

A microbial community growth model for dynamic phenotype predictions

Andrew P. Freiburger¹, Jeffrey A. Dewey², Fatima Foflonker¹, Gyorgy Babnigg², Dionysios A. Antonopoulos², and Christopher Henry^{1*}

¹Argonne National Laboratory, Computing, Environment, and Life Sciences Division, Lemont, IL 60148 USA

²Argonne National Laboratory, Biosciences Division, Lemont, IL 60148 USA
e-mail: cshenry@anl.gov

December 16, 2022

1 Abstract

Microbial communities are increasingly recognized as key drivers in animal health, agricultural productivity, industrial operations, and ecological systems. The abundance of chemical interactions in these complex communities, however, can complicate or evade experimental studies, which hinders basic understanding and limits efforts to rationally design communities for applications in the aforementioned fields. Numerous computational approaches have been proposed to deduce these metabolic interactions – notably including flux balance analysis (FBA) and systems of ordinary differential equations (ODEs) – yet, these methods either fail to capture the dynamic phenotype expression of community members or lack the abstractions required to fit or explain the diverse experimental omics data that can be acquired today.

We therefore developed a dynamic model (CommPhitting) that deduces phenotype abundances and growth kinetics for each community member, concurrent with metabolic concentrations, by coupling flux profiles for each phenotype with experimental growth and -omics data of the community. These data are captured as variables and coefficients within a mixed integer linear optimization problem (MILP) designed to represent the associated biological processes. This problem finds the globally optimized fit to all experimental data of a trial, thereby most accurately computing aspects of the community: (1) species and phenotype abundances over time; (2) a linearized growth kinetic constant for each phenotype; and (3) metabolite concentrations over time. We exemplify CommPhitting by applying it to study batch growth of an idealized two-member community of the model organisms (*Escherichia coli* and *Pseudomonas fluorescens*) that exhibits cross-feeding in maltose media. Measurements of this community from our accompanying experimental studies – including total biomass, species biomass, and metabolite abundances over time – were parameterized into a CommPhitting simulation. The resultant kinetics constants and biomass proportions for each member phenotype would be difficult to ascertain experimentally, yet are important for understanding community responses to environmental perturbations and therefore engineering applications: e.g. for bioproduction. We believe that CommPhitting – which is generalized for a diversity of data types and formats, and is further available and amply documented as a Python API – will augment basic understanding of microbial communities and will accelerate the engineering of synthetic communities for diverse applications in medicine, agriculture, industry, and ecology.

2 Introduction

Microbial communities are ubiquitous on Earth [1], and serve fundamental roles as eukaryotic symbionts [2, 3, 4] and pathogens [5, 6], industrial assets [7] and antagonists [8], and ecological agents of biogeochemical cycling [9, 10, 11]. Microbial communities are therefore essential to understand in fields as diverse as medicine [12, 13] and climatology [14, 15, 16]. The assemblage of microbes into a community, while inviting intimate competition [17], offers a few advantages that make it biologically favorable in most conditions. Firstly, diversification in a) genetic potential, b) biochemical vulnerabilities [18], and c) metabolic machinery [19, 20, 21] allows community members to grow in otherwise inhospitable conditions, such as nutrient-poor or toxic environments, by relying on well-adapted members. Secondly, specialization among the members enables the community to achieve more efficient production of vital nutrients [22] and greater growth rates [23] than is possible by isolated, autonomous, members. These interactions that guide community assemblage are predicated upon a metabolic economy, where the competitive community environment diversifies members

into ecological niches that balance nutritional satiation and metabolic efficiency by leveraging cross-feeding [24, 25]. Metabolic cross-feeding (syntropy) is essential to understand and engineer communities for desirable properties [26], such as bioproduction that exceeds the capabilities of monocultures [7] or sustainable biotechnology by coupling phototrophs and heterotrophs [27], yet it is difficult to experimentally measure.

One impediment to measuring syntropy is that some metabolites do not detectably accumulate in the media, which obviates all but a few experimental methods, such as fluxomics isotope labeling [28]. A second difficulty, which proves to be more formidable, is that the combinatorial quantity of cross-feeding exchanges creates an untenable multitude of experiments to resolve all possible interactions in natural communities. These challenges are further compounded by the numerous metabolic phenotypes of each member that dynamically respond, even in clonal monocultures [29], to environmental conditions. Phenotype variability is necessary to understand community dynamics, and particularly to engineer members for production or exclusion, but this analysis adds another experimental dimension to an already challenging and expensive series of measurements.

Computational biology offers a few tools to resolve the metabolic nuances of microbial communities. Flux Balance Analysis (FBA) [30], for example, is a prominent framework for simulating cellular metabolism and has been applied in numerous software tools that seek to resolve metabolic interactions within microbial communities: e.g. BacArena [31], μ bialSim [32], dOptCom [33], SMETANA [34], CASINO [2], and MICOM [35], to name a few. FBA is attractive as a simple method for simulating metabolism that can operate without or, optionally, with experimental -omics data types [36] such as metabolomics [37, 38], proteomics [39], transcriptomics [40, 41, 42], or multi-omics [43] data. The under-determined linear problem of FBA, however, often lessens the precision and consistency of predictions and FBA moreover cannot natively capture phenotype variability within a given species metabolic model. Machine-learning algorithms [44, 45], in contrast, offer more exploratory flexibility for phenotype abundances from experimental data, but often lack the mechanisms that cultivate basic understanding and design principles. Differential equation models offer precise mechanistic explanations for predictions but has the converse problem [46] of frequently being stiff and insufficiently unbiased to flexibly explore unparameterized spaces. Fitting models, by contrast, as a final popular tool of computational biology, can be flexible enough to acquiesce new information from experimental data [47, 48, 49] while maintaining sufficient mechanistic resolution for basic understanding. Fitting methods have been accordingly applied to deduce kinetics coefficients [50] or phenotype expression [40] from experimental -omics data, but these methods do not convey time-resolved phenotype abundances within context of other experimental dimensions, such as metabolic concentration, that cultivate comprehensive understanding.

