

RESEARCH ARTICLE

A multi-objective constraint-based approach for modeling genome-scale microbial ecosystems

Marko Budinich*, Jérémie Bourdon, Abdelhalim Larhlimi, Damien Eveillard

Computational Biology group, LINA UMR 6241 CNRS, EMN, Université de Nantes, Nantes, France

* marko.budinich@univ-nantes.fr



OPEN ACCESS

Citation: Budinich M, Bourdon J, Larhlimi A, Eveillard D (2017) A multi-objective constraint-based approach for modeling genome-scale microbial ecosystems. PLoS ONE 12(2): e0171744. doi:10.1371/journal.pone.0171744

Editor: Tamir Tuller, Tel Aviv University, ISRAEL

Received: July 28, 2016

Accepted: January 25, 2017

Published: February 10, 2017

Copyright: © 2017 Budinich et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: MB is supported by CNRS and Region Pays de la Loire funding (GRIOTE project, <http://griote.univ-nantes.fr/>). This study is supported by ANR (IMPEKAB, ANR-15-CE02-001-03). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Interplay within microbial communities impacts ecosystems on several scales, and elucidation of the consequent effects is a difficult task in ecology. In particular, the integration of genome-scale data within quantitative models of microbial ecosystems remains elusive. This study advocates the use of constraint-based modeling to build predictive models from recent high-resolution -omics datasets. Following recent studies that have demonstrated the accuracy of constraint-based models (CBMs) for simulating single-strain metabolic networks, we sought to study microbial ecosystems as a combination of single-strain metabolic networks that exchange nutrients. This study presents two multi-objective extensions of CBMs for modeling communities: multi-objective flux balance analysis (MO-FBA) and multi-objective flux variability analysis (MO-FVA). Both methods were applied to a hot spring mat model ecosystem. As a result, multiple trade-offs between nutrients and growth rates, as well as thermodynamically favorable relative abundances at community level, were emphasized. We expect this approach to be used for integrating genomic information in microbial ecosystems. Following models will provide insights about behaviors (including diversity) that take place at the ecosystem scale.

Introduction

Microbial organisms comprise approximately 50% of the Earth's biomass [1, 2] and their interplay drives most biogeochemical cycles [3, 4]. The study of microbial interactions, which occur at the molecular scale, remains crucial to the elucidation of larger-scale processes [5]. Several models have attempted to simulate the quantitative impact of molecular-scale processes at an ecosystem level. Among others, trait-based approaches have gained attention as a precise way to understand and predict the quantitative behaviors of microbial communities [6, 7]. However, such models remain difficult to apply to most communities without the additional expertise required for deciphering particular traits and performing extensive experiments to design accurate parameters [8]; such expertise is often unavailable for the study of natural communities.

In the last decade, great advances have been made in the development of high-throughput techniques that enable the study of the metagenomics, meta-transcriptomics, and meta-metabolomics of natural communities. Such techniques provide ‘omics-scale information for organisms, from which it is possible to identify specific molecules (*e.g.*, DNA, mRNA, metabolites) present in a particular microbial ecosystem. Such studies of microbial ecosystems have facilitated drastic changes in approaches utilized for characterizing microbial communities [9, 10], thus leading to the emergence of the field of microbial systems ecology. Further, advances in bioinformatics and computational techniques have enabled the development of next-generation sequencing technologies for the qualitative analysis of microbial environments by emphasizing *who is there and who is not* [11] and allowing the study of the co-existence of microbial strains under different environmental conditions (see [12] for illustration). However, among the most significant challenges in modeling microbial communities remains the ability to quantitatively predict microbial community composition and functions under specific environmental conditions.

We propose to overcome this challenge by using recent systems biology approaches for the prediction of quantitative behaviors of single organisms based on genome-scale data [13, 14]. This study presents a natural extension of such approaches via their application to the modeling of microbial ecosystems and the elucidation of their quantitative features [15, 16].

Genome-scale descriptions, in this context, are provided by metabolic networks. A metabolic network summarizes the set of biochemical reactions encoded by the genome of a given organism. Two reactions are linked within a metabolic network if the substrate of one reaction is the product of the other. Such genome-scale descriptions of organisms are currently applied in systems biology for the purpose of investigating physiology [17]. In particular, for an increasing number of species, current bioinformatics protocols build genome-scale metabolic networks from genome-scale transcriptomic or metabolomic data [18].

Quantitative analyses utilize such metabolic networks as inputs for constraint-based models (CBMs) in order to infer physiological features based on a genome-scale description [17]. As a central assumption, constraint-based modeling considers the constraints defined by the set of reactions as linked within a metabolic network at steady state, and assume the corresponding model to behave optimally to achieve a given objective [13, 14]. The use of constraint-based modeling for microbial ecosystems, which involves the generation of a framework to perform data integration as well as mathematical descriptions useful for numerical simulations, seems promising [16, 19].

Several attempts have been made to model the metabolic network of microbial communities. Rodríguez *et al.* [20] proposed to use a “supra-organism” assumption, which considers reactions of all members of the community as a single entity. While such an approximation was used in recent studies (see Biggs *et al.* [21] and Perez-Garcia *et al.* [22] for a review), Kiltgord and Segré [23] previously showed that fluxes from a compartmentalized network and its de-compartmentalized counterpart (*i.e.*, supra-organism approach) are significantly different in their predicted FBA and FVA values. Furthermore, they show that fluxes using both assumptions are often not correlated. Such a distinction between both modeling results, along with the indisputable presence of compartments within ecosystems, clearly advocates for the use of compartments in the modeling. Considering so, several modelings have been proposed. However, while they all assume to consider distinct compartment for each microbial strain involved, they differ in their use of choosing the objective function. Stolyar *et al.* [24] first proposed a compartmentalized flux balance approach for modeling a mutualistic co-culture that requires an “ecosystem function”. Such a function is usually a weighted sum of each compartment objective. Nevertheless, the relative weight of each strain objective function remains

herein at the discretion of an empirical expertise that is mostly out of reach for complex or uncharacterized microbial ecosystems.

To overcome such a weakness, more elaborated modeling approaches have been proposed. Zomorodi and collaborators [25, 26] modeled each organism in a microbial community as a single CBM with its own objective function, nested within a global ecosystem model, thereby enabling the maximization of an ecosystem objective function. This approach still requires to design an ecosystem objective function but proposes a multi-level optimization that considers both microbial strain and ecosystem objectives. Meanwhile, Khandelwal and collaborators [27] (followed by [28]) advocates for the use of the “balanced growth” concept, according to which all microorganisms grow at the same rate. Accordingly, this approach considers several compartments with no ecosystem objective function per se but rather introduces community fractions into the formulation, adding new degrees of freedom to the general optimization problem. Worth noticing, such a modeling assumption is justified for microbial communities for which biomass production is monitored and constrained in chemostat, but not necessary for open systems as observed in nature.

In this study, we propose a complementary model, to investigate the general case of microbial ecosystems. Based on Pareto optimality [29], we aim at describing all the feasible solutions considering metabolic constraints from each strain with no design of ecosystem function. Consistent with previous works, the present study considers the community as a compartmentalized system in which each organism (*i.e.*, a compartment) has (i) its own objective to optimize and (ii) shares metabolites through the environment. Contrary to above methods, our approach is based on multi-objective optimization, which allows us to consider the objective function of each organism simultaneously.

Specifically, following previous works, we implemented a multi-objective flux balance analysis method [30], henceforth known as MO-FBA, for microbial communities, which is based on an exact resolution algorithm. Additionally, we introduced a complementary multi-objective flux variability analysis (MO-FVA) method. These analyses emphasize putative metabolic behaviors that are optimal at the community level, while considering metabolic constraints for each strain. Finally, we performed complementary thermodynamics analysis [31], which enabled us to pinpoint (i) favored ecosystem responses to environmental parameters and (ii) the corresponding diversity.

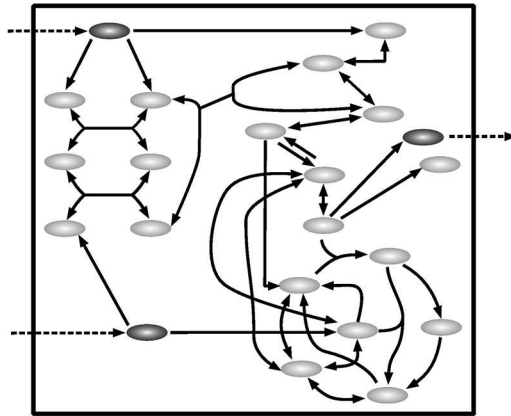
For the sake of MO-FBA and MO-FVA illustration, this study models a microbial ecosystem comprising three distinct phenotypes: a primary producer, *Synecococcus spp.* (SYN), filamentous anoxygenic producers (FAP), namely *Chloroflexus spp.* and *Roseiflexus spp.*; and sulfate-reducing bacteria (SRB, composed by *Thermodesulfovibrio spp.*-like activity, [32]), as described in [33]. Results emphasize trade-offs between distinct bacterial growth rates based not only on environmental conditions and genome-scale descriptions of each strain, but also thermodynamical quantitative predictions that are consistent with experimental knowledge.

Material and methods

Metabolic networks as constraint-based models

The genomic data for a particular microorganism describes a set of genes, allowing the identification of enzymes and related reactions. Reactions produce metabolites that are used as substrates in subsequent reactions; such interplay constitutes a “*metabolic network*” whose size may vary from few tens to several hundreds of reactions [14]. Metabolic networks are modeled (Fig 1A) in order to study the physiology of the relevant microorganism. In particular, metabolic models are used to infer reaction rates, also known as fluxes, without using kinetic parameters. For this purpose, a metabolic model is formally described by its stoichiometric

A



B

	R_1	R_2	R_3	...	R_{r-2}	R_{r-1}	R_r	
Metabolites	1	0	0	0	0	0	M_1
	0	-3	-2	0	0	0	M_2
	-1	-1	0	0	0	0	M_3
	0	0	1	0	0	0	M_4
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	0	3	-1	-1	0	0	M_{n-2}
	0	1	0	0	-1	0	M_{n-1}
	0	0	1	0	0	1	M_n
		Reactions						

C

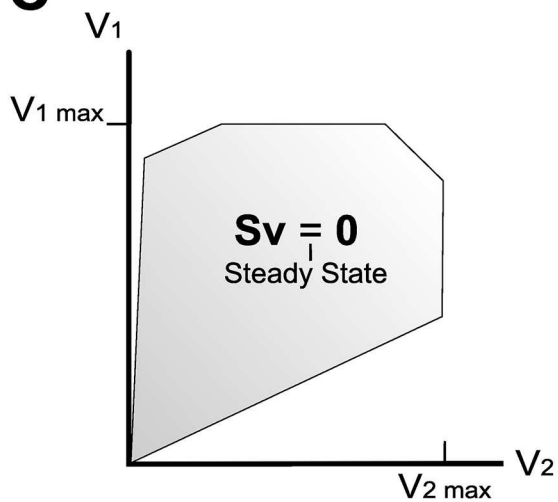


Fig 1. Construction of a Constraint Based Model (CBM). (A) **Metabolic Network** is represented as a chart of metabolites (ellipses) through chemical reactions (arrows); borders represent the system boundary. (B) depicts the **Stoichiometric Matrix**, in which reactions are presented as columns and metabolites as rows. Each coefficient S_{ij} of the matrix corresponds to the stoichiometric coefficient of metabolite M_i in reaction R_j , with reactants as negative and products positive. Exchange reactions and exchange metabolites are placed in the right and inferior section of the matrix, respectively. Therefore, submatrix ζ is in the left and highlighted in light gray while submatrix ξ is highlighted in dark grey (see text). Normal gray depicts a matrix with only zeros. (C) **Flux space**, also known as “solution space”, is defined by the set of restrictions of the CBM (mass balance in steady state, bounded reaction rates, etc.) and contains all possible values of \mathbf{v} .

doi:10.1371/journal.pone.0171744.g001

matrix \mathbf{S} (Fig 1B), where the rows correspond to the metabolites and the columns correspond to the reactions considered in the metabolic network. At steady-state conditions, the rate of formation of internal metabolites is equal to the rate of their consumption. This is expressed by the flux balance equation $\mathbf{S}\mathbf{v} = \mathbf{0}$, where $\mathbf{v} = (v_1, \dots, v_r)$ stands for the flux vector, *i.e.*, v_j is the flux of reaction R_j for all $j = 1, \dots, r$.

