but several hypotheses and open questions could not be answered by the data. This study therefore provides an interesting test case for integration of the experimental data into a mathematical model to get mechanistic understanding of community functioning and eventually could lead to control and steer its performance. The third and last step in our proposed work flow in **Figure 3** of the "Inference problem" would be the construction of a coarse-grained data model describing a limited set of pathways and connect them to each other to create a community (**Figure 4**). Once the simulations with the coarse-grained models match the experimental data, they can be applied to answer questions about

FIGURE 4 | Illustration of the models used to get a mechanistic understanding of the nitrate respiring community studied by Kraft et al. (2014). At the top the coarse-grained models, whereas at the bottom the highly detailed genome-scale models are shown. In the middle is the type of model created using the synergistic approach of the coarse-grained models and the genome-scale models which will contribute to mechanistic understanding of the nitrate respiring community and it's functioning.

the community. The type of research question will influence which parts of models should be described in high detail and which not.

For instance, to address the hypothesis that nitrite affinity determines the major respiration process, we start by creating simple models of the three functional groups (ammonifiers, denitrifiers, and carbon fermenting organisms). The model of the carbon fermenting organism will be the simplest, that is: uptake reactions of glucose and amino-acids and secretion reactions of acetate, formate and hydrogen, because testing the given hypothesis only requires the secretion of fermentation products that can act as substrate for the other functional groups. The ammonifiers model should contain the metabolic pathway of nitrate to ammonium conversion, transport reactions, carbon metabolism and other energy generating pathways. The denitrifiers require the same level of detail in their model as the ammonifiers. Now, these simple coarse-grained models allow to first investigate whether the experimentally observed differences might be explained without requiring the inclusion of the affinities for nitrite. If not, a kinetic description of the nitrite reductase in both functional groups can be included. Subsequently, it is possible to determine how big the difference in nitrite affinity should be to explain the experimental results.

Kraft et al. (2014) also observed large differences between the potential rates of nitrite reduction between the ammonification and denitrification process, in particular when multiple electron donors were fed to the consortium. They hypothesized that this could be caused by a bottleneck in electron supply to the nitrite reductases of ammonification. Those proteins require six electrons per nitrite, where only one electron is required for denitrification. Testing this hypothesis requires an additional level of detail compared to the models presented in the first example: the metabolic pathways of the different electron donors should now be implemented. A medium with multiple electron donors can be simulated and compared with a medium with only one electron donor. The model will show whether the ammonifiers become relatively less active when multiple electron donors are provided to the community compared to a situation in which one carbon source is provided.

Finally, Kraft et al. (2014) observed functional redundancy within both denitrifiers and ammonifiers, and the applied generation time determined which population became dominant within these groups. To better understand the underlying mechanisms, a more detailed model would be required compared to the previous examples. Firstly, each functional group needs to be split in several populations, each capable of either ammonification or denitrification. Secondly, for each population a more detailed metabolic model needs to be constructed, to capture differences in metabolism within the functional groups. These models are informed by the binned metagenomic data of Kraft et al. (2014). The resulting models could subsequently be tested. To establish what causes the difference in dominance in relation to generation time, kinetic information to describe competition should likely be included.

The aforementioned examples show how the question determines the level of detail of the "Inference models." These coarse-grained models of the players in the ecosystem should be based on knowledge of the (genome-derived) metabolic networks of the species. Therefore, parallel to this "Inference model," genome-scale stoichiometric metabolic models are maintained to act as data repositories (shown in Simulation in **Figure 3**). For the key species in the community, genome-scale metabolic reconstructions could be created using genomic, literature, and experimental data. All models can be combined in one metagenome-scale compartmentalized consortium model and allow for the exploitation and exploration of the microbial community.

Finally, a synergistic approach emerges where the interplay between the simple coarse-grained models and the genome-scale metabolic models are determined by the type of question being asked about the community. An important aspect is that the output of models with different levels of description should be consistent with each other and agree with available data. This process of iterative model making should inform the experimentalists about knowledge gaps and suggests new experiments.