We therefore developed a global fitting model (CommPhitting) that deduces time-resolved information from experimental data: such as a) biomass of each community member and member phenotype; b) total concentrations of all metabolites that are used by the phenotypes; c) growth kinetics for each member phenotype; and d) conversion factors from each experimental signal to biomass that may accelerate the interpretation of future experiments. CommPhitting is a multi-species dynamic model of the community member phenotypes, where the phenotype flux profiles are rigorously defined through a sequence of FBA simulations – 1) a minimal growth is constrained (to prevent singularities); 2) carbonaceous non-phenotype consumption is minimized (to isolate the phenotype); 3) consumption of the phenotype source is minimized (to maximize the biomass yield); 4) optionally, specified phenotype excretions are maximized (to mirror experimental phenotypes); and 5) total flux is minimized (to find the most parsimonious flux profile) – that extends previously defined methods [40]. CommPhitting is generalized for various data types into a linear problem [51] that solves simultaneously to ensure that a global optimum is determined [52], and is concisely coded to execute within a minute on a personal computer.

We exemplify CommPhitting with an idealized 2-member community of model organisms (*Escherichia coli* and *Pseudomonas fluorescens*), who exhibit complementary carbon preferences [53] but have not been investigated as a coculture. This benchtop community [54] demonstrated pivotal cross-feeding with acetate excreta from *E. coli* that was able to cultivate growth of *P. fluorescens* in a non-viable media of maltose (Figure 1). CommPhitting fit the experimental growth data remarkably well, and the predictions of phenotype abundances and metabolic concentrations were validated with transcriptomics data from the NCBI GEO database [55] and in-house metabolomics measurements of our community, respectively. We further simulated BIOLOG data from this community, which revealed growth kinetics of this community for each condition and may have broad utility for predicting community behaviors in a variety of conditions. We anticipate using the knowledge from these simulations to guide the engineering of *E. coli* for community bioproduction. CommPhitting is available as an open-source Python module in the ModelSEEDpy and is operable with the KBase [56] ecosystem. We believe that this light-weight yet robust fitting model will illuminate nuances of community interactions and cultivate rational design of microbial communities for myriad fundamental and industrial applications.