Under steady-state conditions, the continuous supply of metabolites from the media is facilitated by exchange reactions at a constant rate (dark gray ellipses and dashed lines in Fig 1A and highlighted dark gray block in Fig 1B). This matter exchange with the media allows the metabolic network to be in a non-equilibrium steady state (NESS). If metabolite exchange were not possible, then for each reaction the only possible state would be the chemical equilibrium, with all net fluxes equal to zero [31]. In the following, ζ and ξ represent, respectively, internal reaction and exchange reaction submatrices (light gray and dark gray blocks in Fig 1B, respectively). Occasionally, exchange rates may be experimentally measured and incorporated into the model as equations of the form $v_i = b$ for reaction i . In addition, maximal and minimal flux values may be expressed as *lower* and *upper* bounds constraints, by equations of the form $l_i \leq v_i \leq u_i$, resulting in a model described as a set of constraints. Such models are termed CBMs. CBMs usually comprise more reactions than metabolites; therefore, these models are undetermined in that when a solution \mathbf{v} exists, it is not unique. All feasible solutions define a “flux space” (Fig 1C) that may be further analyzed through several state-of-the-art approaches. For a detailed review of these methods, the reader may wish to refer to [13] and [14].

Flux balance analysis. Flux balance analysis (FBA) is one of the most widely used approaches for the identification of points of interest in the flux space [14]. Using this method, an objective function (for example, biomass production) is stated and its maximal value within the flux space is determined. In addition to the flux balance constraints, FBA utilizes flux capacity constraints that limit the fluxes of reactions. An optimal flux vector may be obtained by solving the following linear program (LP):

$$\begin{aligned} & \underset{\mathbf{v} \in \mathbb{R}^n}{\text{maximize}} && z = \mathbf{c}^T \mathbf{v} \\ & \text{subject to} && \\ & && \mathbf{S}\mathbf{v} = \mathbf{0} \\ & && l_i \leq v_i \leq u_i \quad i = 1, \dots, n, \end{aligned}$$

where $\mathbf{c}^T \mathbf{v}$ is a linear combination of fluxes that represents the objective function (*i.e.*, biomass production or growth rate). From linear programming theory, it is known that the optimal value z^* of objective function is unique; however, multiple flux distributions (*i.e.*, values of \mathbf{v}) that achieve the same optimal value z^* may exist.

Flux variability analysis. The set of all optimal flux distributions, *i.e.*, those with an optimal objective value of z^* , may be investigated by using Flux Variability Analysis (FVA) to

determine the flux range of each reaction in the metabolic network [14]. Formally, FVA solves the two following LPs for each reaction R_j :

$$\begin{aligned} & \underset{v_j \in \mathbb{R}}{\text{maximize / minimize}} && v_j \\ & \text{subject to} && \\ & && \mathbf{c}^T \mathbf{v} \geq \alpha \cdot z^* \\ & && \mathbf{Sv} = \mathbf{0} \\ & && l_i \leq v_i \leq u_i, \quad i = 1, \dots, n \end{aligned}$$

where $\alpha \in \mathbb{R}, 0 \leq \alpha \leq 1$ represents the fraction of the optimum value with respect to the FBA objective value to be considered. FVA allows the user to infer specific properties of the fluxes involved. For example, *essential* reactions have strictly positive or negative fluxes, whereas *blocked* reactions are constrained to have a flux value equal to zero.

Both FBA and FVA are today state-of-the-art tools to explore CBMs [13]. From a computational viewpoint, several algorithms are available to solve these optimization-based approaches (see section Solving Linear Optimization Problems).

Thermodynamic constraints metabolic networks. FBA and FVA utilize constraints derived from mass conservation laws; however, it is possible to exploit thermodynamic laws to derive constraints in order to obtain further insights into the behavior of a metabolic system [31, 34, 35]. In biochemical systems, each metabolite has an associated chemical potential μ_i (expressed in $\text{J} \cdot \text{mol}^{-1}$), which quantifies the potential to perform chemical work. Chemical potentials depend on metabolite concentration according to $\mu_i = \mu_i^0 + RT \ln(x_i/x_i^0)$, where x_i is the molar concentration, x_i^0 is the standard reference concentration (1 M) and μ_i^0 is the standard chemical potential (dependent on temperature, pressure, and ionic strength); these are usually tabulated [36, 37]. For a reaction j , the stoichiometric sum of the chemical potentials of the metabolites involved is equal to the Gibbs energy of the reaction, *i.e.*, $\Delta_r G_j = \sum_i^n S_{ij} \mu_i$ where $\Delta_r G_j \leq 0$ for a spontaneous reaction. In the following, we note the Gibbs energy of reaction as a difference of potentials, *i.e.*, $\Delta\mu_j \doteq \Delta_r G_j$.

Under NESS conditions, the entropy balance implies that $\Delta\mu^T \mathbf{v}_\zeta = \boldsymbol{\mu}^T \mathbf{v}_\xi$, where \mathbf{v}_ζ represents the internal portion of fluxes, \mathbf{v}_ξ boundary fluxes, and $\Delta\mu$ and $\boldsymbol{\mu}$ are vectors of components $\Delta\mu_j$ and μ_i , respectively. The term $\boldsymbol{\mu}^T \mathbf{v}_\xi$ represents the *chemical motive force* or *cmf* of the network, which accounts for energy related to boundary fluxes [31]. This equation may be interpreted as internal fluxes being driven by the consumption of external chemical potential.

The integration of such equations into general CBMs is not straightforward, as in most of applications, concentrations x_i are not known; therefore, these must be introduced as variables. As a result of non-linear expressions, CBM formulations using these constraints are generally more complex to solve [38–40].

Solving linear optimization problems. In general, optimization problems are aimed at determining $f(\mathbf{v})$ where \mathbf{v} is usually required to satisfy constraints. Linear optimization problems (LPs) are a particular kind of optimization problem where both objective function and constraints may be expressed as linear functions of variables, *i.e.*, $\max f = \mathbf{c}^T \mathbf{v}, \mathbf{Av} = \mathbf{b}$; where \mathbf{v} is a vector of variables, \mathbf{c} is a row vector of n coefficients, \mathbf{A} is a matrix of n columns and m rows, and \mathbf{b} a column vector of m values. The solution space of LP problems are polyhedrons that are characterized by their extreme points.

The first algorithm to solve a LP, which was proposed in 1947 by Dantzig [41], was based on the fact that if the objective function has an optimum value in the feasible region, then it reaches this value in at least one of the extreme points. The algorithm begins its search in one

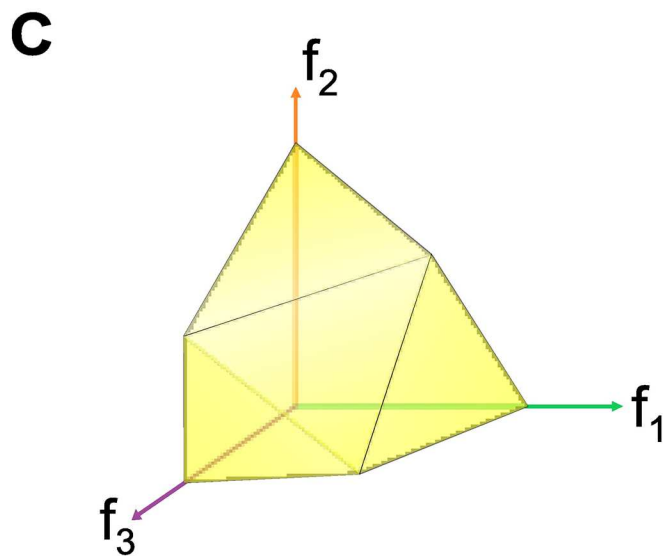
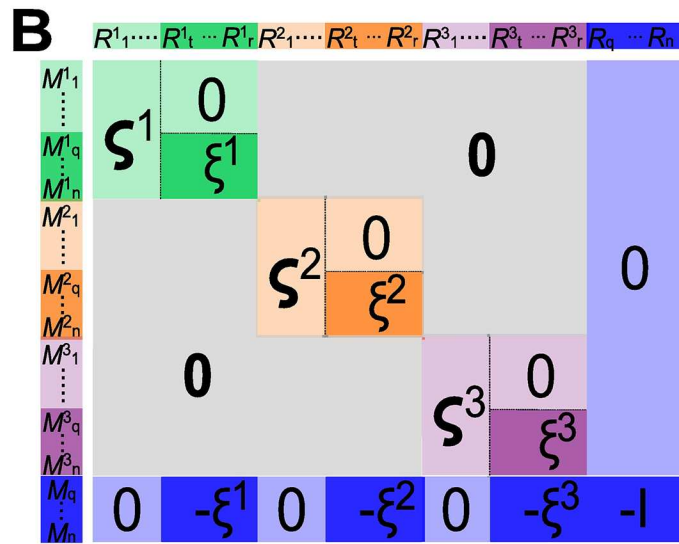
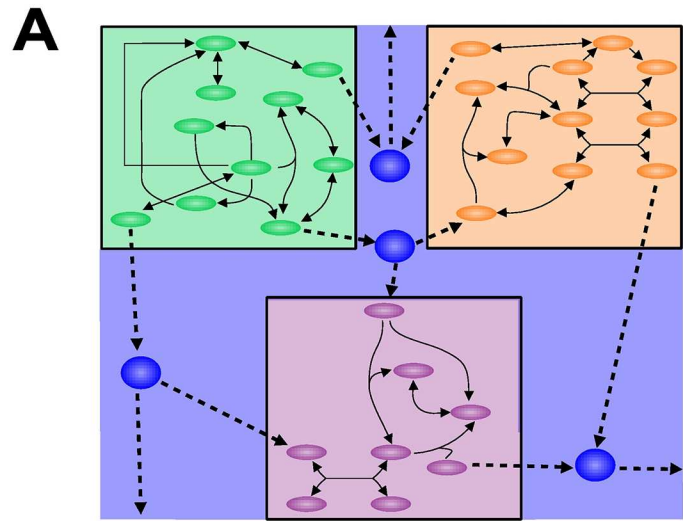


Fig 2. Illustration of microbial ecosystem CBM. For the sake of illustration, an ecosystem may be considered to comprise three microbial strains. (A) According to the metabolic model, each microorganism is considered a separate compartment, depicted here in green, orange, and purple. Metabolic networks are linked via an additional compartment, termed the “pool” (blue), which sums up all external metabolites exchanged between organisms and the environment. (B) depicts the Stoichiometric Matrix S^σ , where each compartment is colored accordingly, with their corresponding ζ and ξ submatrices. (C) Pareto front. When performing an FBA for multiple organisms, a set of points known as the Pareto front (in yellow) is obtained. Objective functions f_1 , f_2 and f_3 define the “objective space”.

doi:10.1371/journal.pone.0171744.g002

vertex of the feasible region and then starts visiting adjoint vertexes until the objective function value cannot be improved. Currently, several solvers such as GUROBI [42] or GLPK are capable of solving LPs and other types of single objective problems (SOPs) efficiently.

