INTERFACE

Headline review

CrossMark
click for updates

Author for correspondence:
Bas Teusink
e-mail: b.teusink@vu.nl

Constraint-based stoichiometric modelling from single organisms to microbial communities

Willi Gottstein, Brett G. Olivier, Frank J. Bruggeman and Bas Teusink

Systems Bioinformatics, Amsterdam Institute for Molecules, Medicines and Systems, VU University Amsterdam, De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands

(iD) WG, 0000-0002-6502-1393

Microbial communities are ubiquitously found in Nature and have direct implications for the environment, human health and biotechnology. The species composition and overall function of microbial communities are largely shaped by metabolic interactions such as competition for resources and cross-feeding. Although considerable scientific progress has been made towards mapping and modelling species-level metabolism, elucidating the metabolic exchanges between microorganisms and steering the community dynamics remain an enormous scientific challenge. In view of the complexity, computational models of microbial communities are essential to obtain systems-level understanding of ecosystem functioning. This review discusses the applications and limitations of constraint-based stoichiometric modelling tools, and in particular flux balance analysis (FBA). We explain this approach from first principles and identify the challenges one faces when extending it to communities, and discuss the approaches used in the field in view of these challenges. We distinguish between steady-state and dynamic FBA approaches extended to communities. We conclude that much progress has been made, but many of the challenges are still open.

1. Introduction

Microbial communities are ubiquitous and of great interest for biotechnological applications [1,2], health [3–6], food production [7] and environmental studies [8]. Key questions in the field of microbial ecology are: 'Who are the community members?', 'What are they doing?', 'How do they interact?' and 'What functionality arises from these interactions?'. Metagenome sequencing is rapidly answering the first question [9–12], and this makes the other questions all the more pressing. Great strides have been made in analysing metabolic interactions between the members of microbial communities [13–17]. Despite this progress, the principles that shape the structure of microbial communities remain largely elusive, making it challenging to understand, describe, predict and (ultimately) design communities with particular behaviour.

Here, we review modelling approaches that take genome information as the primary input and aim to predict community behaviour from a metabolic perspective. Thus, we do not discuss interactions between microorganisms that have no direct metabolic component, such as physical interactions or signalling processes (quorum sensing, for example). The approaches we describe are all constraint-based, stoichiometric, modelling methods for metabolic networks. The associated models are essentially mappings of genes to metabolic enzymes, the reactions of which form a (genome-scale) metabolic network where the edges and nodes represent (enzyme-catalysed) reactions and metabolites, respectively.

Constraint-based modelling has been developed for sequenced microbes in monoculture, and many methods have been developed to reconstruct, query and predict the metabolic capacities of a species based on its genome

sequence [18–24]. The most widely used method is flux balance analysis (FBA) [25,26]. Given a biologically relevant objective function and constraints on input/output rates (e.g. glucose consumption, carbon dioxide excretion), FBA predicts steady-state flux distributions that optimize the chosen objective—typically biomass formation, but also multi-dimensional objectives exist [27–29]. Under the right premises, to be discussed later, these predictions are often surprisingly accurate—which is remarkable considering that these models do not take kinetic information into account [24]. When metagenome sequencing became feasible for entire communities [9–12], the development of constraint-based stoichiometric modelling approaches for microbial communities has become attractive.

As we explain later in more detail, the results of FBA-type approaches mainly depend on (i) the quality of the reconstructed metabolic model, including the biomass composition of the associated microorganism and the assumed gene–protein reaction rules, (ii) the chosen objective function, (iii) the considered constraints for exchange and intracellular rates, and (iv) the considered nutrients that can be used for biomass and product formation. If FBA is extended to analyse metabolism on a community level, then additional layers of complexity are added to all four of these points.

With respect to reconstructions, despite the progress that has been made in automating the reconstruction of metabolic models on the genome scale [30–34], model reconstruction and manual curation is still a very time-consuming process. In particular, detailed biomass compositions are available only for a few microbial species, and their experimental determination requires the isolation and culturing of the species. If one misses physiological characteristics, e.g. certain auxotrophies or transport reactions in the reconstruction process, then it will have a considerable impact on the predicted interactions between organisms.

Once a genome-scale model has been reconstructed, it is essential for community modelling that it can be combined with other models in a meaningful way. The construction of community models is complicated by the fact that, in most cases, such models will have been created by different authors and research groups. Consider, for example, that all models should be encoded to be readable by the same simulation software; model components (reactions, metabolites, flux capacity constraints) must be unambiguously annotated such that they can be used to link models together and community models themselves should be reproducible and exchangeable. Fortunately, with regards to model interoperability, the constraint-based modelling community has recently adopted a standard model encoding format that allows for efficient model exchange [35]. However, the problem of seamless integration, when using models collected from different sources, currently remains open.

Another difficulty is to define a community objective function in both biological and mathematical terms; typically, a combination of the objective functions of individual species is optimized but the exact formulation can differ [15,17,36–39]. This question touches upon fundamental questions in evolutionary biology about group selection versus selection at the level of individuals.

Regarding exchange rates, at a monoculture level, exchange rates can be calculated from extracellular concentration data, and intracellular fluxes can be estimated using, for example, ^{13}C-labelling experiments [40,41]. Such data can then be used as constraints for FBA. Both become much more complicated on
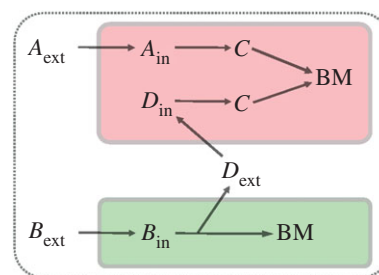

Figure 1. Growth potential depends on the composition of the medium and the metabolic capabilities of members in the community. Organism 1 (red) can only use A_{ext} that can be converted into a precursor of biomass, C. If it grew alone, then it could not make use of B_{ext}. When organism 2 (green) is present, which is capable of using B_{ext} and thereby excreting a growth coupled compound D, organism 1 can increase its biomass production.

a community level; it is not trivial to determine individual contributions of microbial species to measured net fluxes of extracellular compounds, making it difficult to define appropriate constraints on the species level. There are isotopic labelling methods for community-level quantification of fluxes, and meta-transcriptomics and metaproteomics data may be translated to flux constraints under simplifying assumptions [42–48], but these have not reached a sufficiently quantitative level yet.

On a single-species level, the growth potential of organisms is directly determined by the composition of the medium. This is not so on a community level where the available resources and the metabolic capabilities of other community members determine the metabolism and growth behaviour of particular species. Even if species cannot use certain compounds available in their environment directly, they might be able to use those indirectly through the metabolic activity of other members in the community, as illustrated for a minimal example in figure 1. Interactions between species are therefore context and medium dependent [49–54]. Thus, at the moment, models are required to predict and quantify these (hidden) metabolic interactions between microbial species, given different environmental conditions and objectives, as we can presently not measure them.

In this review, we discuss the several methods that have been developed [14,15,17] to use FBA and variants thereof on a community level. The first work using genome-scale models for a microbial community was published in 2007 [36]. Since then, it has been used to estimate interspecies fluxes as well as intracellular flux distributions [36], to classify metabolic interactions [13], to predict the compositions of media that induce interactions between members of a community [55], to predict optimal relative biomass abundances [39] and many more, most of which are summarized in [17]. We view these studies in the light of the complications that we have identified. For a thorough understanding of these issues, however, we first need to derive the classical FBA approach from its basic principles before we move on to the specifics of community FBA (cFBA).

2. Basic principles and assumption of single-organism flux balance analysis: the foundation of microbial-community flux balance analysis

The foundations of classical FBA were developed in the 1980s [56–58] and today it is one of the most used methods to

study metabolism [24–26]. By making assumptions about an objective function (typically biomass formation) and using information about transport fluxes (e.g. sugar uptake, product formation) and intracellular rates, FBA can be used to determine flux distributions that optimize the respective objective; kinetic parameters are not required. FBA and all of its variants rely on a metabolic steady state. Hence, it strictly applies only to cell populations displaying balanced growth, i.e. cells in batch culture growing in the exponential phase, or cells cultured in a chemostat, where the concentration of each metabolic intermediate and the growth rate are constant.

The rationale of FBA is as follows: one considers all metabolic reactions in the metabolic network, including those reactions leading to the macromolecular components of cells such as proteins, RNA, lipids and DNA (or their respective building blocks) and one that specifies the macromolecular component composition of cells, using an artificial biomass-forming reaction. Next, one demands a metabolic steady state, which then directly leads to a constant rate at which new biomass is synthesized. Hence, the balanced growth condition is modelled. To make this explicit and facilitate the discussion of modelling microbial communities, in the following classical FBA is derived from basic principles.