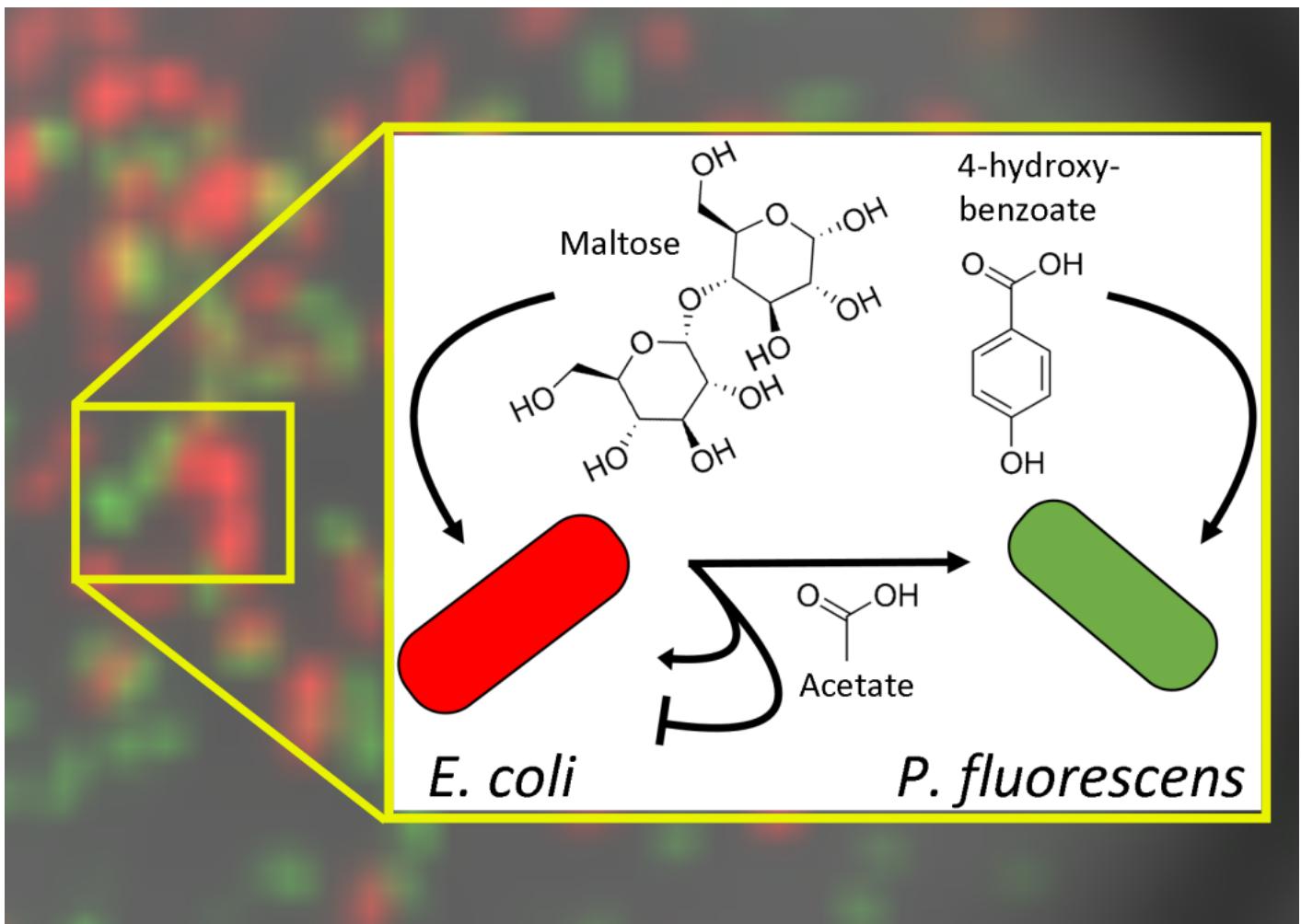


Figure 1: The primary metabolic exchanges that were experimentally elucidated and computationally modeled. The acetate byproduct of *E. coli* is the pivotal exchange of metabolic interest, where this source subsists *P. fluorescens*'s existence in maltose and exhibits interesting dual effects upon *E. coli* as both a secondary carbon source and a growth inhibitor. Additional experiments revealed that *E. coli* is reluctant to grow in pure acetate and most often fails to grow at all, particularly when in cocultures with *P. fluorescens*.

3 Methods

The CommPhitting method transforms the traditional ordinary differential equations that define microbial growth and chemical activity into a mixed integer linear optimization problem (MILP). The MILP variables, constraints, and parameters are described in Table 1, and the MILP objective minimizes error in predicted growth, versus experimental growth, and minimizes overfitting. The processes and calculations that constitute the CommPhitting MILP are detailed in the following sections.

3.1 Metabolic phenotypes

The first step in constructing the CommPhitting MILP defines the metabolic phenotypes for each community member strain, notwithstanding the non-growing stationary phenotype. Experimental genomes of each member are reconstructed into genome-scale metabolic models (GEMs), where the ModelSEED pipeline [57] is advantageous among reconstruction tools for guaranteeing model consistency and compatibility (e.g. matching IDs for extracellular metabolites). The GEMs are then manipulated into specified phenotypes through the following sequence of FBA simulations.

1. The biological and environmental conditions of the GEM are constrained. A minimal biomass growth is defined; hydrogen consumption is prohibited, which forces consumption of the specific carbon source for growth; and oxygen consumption is limited to the total consumption of the defining phenotype carbon source(s), although, this threshold can fluctuate to accommodate the specific species biology and to properly regulate overflow metabolism.
2. The total influx of carbon compounds is minimized, excluding the defined phenotype carbon source(s) and any compounds with unknown formula. This optimally focuses on the metabolism that corresponds with consuming the designated carbon source(s) of the phenotype. The resultant fluxes from this minimization are fixed to the model and are unchanged in subsequent optimizations.
3. The total flux of the phenotype carbon sources is minimized, while fixing the biomass growth flux, which maximizes biomass yield from the designated carbon source(s). This presumably identifies the most efficient and biologically desirable nutrient consumption and byproduct excretion profile for each phenotype. The resultant carbon source fluxes from this minimization are fixed to the model.
4. Optionally, the excretion flux for assigned excreta of a phenotype (where experimental evidence is available) is maximized, and then the excreta flux is fixed to the model.
5. The final step applies parsimonious FBA (pFBA). This FBA method minimizes the sum of all fluxes at a fixed biomass growth and thereby finds the simplest flux solution to reach the prescribed growth. This pFBA simulation presumably emulates the efficiencies that evolution found for the simulated organism. The non-zero exchange fluxes from this final optimization becomes the CommPhitting representation of the specified phenotype.

The above five-step sequence is repeated for each specified phenotype of all species in the community. This robust method creates an irreducible metabolic profile for each phenotype that fosters pure phenotype predictions by CommPhitting.

3.2 Constraints

Each of the following linear constraints of CommPhitting captures a distinct aspect of community biology with the previously defined phenotypes.

3.2.1 Biomass abundances

The experimental biomass ($EB_{s,t}$), over all time points $t \in T$, is determined by converting the experimental fluorescence signal ($E_{s,t}$) via a unique coefficient (EC_s) for each member $s \in S$

$$E_{s,t} * EC_s = EB_{s,t} . \quad (1)$$

The final $E_{s,t}$ is limited to be within 30% of the estimated biomass to consume 90% of the carbon source(s) according to the phenotype metabolic profile, which firmly anchors the EC_s conversion to data. The variance ($EV_{s,t}$) between the converted experimental biomass and the predicted biomass $b_{k,t}$ is determined for each phenotype $k \in K$

$$EB_{s,t} - \sum_{s,k}^{S,K} (es_s * b_{k,t}) = EV_{s,t} , \quad (2)$$

Table 1: A glossary of dimensions, parameters, and variables that comprise the fitting model.