From single microorganisms to microbial ecosystems

In order to model a microbial community, each strain is considered a single compartment [19, 25, 27] that shares metabolites with other strains (see Fig 2A). As the stoichiometric matrix of a single organism, the structure of the ecosystem is described by a stoichiometric matrix S^σ , which is formed by the stoichiometric matrices of each single organism. Accordingly, for a community of k microorganisms, k metabolic models must be considered and represented by their corresponding stoichiometric matrices: $S^l, l = 1, \dots, k$.

As shown in Fig 2B, matrices S^1 to S^k are used to construct a diagonal block matrix. Each block is linked to a *pool compartment*, that mirrors exchange fluxes between each organism and the environment ($-\xi^l$, for $l = 1, \dots, k$ in Fig 2B). A set of exchange reactions R_q to R_n for metabolites M_q to M_n between the Pool and the external environment, is additionally set (bottom right in Fig 2B). Finally, as for single organisms, a steady state hypothesis restricts the solution set by adding a constraint $S^\sigma v = 0$. Together with flux bound constraints l_i and u_i , these constraints describe a solution flux space, as depicted in Fig 1C.

Multi objective flux balance analysis of a microbial ecosystem. Each compartment above corresponds to an organism with a specific objective function c_k . Accordingly, the following multi-objective optimization problem, for analyzing flux balance conditions (MO-FBA), may be defined:

$$\begin{aligned} & \underset{v \in \mathbb{R}^{\bar{n}}}{\text{maximize}} && \begin{pmatrix} f_1 \\ \dots \\ f_k \end{pmatrix} = \begin{pmatrix} c_1^T v \\ \dots \\ c_k^T v \end{pmatrix} \\ & \text{subject to} && S^\sigma v = 0 \\ & && l_i \leq v_i \leq u_i \quad i = 1, \dots, \bar{n} \end{aligned}$$

where $(f_1, \dots, f_k)^T$ are the objective functions of the k organisms and \bar{n} is the total number of reactions (*i.e.*, the sum of reactions of each organism and exchange reactions from the pool compartment). The general class of MO-FBA problems is referred to as the *multi objective problems* (MOP) [29, 43]. Contrary to single objective problems, solution of MOPs is a set of vectors instead of a single value, producing a Pareto front (see section Solving Multi Objective Optimization Problems), defined in the objective space (Fig 2C). In our present formulation, all constraints and objective functions are linear, thereby resulting in a particular type of MOP known as the multi-objective linear problem (MOLP).

Interpretation of MO-FBA can be done in terms of growth rates and resources used to produce such growth. Indeed, if one of the members of the ecosystem decreases its growth rate, more resources are available for other members. According to their particular physiologies, they can use these new available resources to increase their own biomass. A guideline containing three ideal cases for two guilds is provided in [S1 File](#).

Flux variability analysis of a microbial ecosystem. Given a particular point \mathbf{f}^* of the Pareto Front, the multiple optimal flux solutions that achieve the optimal objective values, as given by the Pareto optima \mathbf{f}^* , must be explored. To this end, we propose the use of the multi-objective FVA (MO-FVA) for multiple organisms, which may be considered a straightforward extension of FVA (see Flux Variability Analysis). Indeed, given a reaction R_j with $j = 1, \dots, \bar{n}$, the range of the flux v_j may be determined by solving the following LPs:

$$\begin{aligned} & \underset{v_j \in \mathbb{R}}{\text{maximize / minimize}} \quad v_j \\ & \text{subject to} \\ & \mathbf{C}^T \mathbf{v} \geq \alpha \cdot \mathbf{f}^* \\ & \mathbf{S}^\sigma \mathbf{v} = \mathbf{0} \\ & l_i \leq v_i \leq u_i \quad i = 1, \dots, \bar{n} \end{aligned}$$

where \mathbf{C} is the matrix such as the column j corresponds to objective function \mathbf{c}_j , *i.e.*, \mathbf{C} is column defined as $\mathbf{C} = [\mathbf{c}_1, \dots, \mathbf{c}_k]$. $\alpha \in \mathbb{R}$, $0 \leq \alpha \leq 1$ is the fraction of the optima considered.

Thermodynamics analysis in the context of a microbial ecosystem. Biological systems are hypothesized to favor thermodynamic states where entropy production is maximal [44, 45]. To take into account this hypothesis, given a particular point \mathbf{f}^* of the front, we propose the following: First, a MO-FVA must be applied to determine R_j for each reaction, with $j = 1, \dots, \bar{n}$ and the range $[a_j, b_j]$ of the flux v_j near the Pareto optima \mathbf{f}^* . Next, the following optimization problem must be considered:

$$\begin{aligned} & \underset{i \in \xi}{\text{maximize}} \quad cmf = \sum \mu_i v_i \\ & \text{subject to} \\ & a_i \leq v_i \leq b_i, \quad i \in \xi, \\ & \mu_i^0 - dg_i \leq \mu_i \leq \mu_i^0 + dg_i, \end{aligned}$$

where ξ is the set of exchange reactions and $dg_i = RT \ln(x_i/x_i^0)$. As cmf is non-linear, optimization algorithms based on heuristics must be used in order to obtain a numerical solution to this problem (see Computational Procedures).

Solving multi objective optimization problems. In 1906, Vilfredo Pareto in his *Manuale di Economia Politica*, stated that, while (economic) optima have not been achieved, it is possible to increase the objective of an agent (*i.e.*, welfare) without decreasing that of another [46]. In the following, a formal definition of Pareto optima and efficient solutions is given [43] and approaches to solutions are discussed.

Let $\mathcal{X} \subseteq \mathbb{R}^n$ and $\mathcal{Y} \subseteq \mathbb{R}^p$ represent the flux space and objective space, respectively, where \mathcal{X} is defined by the set of restrictions and $\mathcal{Y} := \{\mathbf{y} \mid \mathbf{y} = \mathbf{f}(\mathbf{x}), \mathbf{x} \in \mathcal{X}\}$, with $\mathbf{f} = (f_1(\mathbf{x}), \dots, f_p(\mathbf{x}))^T$ denoting the objective functions. If both \mathcal{X} and \mathcal{Y} are constructed using linear restrictions and linear objective functions, the MOP represents a MOLP.

A point $\mathbf{y} \in \mathcal{Y}$ is a **Pareto optimum** if there is no $\mathbf{y}^* \in \mathcal{Y}$ such as $y_j^* \geq y_j, j = 1, \dots, p$ and $\mathbf{y} \neq \mathbf{y}^*$. Similarly, \mathbf{y}^w is a **weak Pareto optimum** point if there is no \mathbf{y}^* such as $y_j^* > y_j^w, j = 1, \dots, p$. A point $\mathbf{x} \in \mathcal{X}$ is an **efficient** solution if there is not a $\mathbf{x}^* \in \mathcal{X}$ such that

$\mathbf{f}(\mathbf{x}^*) \geq \mathbf{f}(\mathbf{x})$. A $\mathbf{x}^w \in \mathcal{X}$ is a **weak efficient** solution if there is no $\mathbf{x}^* \in \mathcal{X}$ such as $\mathbf{f}(\mathbf{x}^*) > \mathbf{f}(\mathbf{x}^w)$. Therefore, a (weak) Pareto optimum is the image of a (weak) efficient solution. Note that all efficient solutions are also weakly efficient solutions but no vice-versa. The collection of Pareto optimal points is termed **Pareto Front**.

Approaches for solving MOPs have been reviewed, for example, by [43] and [47]. Traditional approaches makes use of “scalarization techniques”, that involve the transformation of the MOP into a SOP by using a real-valued scalar function of the objective functions. Solution approaches using scalarization techniques aim to find the set of (weak) efficient solutions $\mathbf{x}^* \in \mathcal{X}$.

The most well known approach is the “weighted sum approach”, wherein the weighted sum of the objective functions is optimized, *i.e.*, $\max \sum \lambda_k f_k(\mathbf{x})$, where $\mathbf{x} \in \mathcal{X}$ and $\lambda \in \mathbb{R}^p$ is a given weight vector with components $\lambda_k \geq 0$ and at least one $\lambda_k > 0$. If \mathbf{x}^* is a solution of this SOP then \mathbf{x}^* is an efficient solution of the MOP. Furthermore, if the MOP is convex, the inverse is also true.

Another commonly used approach is the “ ϵ -constraint method”, where only one objective function is retained as the objective and the remaining objective functions are used to introduce new constraints. Then, the j -th ϵ -constraint problem is as follows: $\max f_j(\mathbf{x})$, subject to $f_i(\mathbf{x}) \geq \epsilon_i$, $i \neq j$ and $\mathbf{x} \in \mathcal{X}$. If \mathbf{x}^* is a solution of this SOP, then \mathbf{x}^* is a weak efficient solution of the MOP.

Not all approaches rely on scalarization: for MOLPs, a set of algorithms describing the shape of the image of efficient points, $\mathcal{Y}_E := \{\mathbf{C}\mathbf{x} \mid \mathbf{x} \text{ is efficient}\}$, referred to as “outer approximation” or “Benson type” algorithms, have been described [48–51]. Generally speaking, these type of algorithms calculate \mathcal{Y} and identify their vertices, which correspond to Pareto optimal points; additionally, despite their names, these algorithms provide exact solutions. BENSOLVE [52], a solver based on these approaches, computes a set of directions and points describing the image of the efficient points.