As demonstrated by Kraft et al. culturing communities under well-controlled conditions, for instance in chemostats, enables quantification of community-level fluxes, biomass abundances and gene expression levels of specific populations, and allows for applying well-tractable perturbations. This type of experimental approach will be very helpful to develop and optimize methods for species-specific flux descriptions on basis of CSSMs, to understand the community function and eventually control the microbial ecosystem.

Concluding Remarks

In this paper, we described current methods to model the metabolism of single species and how we think such system approaches could be used at the microbial community level. These methods go from genomic information to understanding and directing microbial community metabolism. We argued that metagenomics data are great, but they need to be augmented with flux measurements and model-based inference methods to identify the community structure and state. We arrive at a proposed approach where models are combined with quantitative data. Such models can vary in their level of coarse-graining depending on the research questions, data and experimental options at hand. Genome-scale stoichiometric models will be the knowledge base with which to integrate genomic, physiological and physicochemical data, and can be developed parallel to coarse-grained models to understand the community structure, state and function. We feel that what is missing -but is within reach- is a description of the functioning of a microbial community in terms of the fluxes through its members. We therefore advocate that flux analysis specialists from the fields of biotechnology, microbial physiology, and systems biology team up with microbial ecologists, as they increasingly do.

Acknowledgments

References

Abate, A., Hillen, R. C., and Aljoscha Wahl, S. (2012). Piecewise affine approximations of fluxes and enzyme kinetics from *in vivo* 13C labeling experiments. *Int. J. Robust Nonlin. Control* 22, 1120–1139. doi: 10.1002/rnc.2798

Blazier, A. S., and Papin, J. A. (2012). Integration of expression data in genome-scale metabolic network reconstructions. *Front. Physiol* 3:299. doi: 10.3389/fphys.2012.00299

Bordbar, A., Monk, J. M., King, Z. A., and Palsson, B. O. (2014). Constraint-based models predict metabolic and associated cellular functions. *Nat. Rev. Genet.* 15, 107–120. doi: 10.1038/nrg3643

Bro, C., Regenberg, B., Förster, J., and Nielsen, J. (2006). *In silico* aided Metab. Eng. of *Saccharomyces cerevisiae* for improved bioethanol production. *Metab. Eng.* 8, 102–111. doi: 10.1016/j.ymben.2005.09.007

De Hollander, J. (1991). The use of stoichiometric relations for the description and analysis of microbial cultures. *Antonie Van Leeuwenhoek* 60, 257–273. doi: 10.1007/BF00430369

Duarte, N. C., Herrgård, M. J., and Palsson, B. Ø. (2004). Reconstruction and validation of *Saccharomyces cerevisiae* iND750, a fully compartmentalized genome-scale metabolic model. *Genome Res.* 14, 1298–1309. doi: 10.1101/gr.2250904

Dumont, M. G., and Murrell, J. C. (2005). Stable isotope probing-linking microbial identity to function. *Nat. Rev. Microbiol.* 3, 499–504. doi: 10.1038/nrmicro1162

Feist, A. M., Scholten, J. C., Palsson, B. Ø., Brockman, F. J., and Ideker, T. (2006). Modeling methanogenesis with a genome-scale metabolic reconstruction of *Methanosarcina barkeri*. *Mol. Syst. Biol.* 2:e2006.0004. doi: 10.1038/msb4100046

Feist, A. M., Herrgård, M. J., Thiele, I., Reed, J. L., and Palsson, B. Ø. (2008). Reconstruction of biochemical networks in microorganisms. *Nat. Rev. Microbiol.* 7, 129–143. doi: 10.1038/nrmicro1949

Fellbaum, C. R., Gachomo, E. W., Beesetty, Y., Choudhari, S., Strahan, G. D., Pfeffer, P. E., et al. (2012). Carbon availability triggers fungal nitrogen uptake and transport in arbuscular mycorrhizal symbiosis. *Proc. Natl. Acad. Sci. U.S.A.* 109, 2666–2671. doi: 10.1073/pnas.1118650109