Term	Type	Description
DIMENSIONS		
s	String	A species in the examined community.
k	String	A growth phenotype of species s .
t	Float	An experimental time point.
i	String	Extracellular metabolite.
z	Integer	A biomass partition.
PARAMETERS		
$E_{s,t}$	Float	The experimental growth signal for a species at instant t .
$es_{s,k}$	Boolean	A designation of truth for $k \in s$
Δt	Float	The seconds per timestep, which determines the amount of biomass growth per timestep.
$n_{k,i}$	Float	The exchange flux of each metabolite i in each strain k .
$cvcf$ & $cvcf$	Float	Conversion coefficients of phenotype biomass to and from the stationary phase, respectively.
bcv_k	Float	The greatest fraction of biomass ($0 < bcv < 1$) of strain k that can transition phenotypes in a timestep.
$cvmn$	Float	The minimal value of variable $cvt_{k,t}$.
$stat$	Float	The optimization penalty for the stationary phenotype of each species.
$kcat_z$	Float	The $kcat$ growth rate constant for the biomass partition z .
$kcat_k$	Float	The $kcat$ growth rate constant for the phenotype k .
VARIABLES		
EC_k	Continuous	The conversion coefficient ($0 < EC < 1000$) from parameter $E_{s,t}$ into biomass, which is unique for each strain k .
$EB_{s,t}$	Continuous	The computed biomass from each experimental datum, as the product of EC_k & $E_{s,t}$.
bin_k^z	Binary	A binary switch that determines whether a given biomass partition of a phenotype is active and contributes to the total $kcat$ of the phenotype.
$b_{k,t}^z$	Continuous	The phenotype biomass partitions, which each exhibit a distinct $kcat$.
$b_{k,t}$	Continuous	The total predicted phenotype biomass from the fitting model.
$EV_{s,t}$	Continuous	The variance between the computed experimental biomass $EB_{s,t}$ and the predicted biomass $b_{k,t}$.
$g_{k,t}$	Continuous	The predicted growth rate for each strain at each datum.
$c_{t,i}$	Continuous	The concentration of metabolite i at an experimental datum.
$cvt_{k,t}$ & $cvf_{k,t}$	Continuous	The quantity of strain k biomass that transitions to and from the stationary phase, respectively, at an experimental datum.

where only the phenotypes of each respective species are considered, per the binary es_s variable. The predicted biomass $b_{k,t}$ is further partitioned into five sub-groups that are weighted along a gradient

$$b_{k,t}^z \leq kcat_z * b_{k,t} \quad \forall z \in Z \quad (3)$$

and that each exhibit a distinct $kcat_z$ value. The expression of these biomass partitions is coordinated by the respective binary variable bin_k^z for each partition through the

$$b_{k,t}^z \leq 1000 - 1000 * bin_k^z \quad \forall z \in Z \quad (4)$$

and

$$kcat_z * b_{k,t} - 1000 * bin_k^z \leq 0 \quad \forall z \in Z, \quad (5)$$

where one of the biomass partitions must be expressed

$$\sum_{z,k}^{Z,K} (bin_k^z) \leq 4. \quad (6)$$

The $kcat_k$ value of each phenotype is estimated by adjusting the $kcat_z$ distributions and utilizing an exterior loop that changes the $kcat_z$ values for each $z \in Z$ partition (e.g. [100, 10, 1, 0.1, 0.01] in the first loop, [4, 2, 1, 0.5, 0.25] in the second loop, [1, 0.5, 0.25, 0.125, 0.0625] in the third loop),

$$kcat_{k,new} = \sum_z^Z (kcat_z * bin_k^z * kcat_{k,old}) \quad (7)$$

where $kcat_{k,old}$ is guessed for the first loop and the solved $kcat_{k,new}$ is substituted into the next loop as $kcat_{k,old}$ for further refinement. This approach approximates $kcat_k$ values for each phenotype while maintaining linearity of the growth constraint in eq. (9).

3.2.2 Phenotype transitions

The biomass change over time is calculated for non-stationary (growing) phenotypes

$$b_{t,k} + \frac{\Delta t}{2} (g_{t,k} + g_{t+1,k}) + cvf_{t,k} - cvt_{t,k} = b_{t+1,k}. \quad (8)$$

The constraint consists of terms for the current biomass ($b_{k,t}$), the biomass growth rate $g_{k,t}$ as the biomass derivative ($\frac{\Delta biomass}{\Delta t}$), the biomass in the next timestep ($b_{k,t+1}$), and the net transition of biomass to non-stationary metabolic phenotypes $cvf_{k,t} - cvt_{k,t}$. The growth rate is constrained

$$kcat_k * \sum_z^Z (b_{k,t}^z) = g_{k,t}, \quad (9)$$

as the product of the current biomass abundance $b_{k,t}$ and the growth rate constant $kcat_k$, which reflects 1st-order kinetics with respect to biomass abundance and zero-order kinetics with respect to substrate concentration. Every species is defined with a stationary phenotype (not growing) that mediates and subtly delays phenotype transitions to reflect the inherent lag in pathway expression changes. Further, the stationary phenotypes allow each species to cease growth once the defined carbon source(s) is exhausted. Future biomass for the stationary phenotype is analogous to eq. (8),

$$b_{k,t} - \sum_s^S (es_s * (cvf_{k,t} - cvt_{k,t})) = b_{k,t+1} \quad (10)$$

except that a) $g_{k,t} = 0$ by definition; b) the net transition to non-stationary phenotypes is negative; and c) the net transition to all of the many non-stationary phenotypes are summed. The fraction of biomass that can transition is constrained

$$cvt_{k,t} \leq bcv * b_{k,t} + cvmin \quad (11)$$

to greater than a minimum limit ($cvmmin$) and lesser than a fraction of biomass (bcv) above this minimum.

This constraint is an application of Heun's integration method [58, 59], which – for an arbitrary function y , its derivative y' , and a timestep Δt – is defined as

$$y_{t+1} = y_t + \frac{1}{2} * \Delta t * (y'_t + y'_{t+1}). \quad (12)$$

Heun's method is a 2nd-order Runge-Kutta formulation [60] that captures dynamic changes with high numerical accuracy [61] while maintaining a linear formulation that is amenable with our MILP.