Existing CBM approaches for communities. The various approaches to studying microbial communities have been recently reviewed by Biggs *et al.* [21] and Perez-Garcia *et al.* [22]. Among the methods reviewed, OptCom most closely resembles the approach presented here, in that each member of the community is considered to maximize its own biomass. OptCom is based on bi-level optimization, where an “outer” maximization problem represents the whole community and each member of the community is represented by a “inner” optimization problem. Inner optimization problems are solved using the primal-dual theorem, which transforms the whole bi-level formulation into a non-convex single-objective form [25]. A second approach that combines compartments and FBA, known as community flux balanced analysis, advocates the application of a “balanced growth” hypothesis, wherein each compartment grows at the same rate. Furthermore, this approach considers the biomass fraction of each member of the community. In general, the approach is non-linear, although it may be made linear by fixing biomass fractions and solving the corresponding FBA. Then, optimal solutions for various combinations of biomass fractions may be explored [27]. For illustration purposes, the application of our approach to the analysis of a microbial ecosystem is discussed below.

Case study: Hot spring mat

In order to illustrate the application of the present approach, we modeled the microbial ecosystem of hot spring microbial mats [33]. Briefly, this ecosystem is composed of three *guilds*, representing three commonly found phenotypes: *Synechococcus spp.* (SYN), *Chloroflexus spp.* and *Roseiflexus spp.* (FAP) sulfate-reducing bacteria (SRB). SYN is a primary producer that fixes

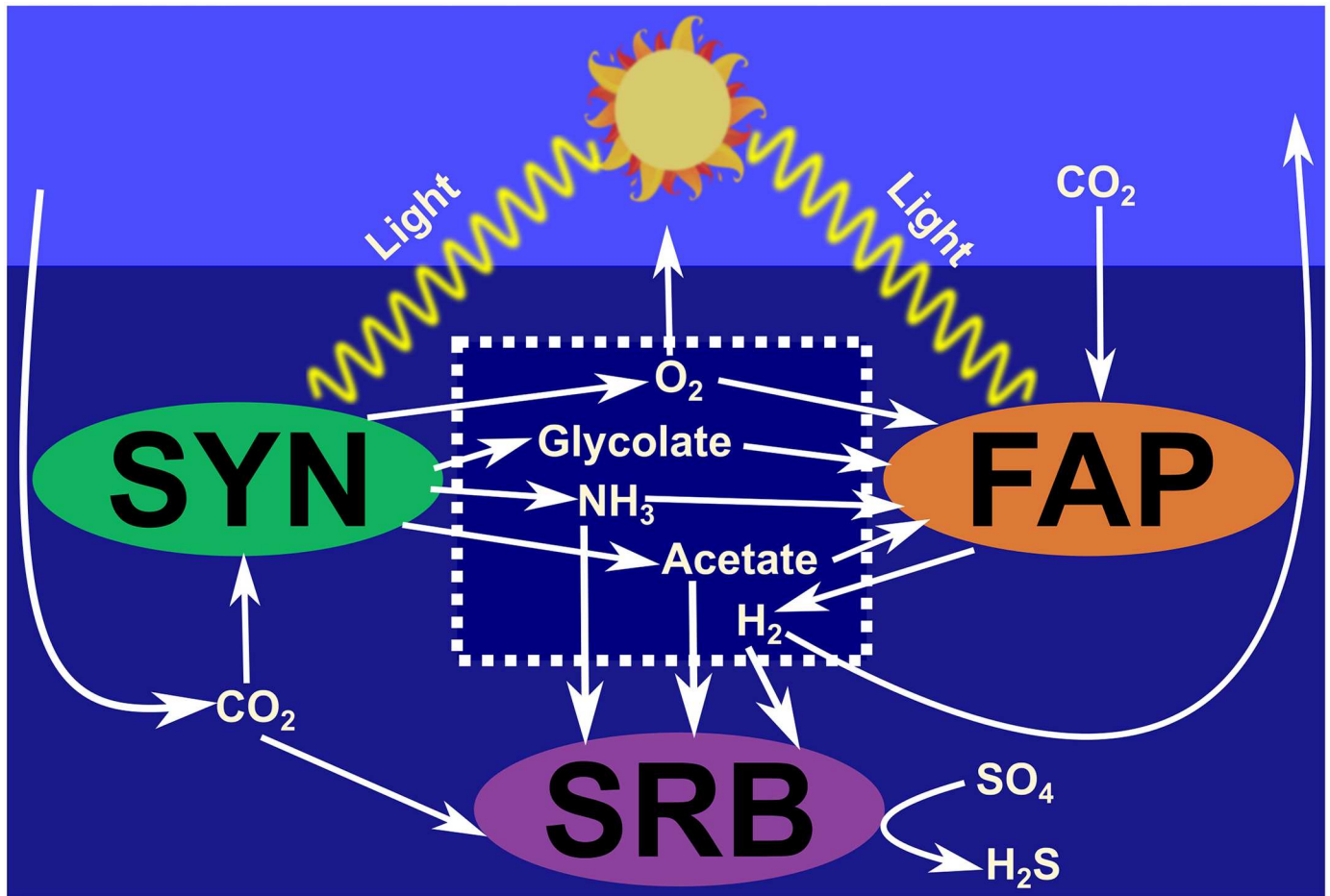


Fig 3. Day Model of the Hot Spring Mat Community. The model comprises three guilds of microorganisms of the SYN, FAP, and SRB phenotypes. Organics acids produced by SYN may be utilized by FAP and SRB. FAP is capable of fixing carbon by anoxygenic photosynthesis. Under anoxygenic fermentation conditions, FAP is additionally capable of producing hydrogen, which, in turn, may be used by SRB.

doi:10.1371/journal.pone.0171744.g003

carbon and nitrogen for further utilization by other strains. The use of these guilds allows simplification of the ecological diversity while capturing essential metabolite-exchange relationships. Under light conditions, the major fate of nutrients involves assimilation into cells [53]; therefore, most of the overall system growth occurs during the daytime. As growth rates are related to biovolumes, predictions may be compared with relative abundance data. Therefore, we will focus on the daytime model as described in [33] (Fig 3), assuming a simplified nighttime behavior, as described below.

Using the available compartment model of this system, as described in [33], we performed a manual curation (*i.e.*, balancing equations and including intermediate reactions) using METACYC [54]. Model equivalent reactions in [33] are provided in S2 File. Nitrogen fixation has been shown to take place at night and in the early morning [55, 56]; therefore, a nitrate assimilation mechanism for SYN was included and considered as functional. Finally, biomass coefficients of each guild were scaled to match 1 (h⁻¹) as maximal growth rate [57].

Glycolate is produced by the use of O₂ instead of CO₂ by the Rubisco enzyme; the flux ratio between the use of O₂ and CO₂ varies between 0.03 and 0.07. This restriction was included

linearly in the model by fixing a ratio of 0.03 between SYN reactions RXN-961 and RIBULOSE-BISPHOSPHATE-CARBOXYLASE-RXN during all calculations, under the hypothesis that the system is in anaerobic state.

Excess photosynthate producing during the day is stored as polyglucose (PG) by SYN. PG is fermented at night, producing several organic acids that accumulate in the media and are integrated as biomass mostly under light conditions [53, 58]. In order to capture this behavior in the daytime model, PG was not allowed to accumulate; therefore, the excess photosynthesis activity is redirected through acetate production. Accordingly, in our model, acetate is interpreted as equivalent to several forms of reduced carbon.

For each of the exchanged metabolites, standard Gibbs energies for biological conditions were obtained from [37], using calculations from [36]. Values used are found in S2 File. For the pseudo-compound $h\nu$ (representing photons), a standard chemical potential was estimated based on glucose synthesis from CO_2 : $6\text{CO}_2 + 6\text{H}_2\text{O} \xrightarrow{48 h\nu} \text{C}_6\text{H}_{12}\text{O}_6$. The assumption that this reaction approaches equilibrium at standard biological conditions (*i.e.*, $\Delta\mu = 0$) implies that $\mu_{h\nu} = 68.6 \text{ kJ}\cdot\text{mol}^{-1}$ (S2 File). The metabolite concentration was allowed to vary between 10^3 and 10^{-3} M, and therefore chemical potential equals $\mu_i = \mu_i^0 \pm dg$, where $dg = RT\ln(10^3) \approx 20 \text{ (kJ}\cdot\text{mol}^{-1})$ for $T = 75^\circ\text{Celsius}$. For water and $h\nu$, concentrations were considered as fixed at 1 M, implying $dg_{\text{H}_2\text{O}} = dg_{h\nu} = 0$.

Computational procedures

For each guild, a metabolic model was built in MATLAB and an ecosystem stoichiometric matrix \mathbf{S}^σ was constructed, as described above. MO-FBA was carried out using BENSOLVE [52]. In order to analyze nitrogen and carbon fluxes through MO-FBA results, a MO-FVA was performed using GUROBI [42] through Python interface over a mesh of 5 151 equally distributed points in the Pareto surface at 90% fraction of optimum. Then, we subdivided the Pareto surface into 225 similar regions; for each of these regions, we calculated their maximum (as well as their minimum) as the average of MO-FVA maxima of mesh points contained (this procedure was repeated for the minima). Thermodynamics calculations were performed over the same mesh as the MO-FVA using a truncated Newton conjugate algorithm [59] contained in scipy optimization module. Heatmaps and surface illustrations were generated using matplotlib [60] with *ad-hoc* scripts.

From methods discussed in Biggs *et al.* [21] and Perez-Garcia *et al.* [22], OptCom [25] was chosen for comparison, as this method resembles the approach applied to the present work. We applied OptCom and Descriptive OptCom to the hot spring mat model, as follows: first, 11 points were calculated using OptCom, as described by [25], each with a different upper boundary value for SYN biomass; these values ranged from 1.0 to 0.0 with a step of 0.1 (*i.e.* 1.0, 0.9, 0.8, . . . , 0.0). Second, Descriptive OptCom was applied three times using SYN to FAP ratios of 1.5, 2.5, and 3.5, respectively. All programs were written in GAMS language and solved using BARON [61] through the NEOS Server [62–64].