Fernandez, A. S., Hashsham, S. A., Dollhopf, S. L., Raskin, L., Glagoleva, O., Dazzo, F. B., et al. (2000). Flexible community structure correlates with stable community function in methanogenic bioreactor communities perturbed by glucose. *Appl. Environ. Microbiol.* 66, 4058–4067. doi: 10.1128/AEM.66.9.4058-4067.2000

Francke, C., Siezen, R. J., and Teusink, B. (2005). Reconstructing the metabolic network of a bacterium from its genome. *Trends Microbiol.* 13, 550–558. doi: 10.1016/j.tim.2005.09.001

Fuhrman, J. A. (1999). Marine viruses and their biogeochemical and ecological effects. *Nature* 399, 541–548. doi: 10.1038/21119

Gianchandani, E. P., Chavali, A. K., and Papin, J. A. (2010). The application of flux balance analysis in systems biology. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 2, 372–382. doi: 10.1002/wsbm.60

Gonzalez, A., King, A., Robeson II, M. S., Song, S., Shade, A., Metcalf, J. L., et al. (2012). Characterizing microbial communities through space and time. *Curr. Opin. Biotechnol.* 23, 431–436. doi: 10.1016/j.copbio.2011.11.017

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. doi: 10.1002/bit.22954

He, R., Wooler, M. J., Pohlman, J. W., Catranis, C., Quensen, J., Tiedje, J. M., et al. (2012). Identification of functionally active aerobic methanotrophs in sediments from an arctic lake using stable isotope probing. *Environ. Microbiol.* 14, 1403–1419. doi: 10.1111/j.1462-2920.2012.02725.x

Henry, C. S., DeJongh, M., Best, A. A., Frybarger, P. M., Linsay, B., Stevens, R. L. (2010). High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.* 28, 977–982. doi: 10.1038/nbt.1672

Herrgård, M. J., Swainston, N., Dobson, P., Dunn, W. B., Arga, K. Y., Arvas, M., et al. (2008). A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nat. Biotechnol.* 26, 1155–1160. doi: 10.1038/nbt1492

Ibarra, R. U., Edwards, J. S., and Palsson, B. O. (2002). *Escherichia coli* K-12 undergoes adaptive evolution to achieve *in silico* predicted optimal growth. *Nature* 420, 186–189. doi: 10.1038/nature01149

Kümmel, A., Panke, S., and Heinemann, M. (2006). Putative regulatory sites unraveled by network-embedded thermodynamic analysis of metabolome data. *Mol. Syst. Biol.* 2:e2006.0034. doi: 10.1038/msb4100074

Kelk, S. M., Olivier, B. G., Stougie, L., and Bruggeman, F. J. (2012). Optimal flux spaces of genome-scale stoichiometric models are determined by a few subnetworks. *Sci. Rep.* 2:e580. doi: 10.1038/srep00580

Khandelwal, R. A., Olivier, B. G., Röling, W. F., Teusink, B., and Bruggeman, F. J. (2013). Community flux balance analysis for microbial consortia at balanced growth. *PLoS ONE* 8:e64567. doi: 10.1371/journal.pone.0064567

Kraft, B., Tegetmeyer, H. E., Sharma, R., Klotz, M. G., Ferdelman, T. G., Hettich, R. L., et al. (2014). The environmental controls that govern the end product of bacterial nitrate respiration. *Science* 345, 676–679. doi: 10.1126/science.1254070

Lee, N., Nielsen, P. H., Andreasen, K. H., Juretschko, S., Nielsen, J. L., Schleifer, K.-H., et al. (1999). Combination of fluorescent *in situ* hybridization and microautoradiography-a new tool for structure-function analyses in microbial ecology. *Appl. Environ. Microbiol.* 65, 1289–1297.