The starting point of this derivation is the definition of the concentration c_i of a metabolic intermediate i in the metabolic network with $c_i = n_i/V$, where n_i is the number of molecules i and V is the cell volume. V and n_i are defined for the entire population of cells; V is the total cell volume and n_i is the total amount of molecules i. Owing to growth, both variables are time dependent, and the temporal change of c_i is given by applying the chain rule for differentiation

$$\frac{\mathrm{d}}{\mathrm{d}t}c_i = \frac{1}{V}\frac{\partial n_i}{\partial t} - \frac{n_i}{V^2}\frac{\partial V}{\partial t}. \tag{2.1}$$

When the volume is constant, i.e. $\partial V/\partial t = 0$, the concentration change is proportional to the change in the number of molecules. When the number of molecules is fixed, $\partial n_i/\partial t$ evaluates to 0 and c_i can be increased when the volume is decreased and vice versa. At steady state, the temporal change of c_i is 0, i.e. $(\mathrm{d}/\mathrm{d}t)c_i = 0$. This yields for equation (2.1)

$$\frac{1}{V}\frac{\partial n_i}{\partial t} = \frac{n_i}{V^2}\frac{\partial V}{\partial t}. \tag{2.2}$$

Multiplying equation (2.2) by V/n_i allows us to define the specific growth rate μ at steady state as

$$\frac{1}{n_i}\frac{\partial n_i}{\partial t} = \frac{1}{V}\frac{\partial V}{\partial t} := \mu. \tag{2.3}$$

Rewriting equation (2.2) by using equation (2.3) illustrates that the synthesis of new molecules equals their dilution by (volume) growth at steady state

$$\frac{1}{V}\frac{\partial n_i}{\partial t} = \mu\frac{n_i}{V} = \mu \cdot c_i. \tag{2.4}$$

At a constant volume (so in the absence of volume growth), the rates of the reactions v_j occurring in metabolism relate to equation (2.4) in the following way:

$$\frac{1}{V}\frac{\partial n_i}{\partial t} = \frac{\mathrm{d}}{\mathrm{d}t}c_i = \sum_j s_{ij}v_j, \tag{2.5}$$

whereby s_{ij} are stoichiometric coefficients which are positive/negative if metabolite i is produced/consumed in reaction j. Equation (2.5) means that the temporal change of c_i is

determined by fluxes v_j that can either produce or degrade metabolite i. The vectorized version of equation (2.5) then reads

$$\frac{\mathrm{d}c}{\mathrm{d}t} = S \cdot v, \tag{2.6}$$

where c is the vector of concentrations, v is a vector of rates and S is the stoichiometric matrix with m rows and r columns representing metabolites and reaction rates, respectively. It contains the stoichiometric coefficients s_{ij}.
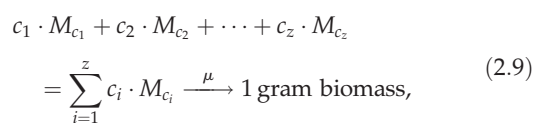
2.1. Biomass and growth dilution

In FBA, one now defines two classes of metabolites that appear in the stoichiometric matrix S: molecules that are required for biomass formation (metabolic end-products such as DNA, RNA, protein, membranes or their respective building blocks) and intermediates which are not. These two classes are treated differently regarding the dilution by growth (equation (2.4)). It is neglected for intermediates as the metabolic rates of intermediate conversions are assumed to be much higher in value than $\mu \cdot c_i$. For biomass components, however, dilution by growth is taken as their main sink, and, therefore, equations (2.4) and (2.5) become

$$\sum_j s_{ij}v_j = \frac{\mathrm{d}}{\mathrm{d}t}c_i = 0, \quad \forall i \in \text{intermediates} \tag{2.7}$$

and

$$\sum_j s_{ij}v_j = \frac{\mathrm{d}}{\mathrm{d}t}c_i = \mu \cdot c_i, \quad \forall i \in \text{biomass components}, \tag{2.8}$$

which means that intermediate concentrations do not change over time, i.e. their production rates equal their consumption rates, whereas biomass components accumulate exponentially with rate μ, reflecting the exponential increase of biomass. In stoichiometric models, the fluxes v_j are usually expressed in mmol $(\mathrm{h} \cdot \mathrm{g})^{-1}$ and growth rate μ has the unit g $(\mathrm{g} \cdot \mathrm{h})^{-1}$ reflecting the increase of biomass per biomass per hour. Considering this, one sees that c_i has the unit mmol g^{-1} (equation (2.8)). Hence, c_i is the amount of biomass component i per gram of biomass. A biomass-forming reaction v_{biomass} can be defined as

$$c_1 \cdot M_{c_1} + c_2 \cdot M_{c_2} + \cdots + c_z \cdot M_{c_z}$$
$$= \sum_{i=1}^{z} c_i \cdot M_{c_i} \xrightarrow{\mu} 1\,\text{gram biomass}, \tag{2.9}$$

where M_{c_i} represents one of the z biomass components (dimensionless) and the coefficients c_i the contribution of biomass component i to 1 g of biomass with unit mmol g^{-1}.

Summarizing, FBA involves the following mass-balance constraints, with the rates v_j as unknowns:

$$\left.\begin{array}{l} \sum_j s_{ij}v_j = 0, \quad \forall i \in \text{intermediates} \\[2mm] \text{and} \quad \sum_j s_{ij}v_j = \mu \cdot c_i, \quad \forall i \in \text{biomass components}, \end{array}\right\} \tag{2.10}$$

or in a more compact form

$$S \cdot v = 0, \tag{2.11}$$

whereby the flux vector v contains v_{biomass} and the stoichiometric matrix S a corresponding column that contains 0s in rows that correspond to intermediates and the factors c_i in rows that correspond to biomass components. So-called boundary species (S_{env} and P_{env} in figure 2) do not appear in
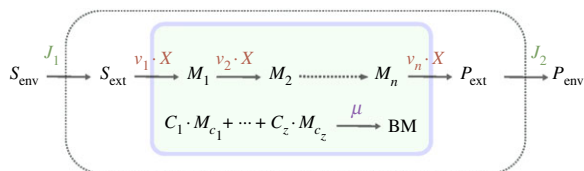
Figure 2. Illustration of the different rates in FBA. One distinguishes environment fluxes J_i with unit mmol h^{-1}, specific rates v_i with unit mmol (h · g)$^{-1}$ and the organism's specific growth rate in h^{-1}. X denotes the organism's biomass in unit g, M_{c_i} are components required for biomass formation and their respective contribution to 1 g of biomass is denoted by c_i with unit mmol g^{-1}.

the stoichiometric matrix; they just serve as parameters and do not have any influence on the actual simulations. In some model formulations, they are omitted, and the environmental fluxes J are expressed as 'hyperspace' reactions. The net units on the left- and right-hand sides of the relations in equation system (2.10) are per unit 'gram cells'. So they remain valid when the biomass increases owing to growth. These relations are therefore the balanced growth condition that was mentioned earlier. The rate of biomass increase equals $dX/dt = \mu \cdot X$ with X as biomass in gram cells. Biomass therefore does not need to be considered explicitly in FBA on monocultures. This is different for communities, as we see!

2.2. Flux distributions have variability

As there are typically more unknown fluxes than metabolites, the linear system defined in equation (2.11) is usually under-determined. The solution space can be reduced (i) by constraining individual rates with so-called capacity constraints, using information about reversibility of reactions and constraints based on experimental data, such as uptake rates, and (ii) by optimizing a particular objective as, for example, v_{biomass}, using linear programming. The optimization problem in FBA can be expressed in a compact form as follows:

$$\left. \begin{array}{c} \max_{v} Z = w^{\mathrm{T}} \cdot v, \\ s.t.\ S \cdot v = 0 \\ \text{and} \qquad v_{\min} \leq v \leq v_{\max}, \end{array} \right\} \quad (2.12)$$

where Z represents the objective function and is expressed as a linear combination of fluxes v with weights contained in vector w and has been discussed extensively in [27,28]; typically, growth rate is maximized, thus $Z = v_{\text{biomass}}$. The $v_{\min}$ and $v_{\max}$ contain the lower and upper bounds of the rates; if a reaction is irreversible, then the corresponding lower bound is set to 0.

Usually, multiple flux distributions exist that optimize the respective objective function. The flexibility of individual fluxes can be examined by using flux variability analysis (FVA), where each flux is minimized and maximized, using the optimal value of the objective function as an additional constraint [59–62]. FVA can be formulated as follows:

$$\left. \begin{array}{c} \max/\min v_i, \quad \forall v_i \in v \\ \text{and} \qquad s.t.\ S \cdot v = 0 \\ v_{\min} \leq v \leq v_{\max} \\ \text{and} \qquad Z = Z_{\text{opt}}, \end{array} \right\} \quad (2.13)$$

whereby Z_{opt} is determined, using FBA. More advanced methods exist to study all the alternative optimal flux distributions [63,64].

2.2.1. A closer look at the units in classical flux balance analysis

In FBA, one can distinguish three different rates, which are depicted in figure 2: environmental (exchange) fluxes J with unit mmol h^{-1}, specific rates v with unit mmol (h · g)$^{-1}$ and a biomass-forming reaction v_{biomass} occurring at the specific growth rate μ which has the unit h^{-1}. In the following, we explain how these different rates are connected, using the small example system from figure 2.