All scripts are available in <https://gitlab.univ-nantes.fr/mbudinich/MultiObjective-FBA-FVA>

Results

Biomass distribution as relative microbial strain abundance

SYN, SRB, and FAP growth rates are represented in a 3-dimensional space, in each axis, respectively, in Fig 4A. MO-FBA solutions are described as a Pareto front, representing a surface with five extreme points of biomass growth: (1, 0, 0), (0, 1, 0), (0, 0, 1); the points

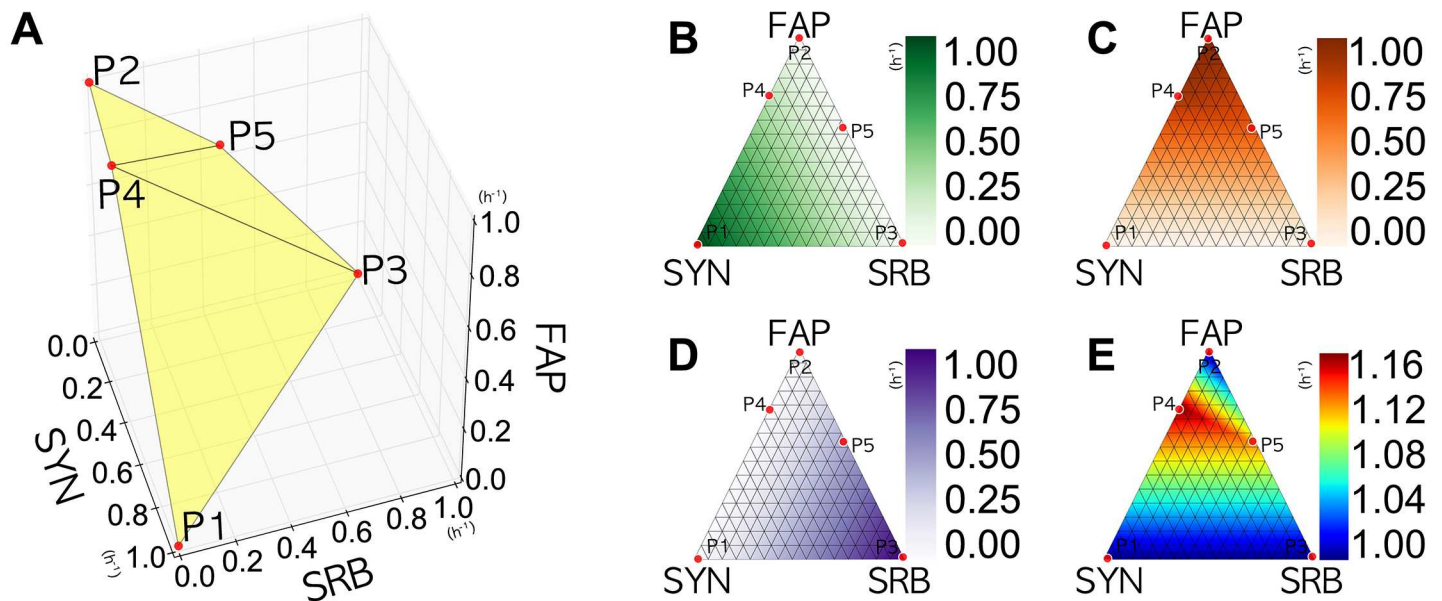


Fig 4. 3D and 2D Projections of Pareto Front. (A) shows a 3D Pareto front, in yellow, describing the maximal growth rates of SYN, FAP, and SRB (in terms of units per hour, h^{-1}), when considered as a system. It is evident that a decrease in the growth rate of one organism results in an increase in that of the other two, but not necessarily in equal proportions (see [S1 Video](#) for an animated view). The sum of the growth rates of all the guilds in P4 and P5 was $1.16 (h^{-1})$ and $1.11 (h^{-1})$, respectively. In (B), (C), (D), and (E), the Pareto front was projected onto the triangular surface formed by P1, P2, and P3. (B), (C), and (D) shows the respective growth rates for SYN, FAP, and SRB, respectively. (E) shows the sum of the three growth rates, which represent the total biomass of the ecosystem.

doi:10.1371/journal.pone.0171744.g004

corresponding to the maximal growth rates of each guild, and points (0.27, 0.00, 0.89) and (0.00, 0.46, 0.65). In the following, these points are designated P1, P2, P3, P4, and P5, respectively. For clarity, this Pareto front is then projected in a two-dimensional space. Therefore, over a triangular surface defined by P1, P2, and P3, heatmaps were produced using the values for the growth rate of SYN, FAP, SRB, as well as their sum, to depict the overall microbial abundance (Fig 4B–4E, respectively). Each vertex of the triangle represents the maximal growth rate of a guild, while its opposing side represents a zero growth rate for that guild.

The results show that when each guild grows at its maximal rate, no biomass is produced by the other guilds. The sum of the growth rates is always minimal in vertices (blue areas in Fig 4E). As the growth rates may be directly related to biovolumes [33], red to yellow areas in Fig 4E represent regions where most of the total biomass of the ecosystem is present. Notably, these regions correspond to guilds growing at sub-optimally rates.

Nitrogen and carbon fluxes between microbial guilds

Multi-objective FVA was performed in the P4 and P5 regions to explore NH_3 import and export fluxes between guilds (Fig 5A, upper and lower panel, respectively). Notably, the growth rate of each strain was found to be related to the use of ammonia; the SYN guild re-oxidized ferredoxins, which were reduced in the photosynthetic reactions, via nitrate assimilation reactions, thereby promoting permanent ammonia production. When growing sub-optimally, NH_3 that is not used to build biomass is excreted. This point is emphasized in Fig 5A, where both maximal and minimal reaction rates are strictly positive for SYN, resulting in an export to the pool.

Nitrogen uptake by FAP and SRB occurs solely from ammonia that is available in the pool compartment; therefore, these strains compete for its intake. When SRB is not growing

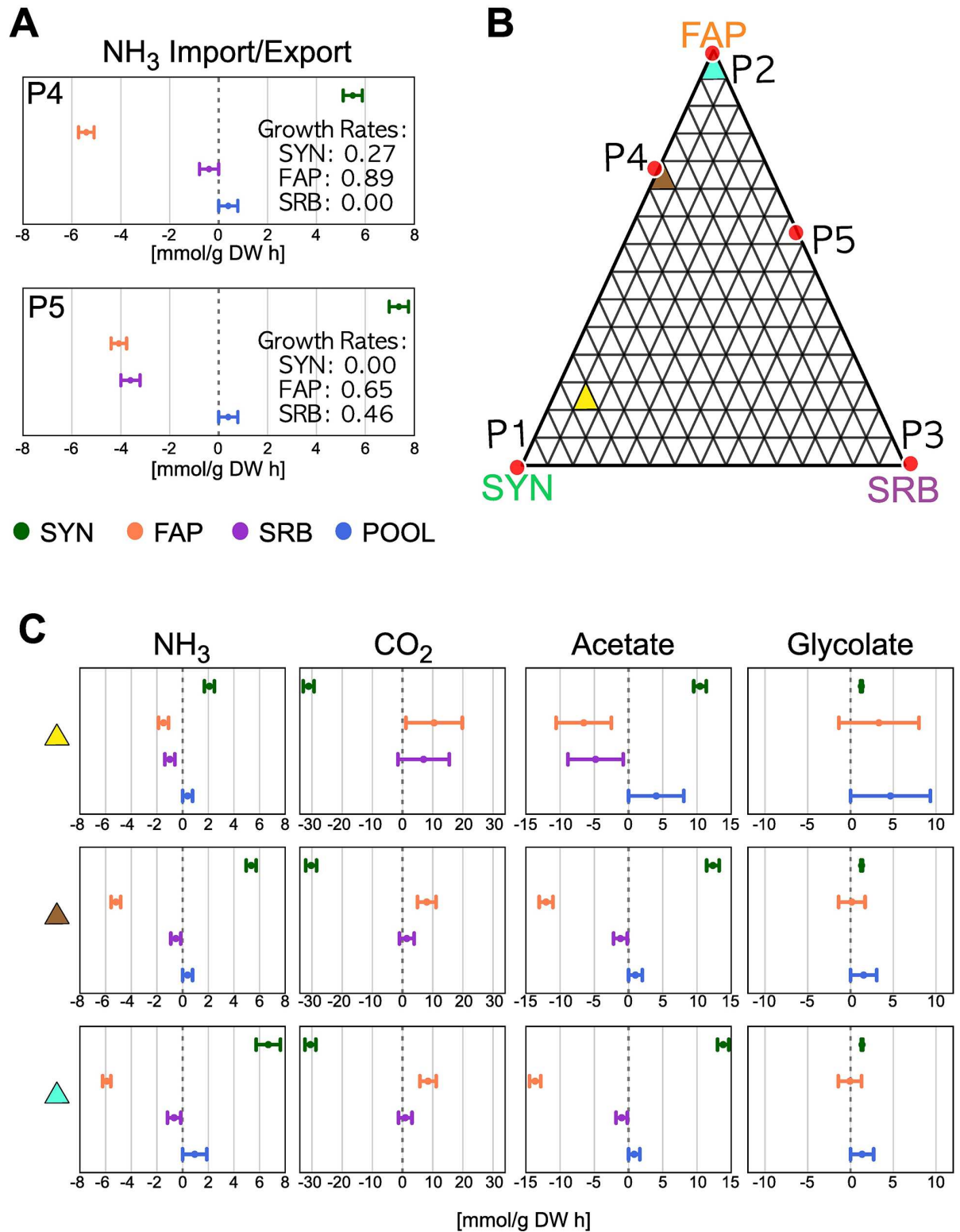


Fig 5. Multi Objective FVA. (A) shows NH₃ maximal and minimal fluxes for SYN, FAP, SRB, and pool compartments (green, yellow, purple, and blue respectively) for extreme points P4 and P5. The export of NH₃ by SYN is correlated with a drop in their growth rate; similarly, increases in NH₃ intake are correlated with increases in the growth rates of FAP and SRB. (B) Three sections selected for the illustration of MO-FVA; (C) Mean values of the minimal and maximal fluxes over selected sections of NH₃, CO₂, acetate, and glycolate (columns) for each section.

doi:10.1371/journal.pone.0171744.g005

(superior panel in Fig 5A), excess of NH_3 is taken up mainly by FAP (both minima and maxima are negative, implying an intake from the pool). Small amounts that are not taken up by FAP may be either taken up by SRB (maximal rate value is null and minimal rate negative, which depicts a possible import) or excreted to the external environment (pool maximal rate value is positive and minimal rate value is null, which depicts a possible export to the media). When SRB is growing (inferior panel of Fig 5A), the uptake rate of ammonia by SRB and FAP is similar, with no export to the external media.

In order to analyze the relationships between the growth rate of each strain and nitrogen- or carbon-related fluxes, we performed a MO-FVA as described in Computational Procedures, focusing on exchange reactions. For the purpose of illustration, we highlighted three sections from 225 calculated, as shown in Fig 5B. These regions were chosen to depict the theoretical interplay between SYN and FAP when the growth rate of SRB is low [65]. Flux variability of exchange fluxes for these regions is shown in Fig 5C (see S1 Fig for an alternative representation and S2 to S5 Figs for a complete MO-FVA for ammonia, acetate, carbon dioxide and glycolate fluxes).

For NH_3 exchange reactions, high growth rates of SYN are related to lower levels of ammonia export, which represents a limiting factor for FAP and SRB growth rates. This results in the two strains competing for its use (S2 Fig). Fig 5C shows that most of the ammonia produced by SYN is captured by FAP, while a small proportion is taken up by SRB. Ammonia that is not captured is released into the pool.

SYN consumes approximately the same amount of CO_2 under all relative abundance conditions (see second column in Fig 5C and S4 Fig), indicating that carbon compounds are involved in reactions that serve functions other than biomass synthesis. Acetate intake by FAP is less restrained at low growth rates of SYN than at high growth rates (see Fig 5C and S3 Fig).

The present results additionally emphasize that FAP and SRB produce relatively small amounts of CO_2 at low growth rates. However, when the growth rate of FAP increases, the maximal excretion of CO_2 reduces, whereas its minimal excretion increases; these data indicate the theoretical efficiency of carbon management, as experimentally reported by [53]. Glycolate metabolism by FAP appears to be reversible as its minimal flux is negative (*i.e.*, intake) while its maximal flux is positive (*i.e.*, excretion), implying that intake or excretion by FAP is related to the relative abundance of other strains (see Fig 5C and S5 Fig for details).