3.2.3 Concentrations

Future concentrations of substrates ($c_{t+1,i}$) for all $i \in I$ substrates are constrained

$$c_{t,i} + \frac{\Delta t}{2} \sum_k^K (n_{i,k}(g_{t,k} + g_{t+1,k})) = c_{t+1,i} \quad (13)$$

with Heun's method from eq. (12). This implementation includes the current concentration ($c_{t,i}$) and concentration changes that are the product of Δt and the inner product of the metabolite flux ($n_{i,k}$) from the phenotype flux profile (section 3.1) and the biomass growth rate for a member.

3.3 Objective

The objective expression

$$\sum_{s,t}^{S,T} (EV_{s,t}^2 + stat * b_{s,k,t}) , k = stationary \quad (14)$$

defines the sum of variance (EV) from eq. (2) – between the predicted (b) and experimental (EB) biomass – and the sum total amount of biomass in the stationary phenotypes. This objective is minimized to optimally fit and recapitulate the experimental growth, and minimally overfit from abusing stationary phenotypes or excessively transitioning phenotypes, where the stationary phenotype mediates phenotype transitions in eq. (8).

3.4 E. coli-Psuedomonas competitive community

3.4.1 Experimental Methods

The exemplary *P. fluorescens* SBW25 and *E. coli* (minimally modified) MG1655 community was selected for several reasons. Firstly, these members are model organisms that encompass a wide range of disciplines, such as a) the human microbiome, b) bioproduction, c) synthetic biology, d) the rhizosphere, e) agriculture, and f) ecology. Secondly, despite these members being robustly studied individually, they have not been studied as a coculture, which offers new experimental information and mitigates biases and preconceptions in hypothesis development and results interpretation.

The coculture was created through the following protocol. The *P. fluorescens* and *E. coli* strains were purchased from ATC, were stored at -80°C before preparation for electrocompetency, and were transformed with a plasmid to constitutively express either mNeongreen or mRuby2 fluorescent proteins (GFP and RFP), respectively [62]. Transformed cells were freshly streaked on a LB agar plate with appropriate antibiotics from -80°C glycerol stocks, and were incubated overnight at 30°C . A single colony from the plate was picked, placed into liquid LB (Lennox) broth with the antibiotics, and shaken @ 250 RPM overnight at 30°C . The 2 mL overnight culture was pelleted (4000x g for 10 min), the supernatant was removed, and the cells were resuspended in 1 mL of M9 media that contains no carbon source. This washing sequence was repeated twice. A 20 μL aliquot of the washed cells was combined with 2 mL of M9 media that contained the appropriate carbon source for each strain – 10 mM D-maltose for *E. coli* and 6 mM 4-hydroxybenzoate for *P. fluorescens* – and was shaken @ 250 RPM for at least 16 hours. These overnight cultures were washed twice, following the same procedure as the overnight culture. These cells were finally analytically examined via optical density (OD 590 nm) and fluorescence using a plate reader (Hidex).

The aforementioned M9 cultured cells were mixed in fresh M9 media to achieve the desired initial cell ratio at OD 0.1 (590 nm) and carbon source concentration/ratios. A 200 μL aliquot of the cell mixture was then added to wells of a sterile, black wall & clear bottom, 96-well imaging plate (Costar). The 96-well plate was then added to the Hidex plate reader and was prewarmed to 30°C . The cells were shaken orbitally at 900 RPM in the plate reader for, at least, 24h while being measured for optical density (600 nm), red fluorescence (544 excitation, 590 emission), and green fluorescence (485 excitation, 535 emission) every 10 minutes.

We utilized two methods to accurately disentangle the composition of a liquid coculture: 12°C and 37°C temperature variability and fluorescence reporter with the green- and red-fluorescence proteins, for *P. fluorescens* and *E. coli* respectively. The precision and mutual validation of these methods is depicted through a combined liquid culture and agar plating experiment (Figure 2), which supports their application in our study as a resolving mechanism for the community members. In this qualitative experiment, 2 mL of M9 media containing either maltose, 4-hydroxybenzoate, or acetate are seeded with mono- or cocultures of *E. coli* and *P. fluorescens* to 0.1 OD. After 24h growth with shaking at 30°C , the liquid cultures are diluted into M9 media and streaked out on LB agar plates. The plates are placed at three different temperatures to enable selective growth of specific organisms. Specifically, *P. fluorescens* grows perceptively after