The external metabolite, S_{ext}, is made available at a rate of $J_1(t)$ and is consumed by the organism at a rate of $v_1 \cdot X(t)$. The temporal change of S_{ext} is therefore given by

$$\frac{\mathrm{d}S_{\text{ext}}}{\mathrm{d}t} = J_1(t) - v_1 \cdot X(t), \quad (2.14)$$

where $X(t)$ denotes the biomass of the organism. In most FBA formulations, S_{ext} is also balanced by uptake and exchange reactions, and because the biomass increases exponentially the steady-state solution to equation (2.14) must read

$$\frac{\mathrm{d}S_{\text{ext}}}{\mathrm{d}t} = 0 = J_1(0) \cdot \mathrm{e}^{\mu \cdot t} - v_1 \cdot X(0) \cdot \mathrm{e}^{\mu \cdot t}.$$

Normalizing this expression with respect to the total biomass given by

$$X_{\text{tot}}(t) = X_{\text{tot}}(0) \cdot \mathrm{e}^{\mu \cdot t},$$

one obtains for the specific rate v_1

$$\frac{\mathrm{d}S_{\text{ext}}}{\mathrm{d}t} = 0 = \frac{J_1(0) \cdot \mathrm{e}^{\mu \cdot t}}{X_{\text{tot}}(t)} - v_1 \cdot \frac{X(0) \cdot \mathrm{e}^{\mu \cdot t}}{X_{\text{tot}}(t)}$$

$$v_1 = v_{e_s},$$

where

$$v_{e_s} := \frac{J_1(0)}{X_{\text{tot}}(0)} \quad (2.15)$$

is the *specific* exchange flux with the environment. v_{e_s} has the unit mmol (h · g)$^{-1}$ consistent with the units of the remaining rates.

For all internal metabolites, the biomass cancels out at steady state, as shown here for metabolite M_1

$$\left. \begin{array}{c} \dfrac{\mathrm{d}M_1}{\mathrm{d}t} = v_1 \cdot X(t) - v_2 \cdot X(t) \\ \text{and} \quad \dfrac{\mathrm{d}M_1}{\mathrm{d}t} = 0 = v_1 - v_2. \end{array} \right\} \quad (2.16)$$

Note that this is not the usual way that FBA and the role of exchange fluxes are presented: rather, exchange fluxes v_e are often 'just' assigned for every external metabolite under the premise of fixed external metabolite concentrations in steady state. Yet, strictly speaking, S_{ext} can only remain constant under exponential growth if the feeding rate increases exponentially with biomass. It illustrates an important point often taken for granted: in FBA, we fix some input flux(es) through capacity constraints. In real systems, this can be achieved in different ways. In fed-batch growth, the feeding rate of S_{ext} is indeed increased exponentially to keep its concentration constant during exponential growth. In the chemostat, we use dilution to achieve a steady state of biomass and nutrient levels; in normal batch growth, we define a region where the inevitable changes in external metabolites do not affect the uptake (and production) rates. For substrates taken up, this implies saturation—or nutrient excess. As we see, a clear distinction of the units of the different fluxes will be essential to understand the FBA approaches applied to communities.

2.2.2. Limitations of single-organism flux balance analysis

FBA has been proven to be a useful tool in several contexts, such as strain design [65–67], the outcome prediction of evolution studies [68,69] and gene deletion studies [70–72] (many more can be found in various reviews and references therein [24–26]). However, because of its stoichiometric nature, and because it uses optimization, FBA comes with inherent limitations. Most prominently, one should be aware of the fact that FBA predicts flux distributions that provide the maximal yield on the limiting nutrient, and does not actually predict kinetic rates. If one sets the uptake rate of a nutrient, e.g. v_{glucose}, to a certain value and μ is maximized by finding optimal values of the remaining v_j, then the fluxes that maximize the ratio μ/v_{glucose} are found. This expression is defined as a yield, $Y_{\mathrm{biomass/glucose}}$, with unit grams of biomass per mol glucose. While there are examples where the predicted flux distributions are in good agreement with experimental data [68,69,73], there is poor agreement under conditions of nutrient excess that favour low-yield strategies such as fermentation [74–76].

As mentioned earlier, owing to the assumption of time-invariant extracellular conditions, classical FBA can only be applied to chemostat cultures and cells in batches that grow exponentially, i.e. populations that display balanced growth. External fluctuations of metabolite concentrations cannot be incorporated. Furthermore, owing to the linear nature of the system, FBA neither allows absolute concentrations to be predicted nor incorporates saturation effects. In addition, regulatory effects cannot be examined. To overcome some of the limitations, classical FBA has been extended in recent years [77–84]. The most exciting extensions from a community point of view are dynamic FBA (dFBA) [85]—discussed in more detail later—and very recent methods that couple metabolic fluxes to enzyme synthesis costs, thus allowing growth-rate-dependent switches in metabolic fluxes according to the principle of limited resource allocation [82–84,86,87]. The latter approaches have not been used in community modelling as far as we are aware, but show much more realistic behaviour than classical FBA.

2.2.3. Summary of the main assumptions for classical flux balance analysis

FBA allows steady-state flux distributions to be determined given a genome-scale stoichiometric model, flux capacity constraints (e.g. sugar uptake) and an objective function (e.g. growth rate). One distinguishes two classes of relevant metabolites for which different mass-balance constraints apply: intermediates for which dilution by growth is not considered and biomass components for which growth dilution is the main sink (equations (2.7), (2.8), (2.10)). There are three different kinds of rates: environmental exchange fluxes with unit mmol h^{-1}, specific rates with unit mmol (h·g)$^{-1}$ and a biomass-forming reaction occurring at the specific growth rate μ which has the unit h^{-1} (figure 2). In classical FBA, biomass does not have to be taken into account explicitly when uptake rates are calculated (§2.1) which is a major difference from FBA on the community level as discussed in the following sections.

3. Steady-state community flux balance analysis

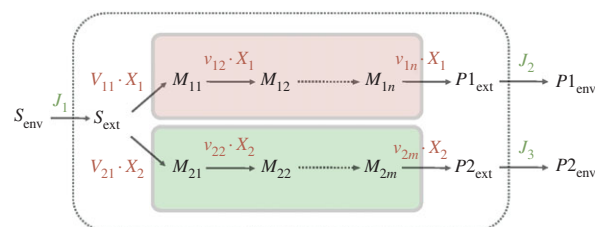In Nature, microorganisms usually do not occur in monocultures but are rather organized in communities where they

Figure 3. Two organisms live in the same environment competing for a substrate S_{ext}. X_i represent their respective biomass, J_i are environment fluxes with unit mmol h^{-1} and v_i are specific rates with unit mmol (h·g)$^{-1}$. Internal metabolites are denoted by M_{ij}.

interact in various ways [88]. Until now, it has remained a great challenge to qualify and quantify metabolic interactions between community members and to understand how these influence the community structure and its dynamics. For this purpose, several approaches have been developed [14,15,17], one of which is FBA on a community level where stoichiometric models of different species are connected [14,17,36,89]. Such a meta-species network allows for an analogue mathematical representation as described for classical FBA. The choice of an appropriate objective function, however, is even less obvious on the community level than it is for monocultures as there has been a long-standing debate on which level natural selection actually occurs [90–97].

Nevertheless, in brave attempts to use the time-invariant constraint-based formalism for communities, several expressions for an objective function have been proposed, such as linear combinations of individual species objectives [36], a community growth rate meaning that all species grow at the same rate [39] as well as a bilevel objective function where individual as well as community objectives are taken into account [37]. We now extend the mathematical formulation of classical FBA to describe the steady-state growth of communities.

For this, a community is considered that consists of only two members, as depicted in figure 3; a more generalized derivation can be found in [39].

First, the steady-state concentration of the external metabolite S_{ext} is considered (the same formalism applies to $P1_{\mathrm{ext}}$ and $P2_{\mathrm{ext}}$). S_{ext} is made available by an environmental flux J_1 and taken up by both organisms. These uptake rates are proportional to their biomass: X_1 and X_2, respectively. The differential equation for S_{ext} therefore reads as follows:

$$\frac{\mathrm{d}S_{\mathrm{ext}}}{\mathrm{d}t} = J_1(t) - v_{11} \cdot X_1(t) - v_{21} \cdot X_2(t),$$

where v_{ij} is the specific uptake/production rate of compound j by species i. At steady state, the temporal change of S_{ext} is 0 and the organisms' biomass increases exponentially with their respective growth rates

$$X_i(t) = X_i(0) \cdot \mathrm{e}^{\mu_i \cdot t},$$

so that one obtains

$$\frac{\mathrm{d}S_{\mathrm{ext}}}{\mathrm{d}t} = 0 = J_1(t) - v_{11} \cdot X_1(0) \cdot \mathrm{e}^{\mu_1 \cdot t} - v_{21} \cdot X_2(0) \cdot \mathrm{e}^{\mu_2 \cdot t}. \quad (3.1)$$

The clearest solution to equation (3.1) can be derived if both organisms grow at the same rate and $J_1(t)$ also increases exponentially with that same rate

$$\frac{\mathrm{d}S_{\mathrm{ext}}}{\mathrm{d}t} = 0 = J_1(0) \cdot \mathrm{e}^{\mu \cdot t} - v_{11} \cdot X_1(0) \cdot \mathrm{e}^{\mu \cdot t} - v_{21} \cdot X_2(0) \cdot \mathrm{e}^{\mu \cdot t}.$$

Normalizing this expression with respect to the total biomass in the system given by

$$X_{tot}(t) = \sum_{i=1}^{S} X_i(0) \cdot e^{\mu \cdot t} = e^{\mu \cdot t} \cdot \sum_{i=1}^{S} X_i(0) = X_{tot}(0) \cdot e^{\mu \cdot t}$$

yields

$$\frac{dS_{ext}}{dt} = 0 = \frac{J_1(0)}{X_{tot}(0)} - v_{11} \cdot \frac{X_1(0)}{X_{tot}(0)} - v_{21} \cdot \frac{X_2(0)}{X_{tot}(0)}.$$