Chemical potentials drive community growth rates

As discussed previously, the direct integration of thermodynamic constraints into MO-FBA and MO-FVA formulations is complex. Instead, we used the thermodynamic optimization problem stated in as a post-treatment analysis. Considering fluxes as computed by MO-FVA in 5 151 points of Pareto front (as a result of which growth rates are also determined), we estimated the corresponding maximal *cmf* for each point (Fig 6A).

Results show that higher *cmf* is associated with SYN growing at its optimal rate. Lower *cmf* rates are related to a higher growth rate of SRB, whereas the impact of the growth rate of FAP on the value of *cmf* appears to be lower than that of SRB.

Given that all surface showed positive values, all regions are feasible from a thermodynamic viewpoint. Under the hypothesis that a biological system prefers configurations in which entropy production is maximal, it is expected that an ecosystem would favor growth rates with higher *cmf* (redder areas in Fig 6A), predicting higher SYN growth rates. This prediction is consistent with *in vivo* field measurements of SYN: FAP relative abundance ratios in the range of 1.5 and 3.5, with a low presence of SRB [33, 65], as shown in Fig 6B.

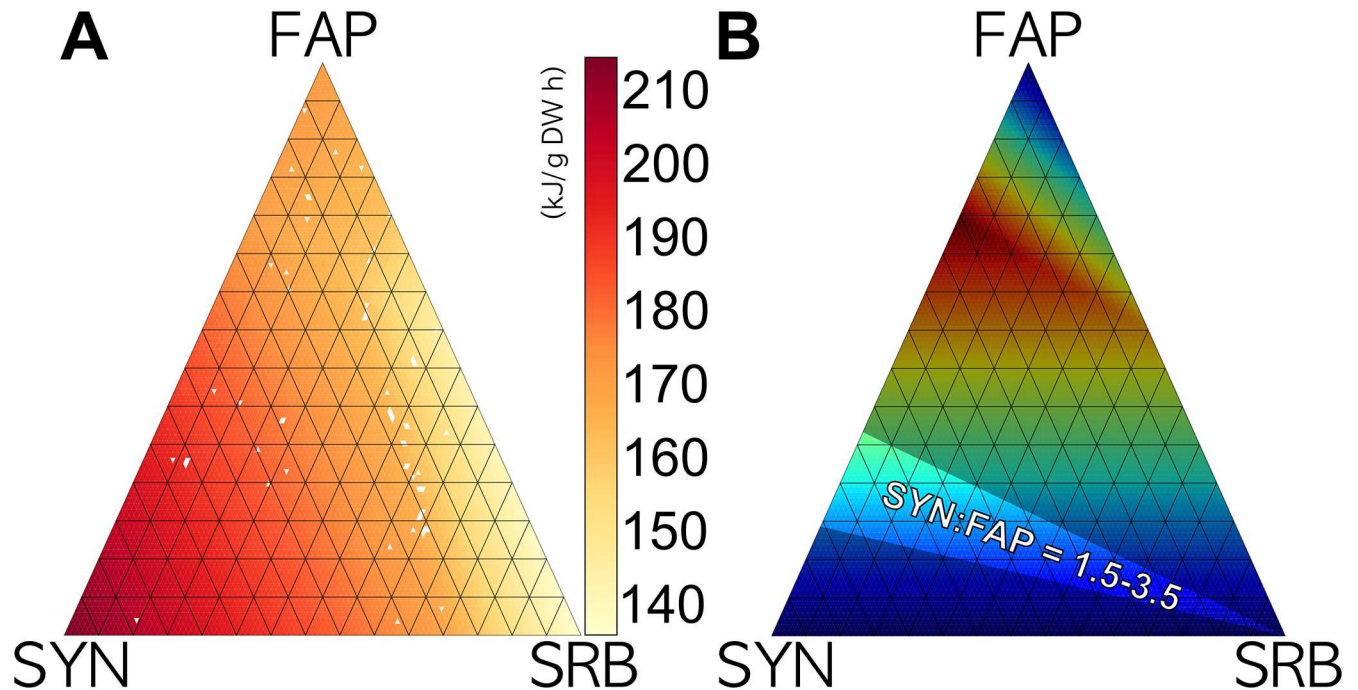


Fig 6. Thermodynamics in the Pareto front. (A) Description of the chemical motive force ($\text{kJ}\cdot\text{gr}^{-1}\cdot\text{DW}^{-1}\cdot\text{h}^{-1}$) for each point of the Pareto front; red regions indicate thermodynamically favored growth rates, while the points where the solver does not reach the optimal criteria are shown in white. The obtained surface appears smooth, without sudden changes in neighboring values. (B) Description of the overall community biomass distribution based on the growth rate of each strain, with a particular emphasis on regions supported by experimental measurements showing a SYN:FAP ratio of between 1.5 and 3.5.

doi:10.1371/journal.pone.0171744.g006

Comparison with previous approaches

We compared growth rates and flux predictions of MO-FBA and MO-FVA with those obtained by a comparable approach (OptCom [25]), as described in Computational Procedures. Predictions obtained were mapped as points in the Pareto front (S6 Fig). Values of growth rates, as well as their corresponding flux values for NH_3 , acetate, glycogen, and CO_2 , are described in S2 File. As expected, all points calculated using the OptCom approach were included in the Pareto front calculated by MO-FBA (S6 Fig). Furthermore, all flux predictions for NH_3 , acetate, glycogen, and CO_2 fall into the range predicted by MO-FVA. Without constraining SYN biomass (point O1), OptCom does not reach the maximal biomass optimum. However, when SYN biomass is increasingly constrained (points O2 to O11), the total biomass increases. This suggests the existence of local optima in the OptCom general formulation for this model.

The composition of a community that function in a constant environment can be also assessed using the approaches proposed in [27] and [28]. Here, we focus on modeling the composition of a community in a changing medium where the considered organisms could grow not necessarily with the same growth rate

Discussion

As reported in previous studies, in particular [25], we extended state-of-the-art systems biology constraint-based approaches to the modeling of microbial ecosystems, by considering a multi-objective optimization framework. Within the ecosystem, each microorganism, with its own

objective function, represents a building block that interacts with others via the exchange metabolites. Furthermore, the genomic knowledge of each microorganism is integrated as a set of metabolic constraints. The main advantage is represented by the capture of trade-offs on objectives and metabolite exchange between members of the ecosystem. While previous works report topological analyses that focus on pathways that promote cross-feeding between strains (see [66, 67] for example), this study quantifies fluxes through these pathways as well as their effect in objective functions, thereby representing a major step towards automatically producing trait-based models. Through the application of MO-FBA, we emphasize a full description of the Pareto front that captures trade-offs in the optimal values of the objective function of each microorganism. Additionally, we introduced MO-FVA as a tool for the analysis of exchange fluxes between members of the community. These fluxes help to characterize the optimal behavior of microorganisms, providing insights into the theoretical relative abundances (*i.e.*, a proxy for microbial diversity) and corresponding nutrients usage, that are based on *omics* descriptions.

Unlike previous works that consider multiple objectives, our approach does not rely either on assumptions about ecosystem behaviors, such as maximization of the total ecosystem biomass, ([25, 26]) nor on the balanced growth ([27, 28]) of microbial strains involved. Instead, we propose to describe all optimal solutions in the sense of Pareto in the objective space. This approach provides several advantages: firstly, it includes any solution for a system objective function expressed as a weighted sum of each compartment objective function (see [43] and section Solving Multi Objective Optimization Problems). Therefore, it comprises all solutions proposed by OptCom as system objectives for microbial communities [25]. Secondly, no additional complementary restrictions are required to focus on given solutions, *i.e.*, imposing an equal growth rate for all members, as proposed by Kandelwal et al. [27]. This restriction remains valid for controlled microbial ecosystems. Third, the set of constraints remains linear, which allows a description of the Pareto front for realistic ecosystems. In [25] and [26], formulations are, in general, non-convex; in [27], the stated general optimization problem is non-linear. However, in order to solve MOLPs, a series of LPs must be solved for which exact algorithms are fast, thereby reducing computational complexity. Note herein that the last two points are mandatory to model natural ecosystems that are by definition composed of a large number of microbial strains and mostly unconstrained.

For illustration purposes, we applied MO-FBA to the daytime part of the diurnal cycle of the microbial hot spring mat system [33]. As most biomass fixation occurs during the day phase [53], we assumed that daytime growth rates dominate overall ecosystem rates. Results show that the maximal total biomass growth rate is achieved when each guild grows at a rate below its theoretical maximum, which may, based on genomic knowledge, be interpreted as an altruistic behavior. Mechanistically, when guilds make resources available to others, they lower their objective value by a certain proportion, based on metabolic pathways used to synthesize those resources and their biomass function. Conversely, the use of new available resources increases the value of the objective functions of the other guilds. Therefore, the growth rate of the global maximal ecosystem, which was designated P4 in our case study, should correspond to the optimal resource allocation scenario from the ecosystem viewpoint. P4 also corresponds to the optimal solution to maximal ecosystem biomass [25].

MO-FVA results show that nitrogen flux is correlated to growth rates, and that the three guilds compete for their usage. In contrast, CO₂ consumption and glycolyte and acetate production by SYN do not seem to be correlated with its growth rate, indicating that these processes are not carbon-limited. Reduced carbon, represented by acetate, appears as being the main carbon flux in the system for FAP and SRB, and becomes a limiting nutrient for FAP at

high growth rates. This result is consistent with those of [53] and [58], in which a high proportion of reduced carbon was shown to be assimilated by FAP.

By coupling MO-FVA results with chemical potentials, we were able to analyze thermodynamic constraints and study favored conditions of the Pareto front by comparing their respective maxima *cmf*. We observed that the SYN:FAP ratio, predicted using this criteria, is closer to the 1.5 to 3.5 value observed in field measurements. Thermodynamic considerations underline relative strain growth rates, or microbial diversities, that are more favorable from an energetic viewpoint, which indicates that an ecosystem behaves according to two different objectives: maximal biomass production and maximization of *cmf*, corroborating previous systems biology studies that advocate the use of distinct concurrent objectives to predict *Escherichia coli* metabolic behaviors [68]. In both cases, observations were possible by general investigation of the Pareto front.

Nevertheless, further refinement of the thermodynamic calculations is warranted. In particular, the calculation of *cmf* does not consider biomass concentration; this may be overcome by considering community fractions as proposed in [27] and [28]. Furthermore, in the current model, biomass generation does not affect the overall ecosystem entropy; however, on an intuitive basis, a larger amount of biomass should increase an entropy term, in terms of Gibbs energy, as a result of mass dispersion [69], thereby affecting *cmf* evaluation. These considerations are out of the scope of the present work; however, they but raise interesting perspectives.

Despite the above limitations, we consider the present form of the modeling approach as fruitful guidance to gain qualitative as well as quantitative data for the metabolic interplay between various species in an ecosystem. This method paves the way for improved contextualization of other -omics datasets in microbial ecology by providing a mechanistic description of species co-occurrence *via* analysis of their metabolic interactions.