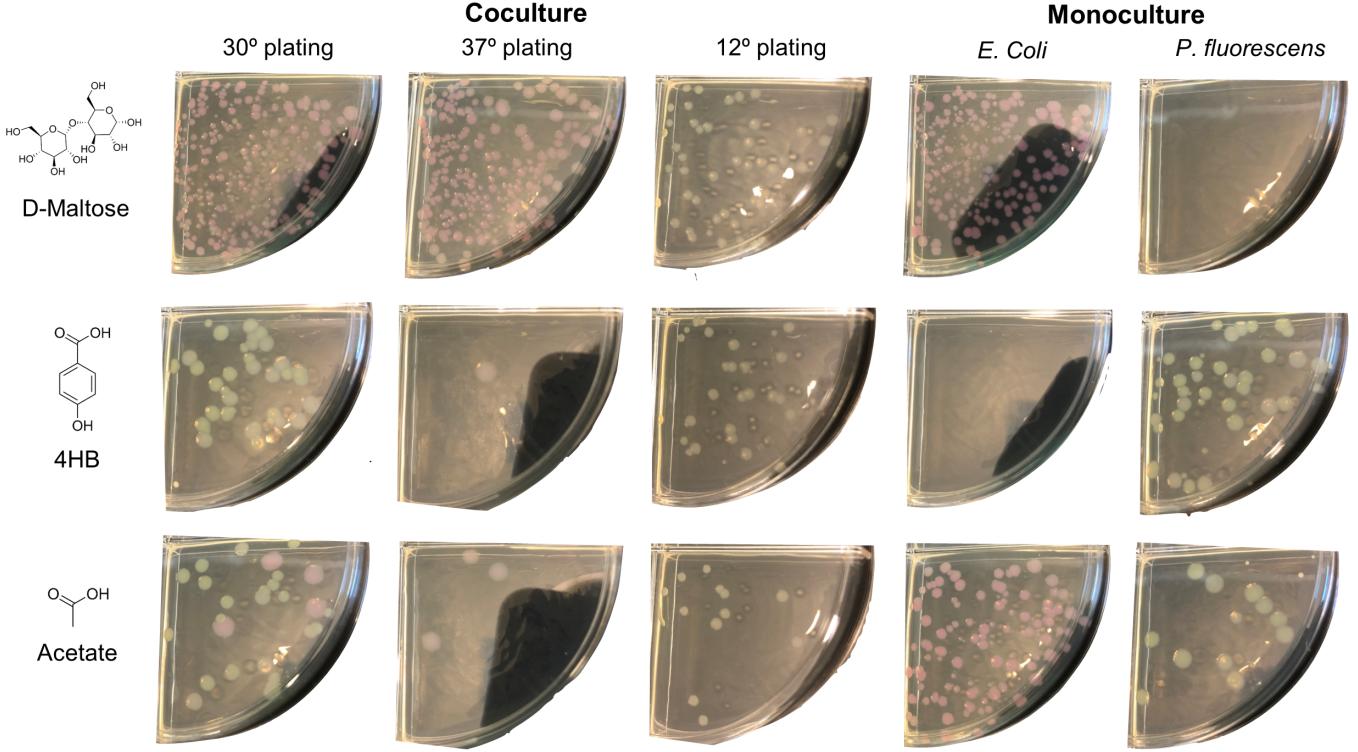


Figure 2: Organism abundance after co- and monoculture experiments in M9 media containing different carbon sources is visualized via plating at select temperatures. Coculture growth patterns replicate those observed in 96-well plate reader assays, confirming that fluorescence reporters accurately represent each organism’s growth/abundance.

60h at 12°C but does not grow at 37°C, *E. coli*. shows the opposite growth trend, and both organisms grow robustly overnight at 30°C. The results and significance of this experiment are discussed below, but the temperature dependent plating method confirms that fluorescence changes observed in coculture accurately represent changes in each organism’s growth/abundance.

3.4.2 Adapting the formulation for growth data

The general formulation in Sections 3.1-3.3 was tailored to this particular 2-member community. First, the signal to biomass conversion constraint in eq. (1) was adapted to this community system

$$\begin{aligned} RFP_t * RFPC &= RFPB_t \\ GFP_t * GFPC &= GFPB_t \\ OD_t * ODC &= ODB_t \end{aligned} \tag{15}$$

where the $RFPC$, $GFPC$, and ODC conversion factors were defined for each member and each RFP_t , GFP_t , and OD_t experimental signal. Second, the variance constraint in eq. (2) is adapted for each experimental signal

$$\begin{aligned} RFPB_t - \sum_k^K (pf_k * b_{t,k}) &= RFPV_t \\ GFPB_t - \sum_k^K (ec_k * b_{t,k}) &= GFPV_t \\ OD_t - \sum_k^K (b_{t,k}) &= ODV_t \end{aligned} \tag{16}$$

where p_{fk} and e_{ck} are binary variables that filter for *Pseudomonas* and *E. coli*, respectively. Third, the objective function of eq. (14) is defined for these members

$$\sum_t^T (EV_{ecoli,t}^2 + EV_{pseudo,t}^2 + EV_{OD,t}^2 + stat * (b_{ecoli,k,t} + b_{pseudo,k,t} + b_{OD,k,t})) , \quad k = stationary . \quad (17)$$

Finally, the data is filtered to remove timepoints beyond when the OD signal plateaus. This filtering mitigates the effects of fluorescent proteins over-production once cells enter their stationary phase and stop growing, by focusing on the lag and growth phases. These data could also be filtered by phenotype-specific conversions of the fluorescence signals into biomass abundance, instead of species-specific conversion. Anomalous timesteps or datums can also be removed ad hoc.

The *P. fluorescens* and *E. coli* genome-scale metabolic models were constructed in KBase Narrative 93465 by leveraging RAST to annotate the genomes [63] and then applying the ModelSEED pipeline [57]. The *E. coli* model derived from the ASM584v2 experimental genome assembly and was gapfilled with acetate and maltose carbon sources. The *P. fluorescens* model derived from the ASM161270v1 experimental genome assembly and was gapfilled with acetate and 4-hydroxybenzoate carbon sources. We attempted to use published metabolic models for *E. coli* [64] and *P. fluorescens* [65], however, these models were a) not sufficiently responsive to the phenotyping methods of 3.1 and b) were developed by different groups, which introduced incompatibilities when simulated together in a community.

4 Results and Discussion

4.1 New data fitting method for iterative experimental design

The CommPhitting method, as detailed in Section 3, was developed to resolve phenotype abundances in microbial communities while accommodating diverse experimental data. The logical workflow of the model is illustrated in Figure 3. First, the experimental data of a trial is parsed and parameterized into a dynamic optimization model of the predicted metabolic growth phenotypes for all community members. Second, this global model is simulated to discover the variable and parameter values that best recapitulate the experimental data while minimizing overfitting. Finally, the simulation predictions – namely phenotype abundances, metabolite concentrations, and growth kinetic parameters – are expressed and exported as figures or spreadsheets.