When we introduce

$$\phi_i := \frac{X_i(0)}{X_{tot}(0)}$$

and

$$v_{e_j} := \frac{J_j(0)}{X_{tot}(0)},$$

for the relative biomass abundances and the specific environmental exchange fluxes, respectively, we obtain the final expression

$$v_{e_1} = v_{11} \cdot \phi_1 - v_{21} \cdot \phi_2. \tag{3.2}$$

Equation (3.2) shows that biomass abundances have to be taken into account when the balances of external metabolites are calculated. This is a major difference from classical FBA where the relative biomass abundance is always 1 by definition (equation (2.15)). For all internal metabolites biomass abundances do not matter as seen before (equation (2.16)) and cancel out as demonstrated for M_{11}

$$\left.\begin{array}{l} \dfrac{dM_{11}}{dt} = v_{11} \cdot X(t) - v_{12} \cdot X(t) \\[2mm] \text{and} \quad \dfrac{dM_{11}}{dt} = 0 = v_{11} - v_{12}. \end{array}\right\} \tag{3.3}$$

This derivation assumed equal growth rates for both organisms. In many cases, this is the only sensible steady state that can be achieved by communities. For communities in chemostats, equal growth rates are obvious, as the dilution rate sets the steady-state growth rate for each organism independently. In fed-batch growth, however, one might consider organisms with different growth rates, combined with a double-exponential feeding regime to keep the nutrients constant, following equation (3.1). In addition, during batch growth, in analogy with monoculture batch FBA, there may be a window in time where all organisms are in balanced growth, i.e. their internal components are in steady state, even when the external concentrations are not. However, this is provided that there is no cross-feeding present. If the two organisms were to have different growth rates, the one with the highest growth rate would rapidly outgrow the other. If this fast-growing organism is dependent on a factor D secreted by the slow organism (as in figure 1), its growth will soon start to become limited by the relatively slow production of D, according to the balance

$$\frac{dD}{dt} = J(t) - v_{1D} \cdot X_1(t) + v_{2D} \cdot X_2(t). \tag{3.4}$$

Here, the signs indicate that species 1 consumes D produced by species 2, and J is the exchange rate of D. A small fraction of species 2 will result in slow production of D, and, unless D is provided from the environment, D will decrease until it affects the specific uptake rate of species 1, i.e. v_{1D}. At that point, growth rate will start to decrease for species 1, and the system will settle into a global steady state only when the growth

rates are equal. Thus, using steady-state FBA for modelling communities is allowed even if organisms grow at different rates, as long as within that regime the conditions allow constant uptake and production rates. For all practical purposes, that means saturation of the uptake systems. For metabolites that are exchanged such as D, saturation is rather unlikely; hence, these models can only represent short snapshots of the system. Either one can resort to dFBA, which glues many such snapshots together (as we discuss later), or one sticks to equal growth rates, as is done in cFBA [39], as discussed in §3.1.

3.1. Community flux balance analysis: applications and limitations

cFBA has a number of advantages. First, it provides an unambiguous objective function for all consortium members: the identical growth rate. Moreover, imposing equal growth rates ensures that the relative biomass abundances are constant, and it was shown that these abundances affect the optimal growth rate [39]. Thus, these models predict species abundance ratios, something that can be readily determined experimentally by (metagenomics) sequencing. Steady-state cFBA is, however, mainly applicable to systems in fairly constant environments such as cells grown in the chemostat or cultures used for waste-water treatment.

It is currently unclear whether organisms in a community are actually capable of adjusting their metabolism to operate in an optimal manner: ideally, long-term community chemostats should be studied. In yoghurt fermentations, which are derived from serial transfer of two microorganisms (*Streptococcus thermophilus* and *Lactobacillus bulgaricus*), we do observe biomass abundances close to the predicted optimum (M Hanemaaijer *et al.* 2016, unpublished results). In our view, predictions made by steady-state cFBA should probably be best seen as idealized states that represent potential final states of evolutionary/adaptation processes. Its main applications are the same as for classical FBA in terms of qualitative exploration of medium compositions, metabolic engineering strategies and—most interestingly perhaps—the prediction of essential interactions between species given a certain medium composition. Additionally, cFBA gives an indication of the corresponding optimal relative biomass abundances.

Essential interactions can be identified by optimizing the system's community growth rate and by performing FVA (equation (2.3)) on all the transport reactions of shared metabolites. If the lower and upper values of a flux have the same sign, then it is unidirectional and can therefore be classified as essential. This way, one can easily identify feeding mechanisms and also competition for resources. Please note that in this framework no assumptions are made about the metabolic interactions between species but that these are a direct outcome of the simulations.

If a community of only two species is examined, then the optimal biomass abundances can be obtained by systematically scanning ratios of biomass abundances and calculating the corresponding maximal community growth rate. A plot of the optimal community growth rate versus the biomass ratio then identifies the optimal ratio of biomass abundances [39]. For communities of larger size, a systematic scan is not feasible and therefore optimization methods as gradient descent or evolutionary algorithms can be used with the constraint that the sum of all biomass fractions is 1.

As for classical FBA, steady-state cFBA also comes with certain limitations: absolute metabolite and biomass concentrations cannot be determined but only flux distributions that optimize the community growth rate. It is also important to keep in mind that the results obtained by this method highly depend on the quality of the reconstructed models, and especially their biomass compositions: these define the community growth rate. In addition, the correct identification of possible auxotrophies of individual species during the reconstruction is crucial as these give rise to potential interactions between organisms.

3.2. Other approaches to model steady-state behaviour of microbial communities

In most studies that have been published using FBA to study metabolism on a community level, equal growth rates are not assumed. For example, in the first paper on community FBA [36] a syntrophic two-species consortium was examined consisting of *Desulfovibrio vulgaris* and *Methanococcus maripaludis* and an objective function of the type

$$
\left.\begin{array}{l}
\max \ w_1 \cdot \mu_1 + w_2 \cdot \mu_2 \\
s.t. \ S \cdot v = 0 \\
\text{and} \quad v_{\min} \leq v \leq v_{\max}
\end{array}\right\} \tag{3.5}
$$

was used for different combinations of weights w_i. Similarly, pairwise interactions of 118 species were investigated and classified as neutral, competitive or cooperative based on a weighted sum of biomass production rates [13]. In [98], growth rates of *Leptospirillum ferriphilum* and *Ferroplasma acidiphilum* are examined in terms of a unidirectional feeding flux from *L. ferriphilum* to *F. acidiphilum* resulting in different specific growth rates. In [99], pairwise interactions of 11 gut bacteria were studied by iteratively fixing the growth rate of one organism and optimizing the other species' growth rate, whereby the resulting Pareto frontiers revealed four different types of interactions. In the OptCom framework, a nested objective function is used whereby a community-level objective is optimized (e.g. maximizing the total biomass of the community) subject to the single-species objectives (e.g. maximizing individual biomass production) [37]. Considering both species-level objective functions and a community-level objective allows cellular and interspecies fluxes to be calculated and trade-offs between species- and community-level objectives to be analysed [37,38,100]. This approach makes no assumption on the relation between individual growth rates that are calculated for each individual species. It is most suitable for well-characterized communities with a defined community objective and, for example, does not include the analysis of communities whose members only compete for resources. However, this approach requires the formulation of a non-convex, bilinear optimization problem that cannot be solved by standard linear programming solvers [37].

Under the constraint that two organisms have to be able to produce biomass at a rate greater than a certain threshold—which allows growth rates to differ—medium compositions that can induce interactions between two species were determined [55]. With this method, known interactions between organisms could be confirmed and new ones predicted, and it allows synthetic communities to be created without modifying the organisms themselves.

These examples show that a lot of valuable information can be gained without assuming equal growth rates, but there are two issues regarding this kind of approach. First, as discussed earlier, these results will describe only a snapshot of the community dynamics, because unequal growth rates will result in rapid changes of shared external metabolites and the biomass abundances of the species. This means that stable relative biomass abundances cannot be inferred when the species' growth rates differ. If this is desired to compare with experimentally determined data, then the equal growth rate constraint needs to be taken into account.

There is a second point to consider when an approach such as equation (3.5) is used, which occurs for communities whose members compete for shared metabolites. If the community species differ in their biomass composition, then a weighted sum of biomass-forming rates will lead to a solution where all available resources are invested in the growth of the organism with the highest biomass yield, i.e. there will be only one biomass formation reaction that carries an actual flux, and which one this is depends on the chosen weights w_i. In other words, all resources will be invested in the least expensive biomass. Technically, this issue can be avoided by assuming equal growth rates [39], by setting lower limits on the growth rates, as in, for example, [55,99], or by coupling the respective transport fluxes to growth rate, as in [101].

One might argue that this problem could be solved by interspecies feeding mechanisms. The reasoning behind this is that then resources cannot only be invested in one organism, but also in the other one because they depend on each other. While this will indeed lead to several metabolically active networks and an exchange of metabolites, still only the biomass-forming reaction of the organism with the highest biomass yield will carry a non-zero flux (if both organisms compete for the same resource and have different biomass yields for this substrate). The network of the other organism will only be used to provide the required metabolite, but resources will not be invested in growth. This is rather artificial as these fluxes require a catalyst—biomass. Coupling of excretion flux to growth rate is one way this problem has been tackled [101].