Supporting information

S1 File. Guidelines for interpreting MO-FBA results.

(PDF)

S2 File. Metabolic Model of Hot Spring Community. A Stoichiometric Matrix of each guild used, along with thermodynamic data considered.

(XLSX)

S1 Video. Animated 3D version of Pareto front.

(MP4)

S1 Fig. Alternative MO-FVA illustration of Fig 5. The convention used is the same for S2–S5 Figs.

(EPS)

S2 Fig. MO-FVA for NH₃ exchange fluxes between SYN, FAP, and SRB.

(PNG)

S3 Fig. MO-FVA for acetate exchange fluxes between SYN, FAP, and SRB.

(PNG)

S4 Fig. MO-FVA for CO₂ exchange fluxes between SYN, FAP, and SRB.

(PNG)

S5 Fig. MO-FVA for glycolate exchange fluxes between SYN, FAP, and SRB.

(PNG)

S6 Fig. OptCom and Descriptive OptCom results mapped in the Pareto front.
(PNG)

Acknowledgments

MB is supported by CNRS and Region Pays de la Loire funding (GRIOTE project). This study was supported by ANR (IMPEKAB, ANR-15-CE02-001-03). We would like to thank Editage (www.editage.com) for English language editing. The authors would like also to thank the anonymous reviewer for his valuable input and comments regarding the manuscript.

Author Contributions

Conceptualization: DE JB MB.

Data curation: MB.

Formal analysis: MB.

Funding acquisition: DE JB.

Software: MB.

Supervision: DE JB AL.

Visualization: MB.

Writing – original draft: DE JB AL MB.

Writing – review & editing: DE JB AL MB.

References

1. Whitman WB, Coleman DC, Wiebe WJ. Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences of the United States of America*. 1998; 95(12):6578–6583. doi: [10.1073/pnas.95.12.6578](https://doi.org/10.1073/pnas.95.12.6578) PMID: [9618454](https://pubmed.ncbi.nlm.nih.gov/9618454/)
2. Kallmeyer J, Pockalny R, Adhikari RR, Smith DC, D'Hondt S. Global distribution of microbial abundance and biomass in subseafloor sediment. *Proceedings of the National Academy of Sciences*. 2012; 109(40):16213–16216. doi: [10.1073/pnas.1203849109](https://doi.org/10.1073/pnas.1203849109)
3. Lin BL, Sakoda A, Shibasaki R, Goto N, Suzuki M. Modelling a global biogeochemical nitrogen cycle in terrestrial ecosystems. *Ecological Modelling*. 2000; 135(1):89–110. doi: [10.1016/S0304-3800\(00\)00372-0](https://doi.org/10.1016/S0304-3800(00)00372-0)
4. Rullkötter J. Organic Matter: The Driving Force for Early Diagenesis. In: *Marine Geochemistry*. Berlin/Heidelberg: Springer Berlin Heidelberg; 2006. p. 125–168.
5. Jessup CM, Kassen R, Forde SE, Kerr B, Buckling A, Rainey PB, et al. Big questions, small worlds: microbial model systems in ecology. *Trends in Ecology & Evolution*. 2004; 19(4):189–197. doi: [10.1016/j.tree.2004.01.008](https://doi.org/10.1016/j.tree.2004.01.008) PMID: [16701253](https://pubmed.ncbi.nlm.nih.gov/16701253/)
6. McGill BJ, Enquist BJ, Weiher E, Westoby M. Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*. 2006; 21(4):178–185. doi: [10.1016/j.tree.2006.02.002](https://doi.org/10.1016/j.tree.2006.02.002) PMID: [16701083](https://pubmed.ncbi.nlm.nih.gov/16701083/)
7. Krause S, Le Roux X, Niklaus PA, Van Bodegom PM, Lennon JT, Bertilsson S, et al. Trait-based approaches for understanding microbial biodiversity and ecosystem functioning. *Frontiers in Microbiology*. 2014; 5(364):251. doi: [10.3389/fmicb.2014.00251](https://doi.org/10.3389/fmicb.2014.00251) PMID: [24904563](https://pubmed.ncbi.nlm.nih.gov/24904563/)
8. Litchman E, Klausmeier CA. Trait-Based Community Ecology of Phytoplankton. *Annual Review of Ecology, Evolution, and Systematics*. 2008; 39(1):615–639. doi: [10.1146/annurev.ecolsys.39.110707.173549](https://doi.org/10.1146/annurev.ecolsys.39.110707.173549)
9. Segata N, Boernigen D, Tickle TL, Morgan XC, Garrett WS, Huttenhower C. Computational meta-omics for microbial community studies. *Molecular Systems Biology*. 2013; 9(666):1–15 doi: [10.1038/msb.2013.22](https://doi.org/10.1038/msb.2013.22) PMID: [23670539](https://pubmed.ncbi.nlm.nih.gov/23670539/)
10. Waldor MK, Tyson G, Borenstein E, Ochman H, Moeller A, Finlay BB, et al. Where Next for Microbiome Research? *PLoS Biology*. 2015; 13(1):e1002050. doi: [10.1371/journal.pbio.1002050](https://doi.org/10.1371/journal.pbio.1002050) PMID: [25602283](https://pubmed.ncbi.nlm.nih.gov/25602283/)

11. Raes J, Bork P. Molecular eco-systems biology: towards an understanding of community function. *Nature Reviews. Microbiology*. 2008; 6(9):693–699. PMID: [18587409](#)
12. Fuhrman JA, Cram JA, Needham DM. Marine microbial community dynamics and their ecological interpretation. *Nature Reviews Microbiology*. 2015; 13(3):133–146. doi: [10.1038/nrmicro3417](#) PMID: [25659323](#)
13. Lewis NE, Nagarajan H, Palsson BO. Constraining the metabolic genotype–phenotype relationship using a phylogeny of in silico methods. *Nature Reviews: Microbiology*. 2012; 10(4):291–305. doi: [10.1038/nrmicro2737](#) PMID: [22367118](#)
14. Bordbar A, Monk JM, King ZA, Palsson BO. Constraint-based models predict metabolic and associated cellular functions. *Nature Reviews. Genetics*. 2014; 15(2):107–120. PMID: [24430943](#)
15. Klitgord N, Segrè D. Ecosystems biology of microbial metabolism. *Current opinion in biotechnology*. 2011; 22(4):541–546. doi: [10.1016/j.copbio.2011.04.018](#) PMID: [21592777](#)
16. Zengler K, Palsson BO. A road map for the development of community systems (CoSy) biology. *Nature Reviews: Microbiology*. 2012; 10(5):366–372. PMID: [22450377](#)
17. Kim TY, Sohn SB, Bin Kim Y, Kim WJ, Lee SY. Recent advances in reconstruction and applications of genome-scale metabolic models. *Current Opinion in Biotechnology*. 2012; 23(4):617–623. doi: [10.1016/j.copbio.2011.10.007](#) PMID: [22054827](#)
18. Thiele I, Palsson BO. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protocols*. 2010; 5(1):93–121. doi: [10.1038/nprot.2009.203](#) PMID: [20057383](#)
19. Hanemaaijer M, Röling WFM, Olivier BG, Khandelwal RA, Teusink B, Bruggeman FJ. Systems modeling approaches for microbial community studies: from metagenomics to inference of the community structure. *Frontiers in Microbiology*. 2015; 6(213):1–12. doi: [10.3389/fmicb.2015.00213](#) PMID: [25852671](#)
20. Rodríguez J, Kleerebezem R, Lema JM, van Loosdrecht MCM. Modeling product formation in anaerobic mixed culture fermentations. *Biotechnology and Bioengineering*. (2006), 93(3), 592–606. doi: [10.1002/bit.20765](#)
21. Biggs MB, Medlock GL, Kolling GL, Papin JA. Metabolic network modeling of microbial communities. *WIREs Systems Biology and Medicine*. 2015; 7:317–334 doi: [10.1002/wsbm.1308](#) PMID: [26109480](#)
22. Perez-Garcia O, Lear G, Singhal N. Metabolic Network Modeling of Microbial Interactions in Natural and Engineered Environmental Systems. *Frontiers in Microbiology*. 2016; 7(673), 1–30. doi: [10.3389/fmicb.2016.00673](#) PMID: [27242701](#)
23. Klitgord N, Segrè D The importance of compartmentalization in metabolic flux models: yeast as an ecosystem of organelles. *Genome Informatics* (2009), 22, 41–55. PMID: [20238418](#)
24. Stolyar S, Van Dien S, Hillesland KL, Pinel N, Lie TJ, Leigh JA, Stahl DA. Metabolic modeling of a mutualistic microbial community. *Molecular Systems Biology* (2007), 3, 92:1–14.
25. Zomorodi AR, Maranas CD. OptCom: a multi-level optimization framework for the metabolic modeling and analysis of microbial communities. *PLoS Computational Biology*. 2012; 8(2):e1002363. doi: [10.1371/journal.pcbi.1002363](#) PMID: [22319433](#)
26. Zomorodi AR, Islam MM, Maranas CD. d-OptCom: Dynamic Multi-level and Multi-objective Metabolic Modeling of Microbial Communities. *ACS Synthetic Biology*. 2014; 3(4):247–257. doi: [10.1021/sb4001307](#) PMID: [24742179](#)
27. Khandelwal RA, Olivier BG, Röling WF, Teusink B, Bruggeman FJ. Community Flux Balance Analysis for Microbial Consortia at Balanced Growth. *PLoS ONE*. 2013; 8(5):e64567. doi: [10.1371/journal.pone.0064567](#) PMID: [23741341](#)
28. Koch S, Benndorf D, Fronk K, Reichl U, Klamt S. Predicting compositions of microbial communities from stoichiometric models with applications for the biogas process. *Biotechnology for Biofuels*. 2016; 9(1):1–16. doi: [10.1186/s13068-016-0429-x](#) PMID: [26807149](#)
29. Ehrgott M. *Multicriteria Optimization*. Berlin, Germany: Springer Science & Business Media; 2005.
30. Vo TD, Greenberg HJ, Palsson BO. Reconstruction and functional characterization of the human mitochondrial metabolic network based on proteomic and biochemical data. *Journal of Biological Chemistry*. 2004; 279(38):39532–39540. doi: [10.1074/jbc.M403782200](#) PMID: [15205464](#)
31. Kschischo M. A gentle introduction to the thermodynamics of biochemical stoichiometric networks in steady state. *The European Physical Journal Special Topics*. 2010; 187(1):255–274. doi: [10.1140/epjst/e2010-01290-3](#)
32. Dillon JG, Fishbain S, Miller SR, Bebout BM, Habicht KS, Webb SM, et al. (2007). High rates of sulfate reduction in a low-sulfate hot spring microbial mat are driven by a low level of diversity of sulfate-respiring microorganisms. *Applied and Environmental Microbiology*. 2007; 73(16), 5218–5226. doi: [10.1128/AEM.00357-07](#) PMID: [17575000](#)