Lee, K. H., Park, J. H., Kim, T. Y., Kim, H. U., and Lee, S. Y. (2007). Systems Metab. Eng. of *Escherichia coli* for L-threonine production. *Mol. Syst. Biol.* 3:e149. doi: 10.1038/msb4100196

Lee, D., Smallbone, K., Dunn, W. B., Murabito, E., Winder, C. L., Kell, D. B., et al. (2012). Improving metabolic flux predictions using absolute gene expression data. *BMC syst. Biol* 6:73. doi: 10.1186/1752-0509-6-73

Li, T., Wu, T.-D., Mazéas, L., Toffin, L., Guerquin-Kern, J.-L., Leblon, G., et al. (2008). Simultaneous analysis of microbial identity and function using NanoSIMS. *Environ. Microbiol.* 10, 580–588. doi: 10.1111/j.1462-2920.2007.01478.x

Link, H., Christodoulou, D., and Sauer, U. (2014). Advancing metabolic models with kinetic information. *Curr. Opin. Biotechnol.* 29, 8–14. doi: 10.1016/j.copbio.2014.01.015

Long, C. P., and Antoniewicz, M. R. (2014). Metabolic flux analysis of *Escherichia coli* knockouts: lessons from the Keio collection and future outlook. *Curr. Opin. Biotechnol.* 28, 127–133. doi: 10.1016/j.copbio.2014.02.006

Maarleveld, T. R., Boele, J., Bruggeman, F. J., and Teusink, B. (2014). A data integration and visualization resource for the metabolic network of *Synechocystis* sp. PCC 6803. *Plant Physiol.* 164, 1111–1121. doi: 10.1104/pp.113.224394

Mahadevan, R., and Schilling, C. (2003). The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab. Eng.* 5, 264–276. doi: 10.1016/j.ymben.2003.09.002

Mahadevan, R., Edwards, J. S., and Doyle III, F. J. (2002). Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys. J.* 83, 1331–1340. doi: 10.1016/S0006-3495(02)73903-9

Mahadevan, R., Bond, D. R., Butler, J. E., Esteve-Nuñez, A., Coppi, M. V., Palsson, B. O., et al. (2006). Characterization of metabolism in the fe (III)-reducing organism geobacter sulfurreducens by constraint-based modeling. *Appl. Environ. Microbiol.* 72, 1558–1568. doi: 10.1128/AEM.72.2.1558-1568.2006

McCloskey, D., Palsson, B. Ø., and Feist, A. M. (2013). Basic and applied uses of genome-scale metabolic network reconstructions of *Escherichia coli*. *Mol. Syst. Biol.* 9:e661. doi: 10.1038/msb.2013.18

Myrold, D. D., Zeglin, L. H., and Jansson, J. K. (2014). The potential of metagenomic approaches for understanding soil microbial processes. *Soil Sci. Soc. Am. J.* 78, 3–10. doi: 10.2136/sssaj2013.07.0287dgs

Nöh, K., Wahl, A., and Wiechert, W. (2006). Computational tools for isotopically instationary 13C labeling experiments under metabolic steady state conditions. *Metab. Eng.* 8, 554–577. doi: 10.1016/j.ymben.2006.05.006

Nöh, K., Grönke, K., Luo, B., Takors, R., Oldiges, M., and Wiechert, W. (2007). Metabolic flux analysis at ultra short time scale: isotopically nonstationary 13C labeling experiments. *J. Biotechnol.* 129, 249–267. doi: 10.1016/j.jbiotec.2006.11.015

Notebaart, R. A., Van Enckevort, F. H., Francke, C., Siezen, R. J., and Teusink, B. (2006). Accelerating the reconstruction of genome-scale metabolic networks. *BMC Bioinformatics* 7:296. doi: 10.1186/1471-2105-7-296

Oberhardt, M. A., Puchałka, J., Fryer, K. E., Dos Santos, V. A. M., and Papin, J. A. (2008). Genome-scale metabolic network analysis of the opportunistic pathogen *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* 190, 2790–2803. doi: 10.1128/JB.01583-07

Oh, Y.-K., Palsson, B. O., Park, S. M., Schilling, C. H., and Mahadevan, R. (2007). Genome-scale reconstruction of metabolic network in *Bacillus subtilis* based on high-throughput phenotyping and gene essentiality data. *J. Biol. Chem.* 282, 28791–28799. doi: 10.1074/jbc.M703759200

Orth, J. D., Thiele, I., and Palsson, B. Ø. (2010). What is flux balance analysis? *Nat. Biotechnol.* 28, 245–248. doi: 10.1038/nbt.1614