Besides using explicit constraints on the growth rates, one can also overcome this issue by merging the species' biomass equations into one, so that all the required macromolecules are produced. Instead of maximizing $w_1 \cdot \mu_1 + w_2 \cdot \mu_2$ (in the case of a two-species community) one defines according to equation (2.9)

$$
\begin{aligned}
& b_{11} \cdot M_{b11} + b_{12} \cdot M_{b12} + \cdots + b_{1z} \cdot M_{b1z} + b_{21} \cdot M_{b21} \\
& + b_{22} \cdot M_{b22} + \cdots + b_{y2} \cdot M_{b2y},
\end{aligned} \tag{3.6}
$$

as an objective function where M_{b1i} and M_{b2k} correspond to the biomass components in organisms 1 and 2, respectively. Such a lumped biomass reaction ensures that biomass components of all species in the community are produced. Merging of models into one network, however, should still take the biomass abundances into account for the exchange fluxes; if not it will implicitly assume a species ratio of $1:1$.

4. Dynamic metabolism-based models of microbial communities: dynamic community flux balance analysis

Steady-state cFBA based on balanced growth is well suited to make predictions about communities in fairly constant environments, such as chemostats, regarding their metabolic
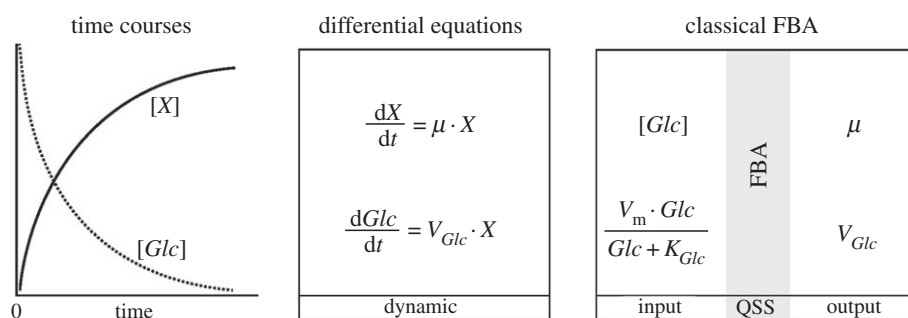
Figure 4. dFBA allows the prediction of time courses for metabolite and biomass concentrations using the quasi-steady-state (QSS) approximation. In this figure, we illustrate the typical results of a dFBA simulation. While a carbon source, *Glc*, is consumed over time, the associated biomass, *X*, increases. The concentration profiles are described by a set of differential equations incorporating the growth rate μ and the net uptake rate for *Glc*, V_{Glc}. These rates are determined using classical FBA whose constraints are dynamically calculated, using Michaelis–Menten kinetics.

interactions and optimal relative biomass abundances. However, in batch cultures and many natural communities, organisms are exposed to dynamically changing conditions and can grow sequentially, something this framework cannot accommodate. For such an analysis, dFBA can be used—under simplifying conditions. dFBA was first used to describe the diauxic growth of *Escherichia coli* [85] and has since also been extended to describe community dynamics [14,15,102–107].

dFBA models are similar to FBA models, except that the constraints on the input fluxes are made dependent on extracellular concentrations through kinetic rate expressions. Usually Michaelis–Menten kinetics are used, which are sometimes extended with inhibitory terms [15],

$$v_j = v_{\max} \cdot \frac{S_{\text{ext},j}}{K_j + S_{\text{ext},j}}. \tag{4.1}$$

When the K_j parameter is equated to the affinity of the transporter, equation (4.1) describes purely substrate-limited uptake, and thus assumes that the transport step is fully controlling the downstream metabolic pathway. From metabolic control analysis, we know that this is not necessary and in fact not very likely, as control over steady-state flux tends to be distributed [108]. On the other hand, K_j can be interpreted as the Monod constant for growth (which tends to be lower than the K_M of the transporter [109]). In practice, K_j is often fitted to data, and the distinction does not have an effect on the simulations, but it is good to be aware of the underlying assumptions. With the uptake kinetics, a system of differential equations can be defined that describes the dynamics of external metabolite concentrations, the consequent dynamics of uptake rates and, via FBA, biomass formation. Thus, at each time integration step, a linear program computes the growth rate from the uptake kinetics, and both biomass and external metabolites are updated according to

$$\frac{dX}{dt} = \mu(v_i(S_{\text{ext},i})) \cdot X, \tag{4.2}$$

for the biomass, and for the external concentration

$$\frac{dS_{\text{ext},i}}{dt} = v_i(S_{\text{ext},i}) \cdot X, \tag{4.3}$$

where v_i is the net uptake or production of an external metabolite and the convention in the field is to assign a negative value for uptake fluxes.

One important assumption in dFBA is that the time constants related to intracellular dynamics are much smaller than those that describe changes of external concentrations. This is the basis for the quasi-steady-state (QSS) approximation inherent in the approach. The assumption that metabolism is in QSS during the relatively slow medium transitions, invoked by the cell's own metabolic activities, is very reasonable and generally accepted in the field. However, for adaptations that require gene-regulatory processes that may act at comparable time scales to the environmental changes, this does not apply, and so cells need not to be in balanced growth, in the sense that *all* biomass components, so also proteins, are in steady state. As long as proteome reallocations are not explicitly modelled in genome-scale metabolic models—but they will be in the future—this complication is not an issue. The concept of dFBA is summarized in figure 4.

Extending dFBA to communities is then relatively straightforward: one defines input fluxes for which kinetic expressions are required, and solves the set of differential equations, using FBA for each organism individually. However, here we also face a challenge in the interactions between the species, and different solutions that researchers have suggested.

4.1. Dynamic community flux balance analysis: applications and limitations

dFBA has been proven to be a very useful tool to analyse metabolism on a community level [14,15,102–107,110]. It allows communities where species grow in succession to be modelled, which cannot be achieved by steady-state cFBA. Multi-species dFBA was first used to describe the competition between *Rhodoferax ferrireducens* and *Geobacter sulfurreducens* in an anoxic subsurface environment leading to predictions about the conditions under which one can outcompete the other [102]. Later on, this approach was applied to a synthetic community consisting of *E. coli* and *Saccharomyces cerevisiae* that—in this study—exclusively consume xylose and glucose, respectively, to engineer a co-culture system in which both sugars are consumed simultaneously [104]. Chiu *et al.* [111] combined FVA and dFBA to examine 6670 two-species communities consisting of 116 species regarding metabolites that can be produced in a community, but not by a single species growing on the same medium. They showed that emergent biosynthetic capacity occurs in most of the communities, and in two phases: once the two organisms are introduced into the same growth medium and when the medium is nutrient-depleted at the end of a growth phase.

In these studies, the important assumption is made that communities are spatially homogeneous. Often, however,

natural occurring communities tend to be structured, which can give rise to interesting dynamics [112–114]. This layer of complexity was addressed in [110], where multi-species community dFBA is coupled with diffusion, allowing the analyses of spatio-temporal effects. This framework was used to predict the species ratio to which a community of two and three members converges, which was found to be in agreement with experimental findings. Similar approaches have also been developed for monocultures [115–117].

As for any application of FBA, the choice of the objective function has certain implications. It has been debated for a long time on which level natural selection occurs: on the species or community level [90–97]. Without specific knowledge about the examined community's objective, it is reasonable to assume that each member of a community attempts to maximize its own growth rate. This is indeed the assumption made in the dynamic multi-species metabolic modelling (DMMM) framework [102], which was the first that allowed dFBA to be performed on a community level.

One very interesting—and challenging—consequence of optimizing organisms individually is that one needs to hard-wire the interactions between the organisms, as they cannot be predicted. The reason for this is that at each time instance FBA will predict the currently most optimal flux distribution for growth. It will not predict the suboptimal excretion of a metabolite (e.g. D), even though, in the long run, providing D to another organism would provide a growth benefit later on. The models cannot predict the optimal strategy over a longer period of time, only at each time instance. If another organism is auxotrophic for this compound D and it is not provided in the medium, the growth dynamics of this organism cannot be described without explicitly setting the model up for this purpose. In contrast, in steady-state cFBA, this interaction would show up without explicitly enforcing it.

The d-OptCom framework [107] tries to capture this kind of behaviour by using not only fitness functions of individual species, but also a community-level objective. Using this, the authors claimed to be, indeed, able to describe interactions that could not be modelled using DMMM [107]. Obviously, this approach assumes that a community-level objective function exists, which does not necessarily have to be the case. Moreover, owing to the nonlinear nature of this optimization problem, it is unclear whether this approach will scale—from a computational point of view—when larger communities are examined.

Scaling up to more complex communities is, in fact, a general problem also for dFBA. The extension to larger naturally occurring communities is non-trivial: sequencing and reconstructing metabolic models—including the determination of the respective biomass compositions—for a representative amount of species is, as for all FBA approaches, very time-consuming. It is however important to have good reconstructions as they determine the inference of the interactions between microorganisms. Moreover, as all the interactions ultimately require some kinetic description of the input fluxes, with increasing size of the community, also the number of parameters required to model exchange rates will increase considerably, making it computationally harder to integrate the system and fit the parameters given experimental data. Not surprisingly therefore, so far, dynamic cFBA has mainly been applied to small well-characterized communities whose members have been sequenced (i.e. genome-scale models exist) and where metabolic interactions are known. In these cases, modelling the community with dFBA provides flux distributions within and between species, and can provide insight into the strength of metabolic interactions that cannot be measured directly.