33. Taffs R, Aston JE, Brileya K, Jay Z, Klatt CG, McGlynn S, et al. In silico approaches to study mass and energy flows in microbial consortia: a syntrophic case study. *BMC Systems Biology*. 2009; 3(1):114. doi: [10.1186/1752-0509-3-114](https://doi.org/10.1186/1752-0509-3-114) PMID: [20003240](https://pubmed.ncbi.nlm.nih.gov/20003240/)
34. Beard DA, Liang Sd, Qian H. Energy balance for analysis of complex metabolic networks. *Biophysical Journal*. 2002; 83(1):79–86. doi: [10.1016/S0006-3495\(02\)75150-3](https://doi.org/10.1016/S0006-3495(02)75150-3) PMID: [12080101](https://pubmed.ncbi.nlm.nih.gov/12080101/)
35. Qian H, Beard DA, Liang Sd. Stoichiometric network theory for nonequilibrium biochemical systems. *European Journal of Biochemistry / FEBS*. 2003; 270(3):415–421. doi: [10.1046/j.1432-1033.2003.03357.x](https://doi.org/10.1046/j.1432-1033.2003.03357.x) PMID: [12542691](https://pubmed.ncbi.nlm.nih.gov/12542691/)
36. Alberty RA. Appendix 2: Tables of Transformed Thermodynamic Properties. *Applications of Mathematica*. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2006.
37. Flamholz A, Noor E, Bar-Even A, Milo R. eQuilibrator—the biochemical thermodynamics calculator. *Nucleic Acids Research*. 2012; 40(Database issue):D770–D775. doi: [10.1093/nar/gkr874](https://doi.org/10.1093/nar/gkr874) PMID: [22064852](https://pubmed.ncbi.nlm.nih.gov/22064852/)
38. Hoppe A, Hoffmann S, Holzhütter HG. Including metabolite concentrations into flux balance analysis: thermodynamic realizability as a constraint on flux distributions in metabolic networks. *BMC Systems Biology*. 2007; 1(1):1–23 doi: [10.1186/1752-0509-1-23](https://doi.org/10.1186/1752-0509-1-23)
39. Henry CS, Broadbelt LJ, Hatzimanikatis V. Thermodynamics-Based Metabolic Flux Analysis. *Biophysical Journal*. 2007; 92(5):1792–1805. doi: [10.1529/biophysj.106.093138](https://doi.org/10.1529/biophysj.106.093138) PMID: [17172310](https://pubmed.ncbi.nlm.nih.gov/17172310/)
40. Fleming RMT, Thiele I, Provan G, Nasheuer HP. Integrated stoichiometric, thermodynamic and kinetic modelling of steady state metabolism. *Journal of Theoretical Biology*. 2010; 264(3):683–692. doi: [10.1016/j.jtbi.2010.02.044](https://doi.org/10.1016/j.jtbi.2010.02.044) PMID: [20230840](https://pubmed.ncbi.nlm.nih.gov/20230840/)
41. Dantzig GB. Reminiscences About the Origins of Linear Programming. In: *Mathematical Programming The State of the Art*. Berlin, Germany: Springer; 1983. p. 78–86. Available from:
42. Gurobi Optimization I. *Gurobi Optimizer Reference Manual*; 2015. Available from: <http://www.gurobi.com>
43. Ehrgott M, Wiecek MM. Multiobjective Programming. In: *Multiple Criteria Decision Analysis: State of the Art Surveys*. New York: Springer-Verlag; 2005. p. 667–708.
44. Aoki I. Entropy and exergy in the development of living systems: a case study of lake-ecosystems. *Journal of the Physical Society of Japan*. 1998; 67(6):2132–2139. doi: [10.1143/JPSJ.67.2132](https://doi.org/10.1143/JPSJ.67.2132)
45. Martyushev LM, Seleznev VD. Maximum entropy production principle in physics, chemistry and biology. *Physics Reports*. 2006; 426(1):1–45. doi: [10.1016/j.physrep.2005.12.001](https://doi.org/10.1016/j.physrep.2005.12.001)
46. Stadler W. A survey of multicriteria optimization or the vector maximum problem, part I: 1776–1960. *Journal of Optimization Theory and Applications*. 1979; 29(1):1–52. doi: [10.1007/BF00932634](https://doi.org/10.1007/BF00932634)
47. Marler RT, Arora JS. Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization*. 2004; 26(6):369–395. doi: [10.1007/s00158-003-0368-6](https://doi.org/10.1007/s00158-003-0368-6)
48. Benson HP. An Outer Approximation Algorithm for Generating All Efficient Extreme Points in the Outcome Set of a Multiple Objective Linear Programming Problem. *Journal of Global Optimization*. 1998; 13(1):1–24.
49. Ehrgott M, Shao L, Schöbel A. An approximation algorithm for convex multi-objective programming problems. *Journal of Global Optimization*. 2010; 50(3):397–416. doi: [10.1007/s10898-010-9588-7](https://doi.org/10.1007/s10898-010-9588-7)
50. Ehrgott M, Löhne A, Shao L. A dual variant of Benson’s “outer approximation algorithm” for multiple objective linear programming. *Journal of Global Optimization*. 2012; 52(4):757–778. doi: [10.1007/s10898-011-9709-y](https://doi.org/10.1007/s10898-011-9709-y)
51. Hamel AH, Löhne A, Rudloff B. Benson type algorithms for linear vector optimization and applications. *Journal of Global Optimization*. 2013; 59(4):811–836. doi: [10.1007/s10898-013-0098-2](https://doi.org/10.1007/s10898-013-0098-2)
52. Löhne A, Weißing B. BENSOLVE—VLP Solver, version 2.0.2; 2015. Available from: <http://www.bensolve.org>
53. Anderson KL, Tayne TA, Ward DM. Formation and Fate of Fermentation Products in Hot Spring Cyanobacterial Mats. *Applied and Environmental Microbiology*. 1987; 53(10):2343–2352. PMID: [16347455](https://pubmed.ncbi.nlm.nih.gov/16347455/)
54. Caspi R, Altman T, Billington R, Dreher K, Foerster H, Fulcher CA, et al. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Research*. 2014; 42(Database issue):D459–471. doi: [10.1093/nar/gkt1103](https://doi.org/10.1093/nar/gkt1103) PMID: [24225315](https://pubmed.ncbi.nlm.nih.gov/24225315/)
55. Steunou AS, Bhaya D, Bateson MM, Melendrez MC, Ward DM, Brecht E, et al. In situ analysis of nitrogen fixation and metabolic switching in unicellular thermophilic cyanobacteria inhabiting hot spring microbial mats. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103(7):2398–2403. doi: [10.1073/pnas.0507513103](https://doi.org/10.1073/pnas.0507513103) PMID: [16467157](https://pubmed.ncbi.nlm.nih.gov/16467157/)

56. Steunou AS, Jensen SI, Brecht E, Becraft ED, Bateson MM, Kilian O, et al. Regulation of nif gene expression and the energetics of N₂ fixation over the diel cycle in a hot spring microbial mat. *The ISME journal*. 2008; 2(4):364–378. doi: [10.1038/ismej.2007.117](https://doi.org/10.1038/ismej.2007.117) PMID: [18323780](https://pubmed.ncbi.nlm.nih.gov/18323780/)
57. Oberhardt MA, Chavali AK, Papin JA. Flux Balance Analysis: Interrogating Genome-Scale Metabolic Networks. In: *Systems Biology*. Totowa, NJ: Humana Press; 2009. p. 61–80.
58. Kim YM, Nowack S, Olsen MT, Becraft ED, Wood JM, Thiel V, et al. Diel metabolomics analysis of a hot spring chlorophototrophic microbial mat leads to new hypotheses of community member metabolisms. *Frontiers in Microbiology*. 2015; 6(209):1–14. doi: [10.3389/fmicb.2015.00209](https://doi.org/10.3389/fmicb.2015.00209) PMID: [25941514](https://pubmed.ncbi.nlm.nih.gov/25941514/)
59. Nash SG. A survey of truncated-Newton methods. *Journal of Computational and Applied Mathematics*. 2000; 124(1-2):45–59. doi: [10.1016/S0377-0427\(00\)00426-X](https://doi.org/10.1016/S0377-0427(00)00426-X)
60. Hunter JD. Matplotlib: A 2D graphics environment. *Computing In Science & Engineering*. 2007; 9(3):90–95. doi: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55)
61. Tawarmalani M, Sahinidis NV. A polyhedral branch-and-cut approach to global optimization. *Mathematical Programming*, 2005; 103(2), 225–249 doi: [10.1007/s10107-005-0581-8](https://doi.org/10.1007/s10107-005-0581-8)
62. Czyzyk J, Mesnier MP, Moré JJ. The NEOS Server. *IEEE Journal on Computational Science and Engineering*. 1998; 5(3), 68–75. doi: [10.1109/99.714603](https://doi.org/10.1109/99.714603)
63. Dolan E. The NEOS Server 4.0 Administrative Guide. Technical Memorandum ANL/MCS-TM-250, Mathematics and Computer Science Division, Argonne National Laboratory. 2001.
64. Gropp W, Moré JJ. Optimization Environments and the NEOS Server. In: *Approximation Theory and Optimization*, Buhmann MD and Iserles A, eds., Cambridge University Press; 1997, p167–182.
65. Klatt CG, Liu Z, Ludwig M, Kuhl M, Jensen SI, Bryant DA, et al. Temporal metatranscriptomic patterning in phototrophic *Chloroflexi* inhabiting a microbial mat in a geothermal spring. *The ISME Journal*. 2013; 7(9):1775–1789. doi: [10.1038/ismej.2013.52](https://doi.org/10.1038/ismej.2013.52) PMID: [23575369](https://pubmed.ncbi.nlm.nih.gov/23575369/)
66. Borenstein E, Kupiec M, Feldman MW, Ruppin E. Large-scale reconstruction and phylogenetic analysis of metabolic environments. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105(38):14482–14487. doi: [10.1073/pnas.0806162105](https://doi.org/10.1073/pnas.0806162105) PMID: [18787117](https://pubmed.ncbi.nlm.nih.gov/18787117/)
67. Bordron P, Latorre M, Cortés MP, González M, Thiele S, Siegel A, et al. Putative bacterial interactions from metagenomic knowledge with an integrative systems ecology approach. *MicrobiologyOpen*. 2016; 5(1):106–117. doi: [10.1002/mbo3.315](https://doi.org/10.1002/mbo3.315) PMID: [26677108](https://pubmed.ncbi.nlm.nih.gov/26677108/)
68. Schuetz R, Zamboni N, Zampieri M, Heinemann M, Sauer U. Multidimensional Optimality of Microbial Metabolism. *Science (New York, NY)*. 2012; 336(6081):601–604. doi: [10.1126/science.1216882](https://doi.org/10.1126/science.1216882)
69. England JL. Statistical physics of self-replication. *The Journal of Chemical Physics*. 2013; 139(12):121923. doi: [10.1063/1.4818538](https://doi.org/10.1063/1.4818538) PMID: [24089735](https://pubmed.ncbi.nlm.nih.gov/24089735/)