Orth, J. D., Conrad, T. M., Na, J., Lerman, J. A., Nam, H., Feist, A. M., et al. (2011). A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism-2011. *Mol. Syst. Biol.* 7:e535. doi: 10.1038/msb.2011.65

Park, J. H., Lee, K. H., Kim, T. Y., and Lee, S. Y. (2007). Metab. Eng. of *Escherichia coli* for the production of L-valine based on transcriptome analysis and *in silico* gene knockout simulation. *Proc. Natl. Acad. Sci. U.S.A.* 104, 7797–7802. doi: 10.1073/pnas.0702609104

Pharkya, P., and Maranas, C. D. (2006). An optimization framework for identifying reaction activation/inhibition or elimination candidates for overproduction in microbial systems. *Metab. Eng.* 8, 1–13. doi: 10.1016/j.ymben.2005.08.003

Price, N. D., Reed, J. L., and Palsson, B. Ø. (2004). Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat. Rev. Microbiol.* 2, 886–897. doi: 10.1038/nrmicro1023

Röling, W. F., and Van Bodegom, P. M. (2014). Toward quantitative understanding on microbial community structure and functioning: a modeling-centered approach using degradation of marine oil spills as example. *Front. Microbiol* 5:125. doi: 10.3389/fmicb.2014.00125

Röling, W. F., Ferrer, M., and Golyshin, P. N. (2010). Systems approaches to microbial communities and their functioning. *Curr. Opin. Biotechnol.* 21, 532–538. doi: 10.1016/j.copbio.2010.06.007

Rühl, M., Hardt, W.-D., and Sauer, U. (2011). Subpopulation-specific metabolic pathway usage in mixed cultures as revealed by reporter protein-based 13C analysis. *Appl. Environ. Microbiol.* 77, 1816–1821. doi: 10.1128/AEM.02696-10

Raghoebarsing, A. A., Pol, A., Van de Pas-Schoonen, K. T., Smolders, A. J., Ettwig, K. F., Rijpstra, W. I. C., et al. (2006). A microbial consortium couples anaerobic methane oxidation to denitrification. *Nature* 440, 918–921. doi: 10.1038/nature04617

Raman, K., and Chandra, N. (2009). Flux balance analysis of biological systems: applications and challenges. *Brief. Bioinformatics* 10, 435–449. doi: 10.1093/bib/bbp011

Reed, J. L. (2012). Shrinking the metabolic solution space using experimental datasets. *PLoS Comput. Biol.* 8:e1002662. doi: 10.1371/journal.pcbi.1002662

Rodríguez, J., Kleerebezem, R., Lema, J. M., and van Loosdrecht, M. (2006). Modeling product formation in anaerobic mixed culture fermentations. *Biotechnol. Bioeng.* 93, 592–606. doi: 10.1002/bit.20765

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. doi: 10.1002/biot.201000159

Santos, F., Boele, J., and Teusink, B. (2011). A practical guide to genome-scale metabolic models and their analysis. *Meth. Enzymol.* 500, 509. doi: 10.1016/B978-0-12-385118-5.00024-4

Shaikh, A. S., Tang, Y. J., Mukhopadhyay, A., and Keasling, J. D. (2008). Isotopomer distributions in amino acids from a highly expressed protein as a proxy for those from total protein. *Anal. Chem.* 80, 886–890. doi: 10.1021/ac071445+

Stolyar, S., Van Dien, S., Hillesland, K. L., Pinel, N., Lie, T. J., Leigh, J. A., et al. (2007). Metabolic modeling of a mutualistic microbial community. *Mol. Syst. Biol.* 3:e92. doi: 10.1038/msb4100131

Sun, J., Sayyar, B., Butler, J., Pharkya, P., Fahland, T., Famili, I., et al. (2009). Genome-scale constraint-based modeling of *Geobacter metallireducens*. *BMC Syst. Biol.* 3:15. doi: 10.1186/1752-0509-3-15

Szyperski, T. (1998). 13C-NMR, MS and metabolic flux balancing in biotechnology research. *Q. Rev. Biophys.* 31, 41–106. doi: 10.1017/S0033583598003412