5. Summary and outlook

In this paper, we illustrated and discussed how steady-state cFBA can be derived from classical FBA without any additional assumptions. We made a clear distinction between different types of fluxes—specific internal fluxes, exchange fluxes and fluxes towards biomass—as they have different units that are relevant in a community setting. We concluded that stable biomass abundances in communities can only be modelled if (i) the species' relative biomass abundances are explicitly considered and (ii) all organisms grow with the same growth; this 'community growth rate' can then serve as the objective function.

cFBA allows optimal biomass abundances and the associated community growth rate at a steady state of balanced growth to be calculated. Furthermore, this approach can be used to infer essential interactions between species in a certain environment without making any assumption about how and whether species interact. However, the cases in which community members all grow at the same rate may be rather limited because that can only be achieved in fairly stable environments. In natural occurring communities, however, dynamically changing environmental conditions are prevalent which cannot be incorporated into cFBA. Here, dcFBA can be used. This approach is, however, limited in its predictive power of unknown interactions and usually requires that the entire system structure is defined beforehand. This is, in particular, true for mutual interactions based on suboptimal growth strategies, i.e. if the excretion of a by-product comes at a growth cost but would benefit an organism in the long run, as it would allow growth of another organism that conversely provides growth-promoting metabolites. While dcFBA is a very powerful tool to describe small synthetic well-characterized communities, its extension to larger communities is not straightforward, owing to the number of parameters required to model transport rate dynamics and the associated numerical instability. One way to keep the number of parameters in a reasonable range is to use explicit exchange kinetics only for metabolites that are essential for the community to grow optimally. This information could be revealed by performing FVA on the solution computed by steady-state cFBA. This combination of cFBA and dcFBA therefore seems promising.

We still face many challenges in modelling metabolism in communities that go significantly beyond insufficient experimental data or reconstructed genomes. We require advances in how to define the objective function of communities, how to deal with dynamics and subsequent (sub)optimal dynamic strategies within the constraint-based modelling format, and how to incorporate the recent advances made in modelling monocultures [82–84]. It seems clear however that better use of metagenomics data towards understanding ecosystem functioning will require models that incorporate such genomics data and ways to make these models useful.

References

1. Brenner K, You L, Arnold FH. 2008 Engineering microbial consortia: a new frontier in synthetic biology. *Trends Biotechnol*. **26**, 483–489. (doi:10.1016/j.tibtech.2008.05.004)

2. Alper H, Stephanopoulos G. 2009 Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential? *Nat. Rev. Microbiol*. **7**, 715–723. (doi:10.1038/nrmicro2186)

3. Ley RE, Turnbaugh PJ, Klein S, Gordon JI. 2006 Microbial ecology: human gut microbes associated with obesity. *Nature* **444**, 1022–1023. (doi:10.1038/4441022a)

4. Fukuda S, Ohno H. 2014 Gut microbiome and metabolic diseases. *Semin. Immunopathol*. **36**, 103–114. (doi:10.1007/s00281-013-0399-z)

5. Heintz C, Mair W. 2014 You are what you host: microbiome modulation of the aging process. *Cell* **156**, 408–411. (doi:10.1016/j.cell.2014.01.025)

6. Moloney RD, Desbonnet L, Clarke G, Dinan TG, Cryan JF. 2014 The microbiome: stress, health and disease. *Mamm. Genome* **25**, 49–74. (doi:10.1007/s00335-013-9488-5)

7. Wolfe BE, Dutton RJ. 2015 Fermented foods as experimentally tractable microbial ecosystems. *Cell* **161**, 49–55. (doi:10.1016/j.cell.2015.02.034)

8. Follows MJ, Dutkiewicz S, Grant S, Chisholm SW. 2007 Emergent biogeography of microbial communities in a model ocean. *Science* **315**, 1843–1846. (doi:10.1126/science.1138544)

9. Handelsman J. 2005 Sorting out metagenomes. *Nat. Biotechnol*. **23**, 38–39. (doi:10.1038/nbt0105-38)

10. Kunin V, Copeland A, Lapidus A, Mavromatis K, Hugenholtz P. 2008 A bioinformaticians guide to metagenomics. *Microbiol. Mol. Biol. Rev*. **72**, 557–578. (doi:10.1128/MMBR.00009-08)

11. Sangwan N, Xia F, Gilbert JA. 2016 Recovering complete and draft population genomes from metagenome datasets. *Microbiome* **4**, 8. (doi:10.1186/s40168-016-0154-5)

12. Anantharaman K *et al*. 2016 Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat. Commun*. **7**, 13219. (doi:10.1038/ncomms13219)

13. Freilich S, Zarecki R, Eilam O, Segal ES, Henry CS, Kupiec M, Gophna U, Sharan R, Ruppin E. 2011 Competitive and cooperative metabolic interactions in bacterial communities. *Nat. Commun*. **2**, 589. (doi:10.1038/ncomms1597)

14. Song HS, Cannon W, Beliaev A, Konopka A. 2014 Mathematical modeling of microbial community dynamics: a methodological review. *Processes* **2**, 711–752. (doi:10.3390/pr2040711)

15. Henson MA, Hanly TJ. 2014 Dynamic flux balance analysis for synthetic microbial communities. *IET Syst. Biol*. **8**, 214–229. (doi:10.1049/iet-syb.2013.0021)

16. Ponomarova O, Patil KR. 2015 Metabolic interactions in microbial communities: untangling the Gordian knot. *Curr. Opin. Microbiol*. **27**, 37–44. (doi:10.1016/j.mib.2015.06.014)

17. Biggs MB, Medlock GL, Kolling GL, Papin JA. 2015 Metabolic network modeling of microbial communities. *Wiley Interdiscip. Rev. Syst. Biol. Med*. **7**, 317–334. (doi:10.1002/wsbm.1308)

18. Forster J. 2003 Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res*. **13**, 244–253. (doi:10.1101/gr.234503)

19. Famili I, Forster J, Nielsen J, Palsson BØ. 2003 Saccharomyces cerevisiae phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network. *Proc. Natl Acad Sci. USA* **100**, 13 134–13 139. (doi:10.1073/pnas.2235812100)

20. Price ND, Reed JL, Palsson BØ. 2004 Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat. Rev. Microbiol*. **2**, 886–897. (doi:10.1038/nrmicro1023)

21. Feist AM, Henry CS, Reed JL, Krummenacker M, Joyce AR, Karp PD, Broadbelt LJ, Hatzimanikatis V, Palsson BØ. 2007 A genome scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol. Syst. Biol*. **3**, 121. (doi:10.1038/msb4100155)

22. Herrgård MJ *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)

23. Österlund T, Nookaew I, Nielsen J. 2012 Fifteen years of large scale metabolic modeling of yeast: developments and impacts. *Biotechnol. Adv*. **30**, 979–988. (doi:10.1016/j.biotechadv.2011.07.021)

24. O'Brien EJ, Monk JM, Palsson BØ. 2015 Using genome-scale models to predict biological capabilities. *Cell* **161**, 971–987. (doi:10.1016/j.cell.2015.05.019)

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

26. Orth JD, Thiele I, Palsson BØ. 2010 What is flux balance analysis? *Nat. Biotechnol*. **28**, 245–248. (doi:10.1038/nbt.1614)

27. Schuetz R, Kuepfer L, Sauer U. 2007 Systematic evaluation of objective functions for predicting intra cellular fluxes in *Escherichia coli*. *Mol. Syst. Biol*. **3**, 119. (doi:10.1038/msb4100162)

28. Feist AM, Palsson BØ. 2010 The biomass objective function. *Curr. Opin. Microbiol*. **13**, 344–349. (doi:10.1016/j.mib.2010.03.003)

29. Schuetz R, Zamboni N, Zampieri M, Heinemann M, Sauer U. 2012 Multidimensional optimality of microbial metabolism. *Science* **336**, 601–604. (doi:10.1126/science.1216882)

30. Overbeek R. 2005 The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res*. **33**, 5691–5702. (doi:10.1093/nar/gki866)

31. Thiele I, Palsson BØ. 2010 A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc*. **5**, 93–121. (doi:10.1038/nprot.2009.203)

32. Henry CS, DeJongh M, Best AA, Frybarger PM, Linsay B, Stevens RL. 2010 High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol*. **28**, 977–982. (doi:10.1038/nbt.1672)

33. King ZA, Lu J, Drger A, Miller P, Federowicz S, Lerman JA, Ebrahim A, Palsson BO, Lewis NE. 2015 BiGG models: a platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res*. **44**, D515–D522. (doi:10.1093/nar/gkv1049)

34. van Heck RGA, Ganter M, Martins dos Santos VAP, Stelling J. 2016 Efficient reconstruction of predictive consensus metabolic network models. *PLoS Comput. Biol*. **12**, e1005085. (doi:10.1371/journal.pcbi.1005085)

35. Olivier BG, Bergmann FT. 2015 The systems biology markup language (SBML) level 3 package: flux balance constraints. *J. Integr. Bioinform*. **12**, 269. (doi:10.2390/biecoll-jib-2015-269)

36. Stolyar S, Van Dien S, Hillesland KL, Pinel N, Lie TJ, Leigh JA, Stahl DA. 2007 Metabolic modeling of a mutualistic microbial community. *Mol. Syst. Biol*. **3**, 92. (doi:10.1038/msb4100131)