CommPhitting is designed as an open-source Python package – which leverages the Optlang optimization module [66], the GLPK solver, and ReadTheDocs documentation – to make the method as broadly accessible as is possible. The CommPhitting package further offers robust graphing capabilities through a simulation parameter that can consolidate the high-dimensional simulation results without requiring user proficiency with data science methods, which accelerates discovery by immediately gleaning biological insights after each simulation.

4.2 Simulation insights

The CommPhitting simulations of our 2-member community fit the growth experimental data in Figure 4 remarkably well, and thus instills confidence in the model predictions. A slight anomaly in the experimental data of the maltose + 4HB trials is compensated by CommPhitting in Figure 5, which demonstrates its flexibility while also communicating an area for experimental refinement. The simulations in maltose 6 elucidated the pure expression of the acetate phenotype in *Pseudomonas*, which kept the acetate concentration below 2mM throughout the simulation per Figure 7, while dynamic expression of *E. coli* between its three defined phenotypes: maltose, acetate, and stationary. The consumption of acetate is primarily predicted by *Pseudomonas*, but *E. coli* briefly becomes the greater acetate consumer as it begins to exhaust maltose and reaches the stationary phase. The contrast between *E. coli* in the coculture and as monoculture in Figure 6 exhibits a few interesting differences that illuminate the effect of *Pseudomonas* on *E. coli*. The first observation is that *E. coli* spends much more time in its stationary phenotype throughout the time-course, perhaps because of the inhibitory effects of acetate upon its growth, and *E. coli* enters the stationary phase a few hours sooner than in a coculture. A second observation is that *E. coli* reluctantly delays acetate consumption until the very end of the simulation, although it ultimately consumes more. A third distinction between *E. coli* as a monocultural versus as a coculture with *Pseudomonas* is the predicted $k_{cat,acetate}$, where the monocultural value is $8.89 \frac{1}{hour}$ while the cocultural value is $0.23 \frac{1}{hour}$. The lesser growth rate, or greater consumption of acetate without growth, may be a competitive adaptation of *E. coli* in the coculture to starve a competitor by consuming its food without necessarily investing the resources to grow itself.