Teusink, B., Wiersma, A., Molenaar, D., Francke, C., de Vos, W. M., Siezen, R. J., et al. (2006). Analysis of growth of *Lactobacillus plantarum* WCFS1 on a complex medium using a genome-scale metabolic model. *J. Biol. Chem.* 281, 40041–40048. doi: 10.1074/jbc.M606263200

Teusink, B., Wiersma, A., Jacobs, L., Notebaart, R. A., and Smid, E. J. (2009). Understanding the adaptive growth strategy of *Lactobacillus plantarum* by *in silico* optimisation. *PLoS Comput. Biol.* 5:e1000410. doi: 10.1371/journal.pcbi.1000410

Thiele, I., Swainston, N., Fleming, R. M., Hoppe, A., Sahoo, S., Aurich, M. K., et al. (2013). A community-driven global reconstruction of human metabolism. *Nat. Biotechnol.* 31, 419–425. doi: 10.1038/nbt.2488

Toya, Y., and Shimizu, H. (2013). Flux analysis and metabolomics for systematic Metab. Eng. of microorganisms. *Biotechnol. Adv.* 31, 818–826. doi: 10.1016/j.biotechadv.2013.05.002

van Heerden, J. H., Wortel, M. T., Bruggeman, F. J., Heijnen, J. J., Bollen, Y. J., Planqué, R., et al. (2014). Lost in transition: start-up of glycolysis yields subpopulations of nongrowing cells. *Science* 343:1245114. doi: 10.1126/science.1245114

Van Gulik, W., and Heijnen, J. (1995). A metabolic network stoichiometry analysis of microbial growth and product formation. *Biotechnol. Bioeng.* 48, 681–698. doi: 10.1002/bit.260480617

Varma, A., and Palsson, B. O. (1994). Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Appl. Environ. Microbiol.* 60, 3724–3731.

Wahl, S. A., Nöh, K., and Wiechert, W. (2008). 13C labeling experiments at metabolic nonstationary conditions: an exploratory study. *BMC Bioinformatics* 9:152. doi: 10.1186/1471-2105-9-152

Wisselink, H. W., Cipollina, C., Oud, B., Crimi, B., Heijnen, J. J., Pronk, J. T., et al. (2010). Metabolome, transcriptome and metabolic flux analysis of arabinose fermentation by engineered *Saccharomyces cerevisiae*. *Metab. Eng.* 12, 537–551. doi: 10.1016/j.ymben.2010.08.003

Zamboni, N. (2011). 13C metabolic flux analysis in complex systems. *Curr. Opin. Biotechnol.* 22, 103–108. doi: 10.1016/j.copbio.2010.08.009

Zengler, K., and Palsson, B. O. (2012). A road map for the development of community systems (CoSy) biology. *Nat. Rev. Microbiol.* 10, 366–372. doi: 10.1038/nrmicro2763

Zhuang, K., Izallalen, M., Mouser, P., Richter, H., Risso, C., Mahadevan, R., et al. (2010). Genome-scale dynamic modeling of the competition between Rhodoferax and Geobacter in anoxic subsurface environments. *ISME J.* 5, 305–316. doi: 10.1038/ismej.2010.117

Glossary

Community-scale stoichiometric model (CSSM): Stoichiometric model that describes the microbial community.

Community state: The full set of concentrations, abundances of species and process rates within a community.

Community structure: Topology of the network of species interactions, as selected by the environment.

Flux Balance Analysis (FBA): Method used to calculate flux distributions through the cell. Measured flux data are not required, so FBA can be used purely computational. FBA is performed on genome-scale stoichiometric metabolic networks.

Isotopomer: Molecules which have the same constitution and the same configuration but differ in isotope substitution.

Isotopomer-based flux analysis: Advanced method for Metabolic Flux Analysis, using isotopomer data to infer intracellular fluxes from isotopomer enriched metabolites.

Metabolic Flux Analysis (MFA): Method to infer fluxes from measured flux data. The stoichiometric models are coarse-grained and are used by the experimentalists.