37. Zomorrodi AR, Maranas CD. 2012 OptCom: a multi-level optimization framework for the metabolic modeling and analysis of microbial communities. *PLoS Comput. Biol*. **8**, e1002363. (doi:10.1371/journal.pcbi.1002363)

38. Shoaie S, Karlsson F, Mardinoglu A, Nookaew I, Bordel S, Nielsen J. 2013 Understanding the interactions between bacteria in the human gut through metabolic modeling. *Sci. Rep*. **3**, 2532. (doi:10.1038/srep02532)

39. Khandelwal RA, Olivier BG, Röling WFM, Teusink B, Bruggeman FJ. 2013 Community flux balance analysis for microbial consortia at balanced growth. *PLoS ONE* **8**, e64567. (doi:10.1371/journal.pone.0064567)

40. Wiechert W, Nöh K. 2005 From stationary to instationary metabolic flux analysis. *Adv. Biochem. Eng. Biotechnol*. **92**, 145–172. (doi:10.1007/b98921)

41. Sauer U. 2006 Metabolic networks in motion: 13C-based flux analysis. *Mol. Syst. Biol*. **2**, 62. (doi:10.1038/msb4100109)

42. Neufeld JD, Wagner M, Murrell JC. 2007 Who eats what, where and when? Isotope-labelling experiments are coming of age. *ISME J*. **1**, 103–110. (doi:10.1038/ismej.2007.30)

43. Kovatcheva-Datchary P, Egert M, Maathuis A, Rajili-Stojanovi M, de Graaf AA, Smidt H, de Vos WM, Venema K. 2009 Linking phylogenetic identities of bacteria to starch fermentation in an *in vitro* model of the large intestine by RNA-based stable isotope probing. *Environ. Microbiol*. **11**, 914–926. (doi:10.1111/j.1462-2920.2008.01815.x)

44. Paterson E. 2013 The use of stable isotope labeling and compound-specific analysis of microbial phospholipid fatty acids to quantify the influences

of rhizodeposition on microbial community structure and function. In *Molecular microbial ecology of the rhizosphere* (ed. FJ de Bruijn), pp. 141–148. Hoboken, NJ: John Wiley & Sons.

45. von Bergen M, Jehmlich N, Taubert M, Vogt C, Bastida F, Herbst FA, Schmidt F, Richnow H-H, Seifert J. 2013 Insights from quantitative metaproteomics and protein-stable isotope probing into microbial ecology. *ISME J.* **7**, 1877–1885. (doi:10.1038/ismej.2013.78)

46. Hawley AK, Brewer HM, Norbeck AD, Pa a-Toli L, Hallam SJ. 2014 Metaproteomics reveals differential modes of metabolic coupling among ubiquitous oxygen minimum zone microbes. *Proc. Natl Acad. Sci. USA* **111**, 11 395–11 400. (doi:10.1073/pnas. 1322132111)

47. Verastegui Y *et al.* 2014 Multisubstrate isotope labeling and metagenomic analysis of active soil bacterial communities. *mBio* **5**, e0115714. (doi:10. 1128/mBio.01157-14)

48. Ghosh A, Nilmeier J, Weaver D, Adams PD, Keasling JD, Mukhopadhyay A, Petzold CJ, Martin HG. 2014 A peptide-based method for ^{13}C metabolic flux analysis in microbial communities. *PLoS Comput. Biol.* **10**, e1003827. (doi:10.1371/journal.pcbi. 1003827)

49. Hibbing ME, Fuqua C, Parsek MR, Peterson SB. 2010 Bacterial competition: surviving and thriving in the microbial jungle. *Nat. Rev. Microbiol.* **8**, 15–25. (doi:10.1038/nrmicro2259)

50. Elias S, Banin E. 2012 Multi-species biofilms: living with friendly neighbors. *FEMS Microbiol. Rev.* **36**, 990–1004. (doi:10.1111/j.1574-6976.2012.00325.x)

51. Foster KR, Bell T. 2012 Competition, not cooperation, dominates interactions among culturable microbial species. *Curr. Biol.* **22**, 1845–1850. (doi:10.1016/j.cub.2012.08.005)

52. Oliveira NM, Niehus R, Foster KR. 2014 Evolutionary limits to cooperation in microbial communities. *Proc. Natl Acad. Sci. USA* **111**, 17 941–17 946. (doi:10.1073/pnas.1412673111)

53. Campbell K *et al.* 2015 Self establishing communities enable cooperative metabolite exchange in a eukaryote. *eLife* **4**, e09943. (doi:10. 7554/eLife.09943)

54. Deng YJ, Wang SY. 2016 Synergistic growth in bacteria depends on substrate complexity. *J. Microbiol.* **54**, 23–30. (doi:10.1007/s12275-016-5461-9)

55. Klitgord N, Segrè D. 2010 Environments that induce synthetic microbial ecosystems. *PLoS Comput. Biol.* **6**, e1001002. (doi:10.1371/journal. pcbi.1001002)

56. Papoutsakis ET. 1984 Equations and calculations for fermentations of butyric acid bacteria. *Biotechnol. Bioeng.* **26**, 174–187. (doi:10.1002/bit.260260210)

57. Watson MR. 1984 Metabolic maps for the Apple II. *Biochem. Soc. Trans.* **12**, 1093–1094. (doi:10.1042/ bst0121093)

58. Fell DA, Small JR. 1986 Fat synthesis in adipose tissue. An examination of stoichiometric constraints. *Biochem. J.* **238**, 781–786. (doi:10.1042/ bj2380781)

59. Burgard AP, Maranas CD. 2001 Probing the performance limits of the *Escherichia coli* metabolic network subject to gene additions or deletions. *Biotechnol. Bioeng.* **74**, 364–375. (doi:10.1002/bit.1127)

60. Mahadevan R, Schilling CH. 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)

61. Reed JL, Palsson BØ. 2004 Genome-scale in silico models of E. coli have multiple equivalent phenol typic states: assessment of correlated reaction subsets that comprise network states. *Genome Res.* **14**, 1797–1805. (doi:10.1101/gr.2546004)

62. Llaneras F, Picó J. 2007 An interval approach for dealing with flux distributions and elementary modes activity patterns. *J. Theor. Biol.* **246**, 290–308. (doi:10.1016/j.jtbi.2006.12.029)

63. Kelk SM, Olivier BG, Stougie L, Bruggeman FJ. 2012 Optimal flux spaces of genome-scale stoichiometric models are determined by a few subnetworks. *Sci. Rep.* **2**, 580. (doi:10.1038/srep00580)

64. Maarleveld TR, Wortel MT, Olivier BG, Teusink B, Bruggeman FJ. 2015 Interplay between constraints, objectives, and optimality for genome-scale stoichiometric models. *PLoS Comput. Biol.* **11**, e1004166. (doi:10.1371/journal.pcbi.1004166)

65. Burgard AP, Pharkya P, Maranas CD. 2003 Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol. Bioeng.* **84**, 647–657. (doi:10.1002/bit.10803)

66. Kim J, Reed JL. 2010 OptORF: optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains. *BMC Syst. Biol.* **4**, 53. (doi:10.1186/1752-0509-4-53)

67. Tepper N, Shlomi T. 2010 Predicting metabolic engineering knockout strategies for chemical production: accounting for competing pathways. *Bioinformatics* **26**, 536–543. (doi:10.1093/ bioinformatics/btp704)

68. Ibarra RU, Edwards JS, Palsson BØ. 2002 *Escherichia coli* K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* **420**, 186–189. (doi:10.1038/nature01149)

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

70. Segrè D, Vitkup D, Church GM. 2002 Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl Acad. Sci. USA* **99**, 15 112–15 117. (doi:10.1073/pnas.232349399)

71. Joyce AR, Palsson BØ. 2008 Predicting gene essentiality using genome-scale in silico models. *Methods Mol. Biol.* **416**, 433–457. (doi:10.1007/ 978-1-59745-321-9_30)

72. Suthers PF, Zomorrodi A, Maranas CD. 2009 Genome-scale gene/reaction essentiality and synthetic lethality analysis. *Mol. Syst. Biol.* **5**, 301. (doi:10.1038/msb.2009.56)

73. Edwards JS, Ibarra RU, Palsson BØ. 2001 In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nat. Biotechnol.* **19**, 125–130. (doi:10.1038/ 84379)

74. Pfeiffer T, Schuster S, Bonhoeffer S. 2001 Cooperation and competition in the evolution of ATP producing pathways. *Science* **292**, 504–507. (doi:10.1126/science.1058079)

75. Teusink B, Wiersma A, Molenaar D, Francke C, de Vos WM, Siezen RJ, Smid EJ. 2006 Analysis of growth of *Lactobacillus plantarum* WCFS1 on a complex medium using a genome-scale metabolic model. *J. Biol. Chem.* **281**, 40 041–40 048. (doi:10. 1074/jbc.M606263200)

76. Schuster S, Pfeiffer T, Fell DA. 2008 Is maximization of molar yield in metabolic networks favoured by evolution? *J. Theor. Biol.* **252**, 497–504. (doi:10. 1016/j.jtbi.2007.12.008)

77. Beg QK, Vazquez A, Ernst J, de Menezes MA, Bar-Joseph Z, Barabási AL, Oltvai ZN. 2007 Intracellular crowding defines the mode and sequence of substrate uptake by *Escherichia coli* and constrains its metabolic activity. *Proc. Natl Acad. Sci. USA* **104**, 12 663–12 668. (doi:10.1073/pnas. 0609845104)

78. Vazquez A, Beg QK, Demenezes MA, Ernst J, Bar-Joseph Z, Barabási AL, Boros LG, Oltvai ZN. 2008 Impact of the solvent capacity constraint on E. coli metabolism. *BMC Syst. Biol.* **2**, 7. (doi:10.1186/ 1752-0509-2-7)

79. Shlomi T, Benyamini T, Gottlieb E, Sharan R, Ruppin E. 2011 Genome-scale metabolic modeling elucidates the role of proliferative adaptation in causing the Warburg effect. *PLoS Comput. Biol.* **7**, e1002018. (doi:10.1371/journal. pcbi.1002018)

80. Zhuang K, Vemuri GN, Mahadevan R. 2011 Economics of membrane occupancy and respiro fermentation. *Mol. Syst. Biol.* **7**, 500. (doi:10.1038/ msb.2011.34)

81. Lerman JA *et al.* 2012 In silico method for modelling metabolism and gene product expression at genome scale. *Nat. Commun.* **3**, 929. (doi:10. 1038/ncomms1928)

82. O'Brien EJ, Lerman JA, Chang RL, Hyduke DR, Palsson BØ. 2013 Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Mol. Syst. Biol.* **9**, 693. (doi:10.1038/msb.2013.52)

83. Nilsson A, Nielsen J. 2016 Metabolic trade-offs in yeast are caused by F1F0-ATP synthase. *Sci. Rep.* **6**, 22264. (doi:10.1038/srep22264)

84. Mori M, Hwa T, Martin OC, De Martino A, Marinari E. 2016 Constrained allocation flux balance analysis. *PLoS Comput. Biol.* **12**, e1004913. (doi:10.1371/ journal.pcbi.1004913)

85. Mahadevan R, Edwards JS, Doyle III FJ. 2002 Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. *Biophys. J.* **83**, 1331–1340. (doi:10. 1016/S0006-3495(02)73903-9)

86. Molenaar D, van Berlo R, de Ridder D, Teusink B. 2009 Shifts in growth strategies reflect tradeoffs in cellular economics. *Mol. Syst. Biol.* **5**, 323. (doi:10. 1038/msb.2009.82)

87. Gottstein W, Müller S, Herzel H, Steuer R. 2014 Elucidating the adaptation and temporal coordination of metabolic pathways using *in-silico* evolution. *Biosystems* **117**, 68–76. (doi:10.1016/j.biosystems.2013.12.006)

88. Grosskopf T, Soyer OS. 2014 Synthetic microbial communities. *Curr. Opin Microbiol*. **18**, 72–77. (doi:10.1016/j.mib.2014.02.002)

89. Mahadevan R, Henson MA. 2012 Genome-based modeling and design of metabolic interactions in microbial communities. *Comput. Struct. Biotechnol. J*. **3**, e201210008. (doi:10.5936/csbj.201210008)

90. Smith JM. 1964 Group selection and kin selection. *Nature* **201**, 1145–1147. (doi:10.1038/2011145a0)

91. Wilson DS. 1975 A theory of group selection. *Proc. Natl Acad. Sci. USA* **72**, 143–146. (doi:10.1073/pnas.72.1.143)

92. Rainey PB, Travisano M. 1998 Adaptive radiation in a heterogeneous environment. *Nature* **394**, 69–72. (doi:10.1038/27900)

93. Griffin AS, West SA, Buckling A. 2004 Cooperation and competition in pathogenic bacteria. *Nature* **430**, 1024–1027. (doi:10.1038/nature02744)

94. Wilson DS, Wilson EO. 2007 Rethinking the theoretical foundation of sociobiology. *Q. Rev. Biol*. **82**, 327–348. (doi:10.1086/522809)

95. Strassmann JE, Queller DC, Avise JC, Ayala FJ. 2011 In the light of evolution V: cooperation and conflict. *Proc. Natl Acad. Sci. USA* **108**, 10 787–10 791. (doi:10.1073/pnas.1100289108)

96. Liao X, Rong S, Queller DC. 2015 Relatedness, conflict, and the evolution of eusociality. *PLoS Biol*. **13**, e1002098. (doi:10.1371/journal.pbio.1002098)

97. Hammerstein P, Noë R. 2016 Biological trade and markets. *Phil. Trans. R. Soc. B* **371**, 20150101. (doi:10.1098/rstb.2015.0101)

98. Merino MP, Andrews BA, Asenjo JA. 2014 Stoichiometric model and flux balance analysis for a mixed culture of *Leptospirillum ferriphilum* and *Ferroplasma acidiphilum*. *Biotechnol. Prog*. **31**, 307–315. (doi:10.1002/btpr.2028)

99. Heinken A, Thiele I. 2015 Anoxic conditions promote species-specific mutualism between gut microbes *in silico*. *Appl. Environ. Microbiol*. **81**, 4049–4061. (doi:10.1128/AEM.00101-15)

100. El-Semman IE, Karlsson FH, Shoaie S, Nookaew I, Soliman TH, Nielsen J. 2014 Genome-scale metabolic reconstructions of *Bifidobacterium adolescentis* L2-32 and *Faecalibacterium prausnitzii* A2-165 and their interaction. *BMC Syst. Biol*. **8**, 41. (doi:10.1186/1752-0509-8-41)

101. Heinken A, Sahoo S, Fleming RMT, Thiele I. 2013 Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. *Gut Microbes* **4**, 28–40. (doi:10.4161/gmic.22370)

102. Zhuang K, Izallalen M, Mouser P, Richter H, Risso C, Mahadevan R, Lovley DR. 2011 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)

103. Zhuang K, Ma E, Lovley DR, Mahadevan R. 2012 The design of long-term effective uranium biore mediation strategy using a community metabolic model. *Biotechnol. Bioeng*. **109**, 2475–2483. (doi:10.1002/bit.24528)

104. Hanly TJ, Urello M, Henson MA. 2012 Dynamic flux balance modeling of *S. cerevisiae* and *E. coli* co-cultures for efficient consumption of glucose/xylose mixtures. *Appl. Microbiol. Biotechnol*. **93**, 2529–2541. (doi:10.1007/s00253-011-3628-1)

105. Zhuang K, Yang L, Cluett WR, Mahadevan R. 2013 Dynamic strain scanning optimization: an efficient strain design strategy for balanced yield, titer, and productivity. DySScO strategy for strain design. *BMC Biotechnol*. **13**, 8. (doi:10.1186/1472-6750-13-8)

106. Hanly TJ, Henson MA. 2013 Dynamic metabolic modeling of a microaerobic yeast co-culture: predicting and optimizing ethanol production from glucose/xylose mixtures. *Biotechnol. Biofuels* **6**, 44. (doi:10.1186/1754-6834-6-44)

107. Zomorrodi AR, Islam MM, Maranas CD. 2014 d-OptCom: dynamic multi-level and multi-objective metabolic modeling of microbial communities. *ACS Synth. Biol*. **3**, 247–257. (doi:10.1021/sb4001307)

108. Fell DA. 2005 Metabolic control analysis. *Syst. Biol*. **13**, 69–80. (doi:10.1007/b137745)

109. Bachmann H, Bruggeman FJ, Molenaar D, Branco dos Santos F, Teusink B. 2016 Public goods and metabolic strategies. *Curr. Opin Microbiol*. **31**, 109–115. (doi:10.1016/j.mib.2016.03.007)

110. Harcombe WR *et al.* 2014 Metabolic resource allocation in individual microbes determines ecosystem interactions and spatial dynamics. *Cell Rep*. **7**, 1104–1115. (doi:10.1016/j.celrep.2014.03.070)

111. Chiu HC, Levy R, Borenstein E. 2014 Emergent biosynthetic capacity in simple microbial communities. *PLoS Comput. Biol*. **10**, e1003695. (doi:10.1371/journal.pcbi.1003695)

112. Kerr B, Riley MA, Feldman MW, Bohannan BJM. 2002 Local dispersal promotes biodiversity in a real life game of rockpaperscissors. *Nature* **418**, 171–174. (doi:10.1038/nature00823)

113. Fuhrman JA. 2009 Microbial community structure and its functional implications. *Nature* **459**, 193–199. (doi:10.1038/nature08058)

114. Liu J, Prindle A, Humphries J, Gabalda-Sagarra M, Asally M, Lee D-yD, Ly S, Garcia-Ojalvo J, Süel GM. 2015 Metabolic co dependence gives rise to collective oscillations within biofilms. *Nature* **523**, 550–554. (doi:10.1038/nature14660)

115. Biggs MB, Papin JA. 2013 Novel multiscale modeling tool applied to *Pseudomonas aeruginosa* biofilm formation. *PLoS ONE* **8**, e78011. (doi:10.1371/journal.pone.0078011)

116. Cole JA, Kohler L, Hedhli J, Luthey-Schulten Z. 2015 Spatially-resolved metabolic cooperativity within dense bacterial colonies. *BMC Syst. Biol*. **9**, 15. (doi:10.1186/s12918-015-0155-1)

117. Chen J, Gomez JA, Höffner K, Phalak P, Barton PI, Henson MA. 2016 Spatiotemporal modeling of microbial metabolism. *BMC Syst. Biol*. **10**, 21. (doi:10.1186/s12918-016-0259-2)