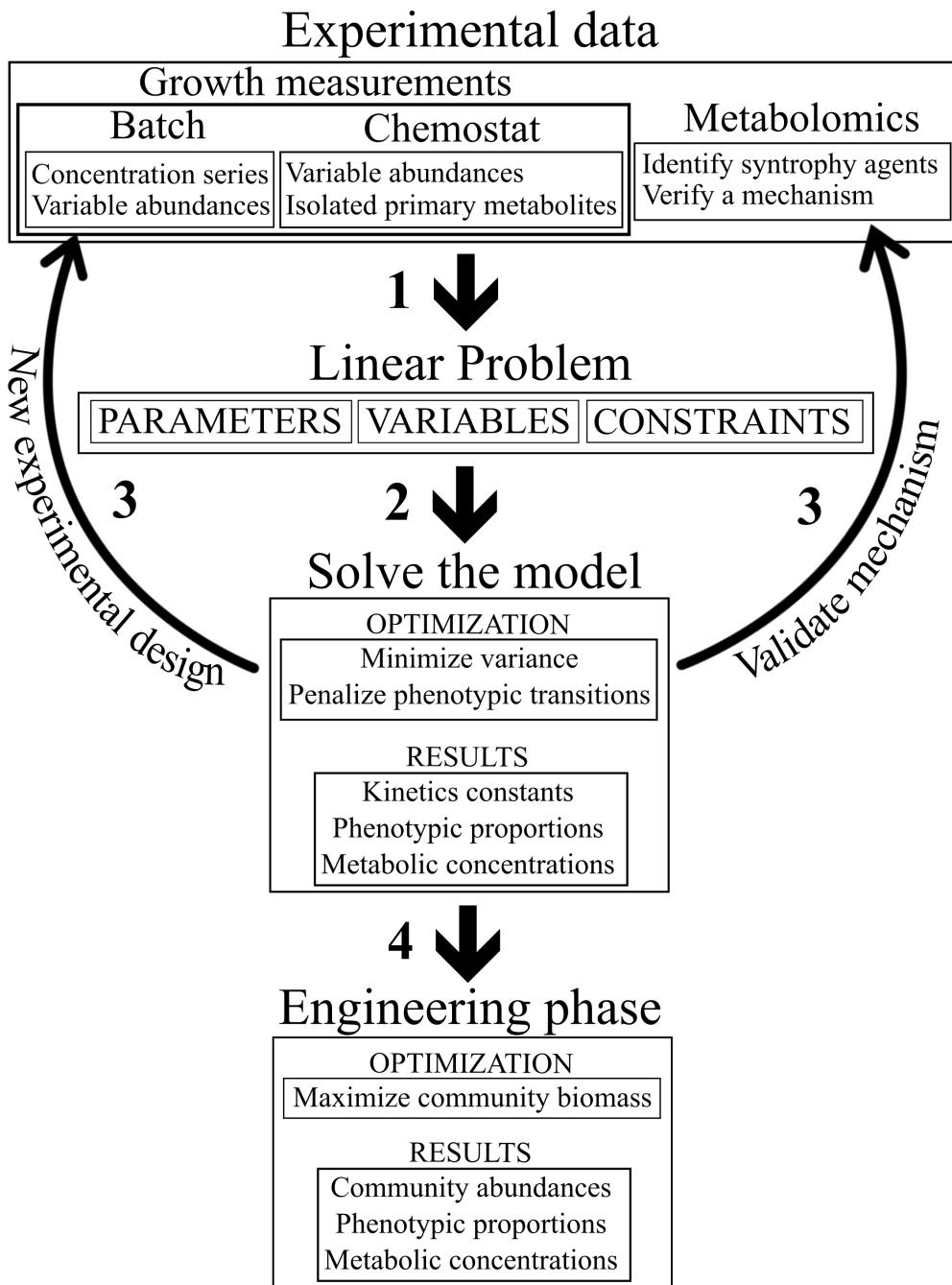


Figure 3: A workflow of the fitting model. **Step 1:** experimental data – from growth and possibly metabolomics measurements – is parsed into a MILP problem that consists of the parameters and variables that are detailed in Table 1 and the constraints that are explained in Section 3. **Step 2:** the linear problem is executed with the objective function of eq. (14). **Step 3:** the simulation results are interpreted to either identify experimental changes that will improve modeling fit or to propose select additional omics measurements that can further solidify mechanistic insights from the fit. Steps 1-3 repeat until a satisfactory fit and mechanistic resolution is achieved. **Step 4:** the fitted model can be used in a forward design-phase, instead of a purely retrospective fitting-phase, by replacing the objective function of the fitted model with one that maximizes community growth or targets other potential ecological objective functions. The system of steps 1-4 create an integrated method for gleaned mechanistic insights of microbial communities and then immediately using these insights to rationally design a community with desirable activity.

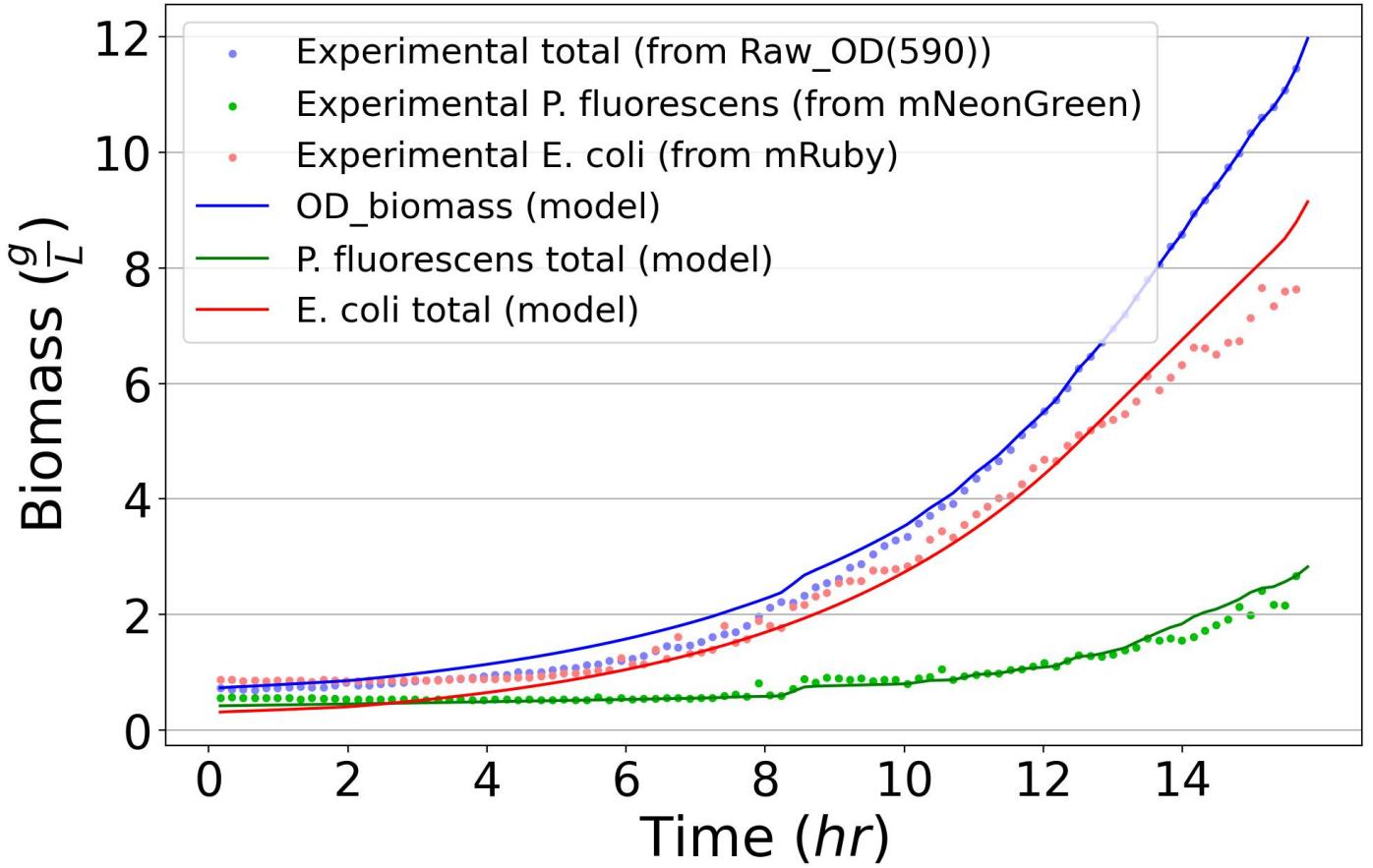


Figure 4: The converted experimental and predicted biomasses of the total community (OD), *E. coli* (RFP), and *P. fluorescens* (GFP) for the coculture experiment on maltose media (Table S1). Tight agreement between the experimental and predicted biomass values improves confidence in the predicted community behaviors and underlying chemical parameters.

4.3 Validation

The CommPhitting simulations predicted minimal *E. coli* acetate consumption during coculture with *P. fluorescens*. This motivated our efforts to find *E. coli* mutants that lack this acetate absorption phenotype in order to corroborate: a) our experimental evidence that acetate is the predominate cross-feeding agent; and b) the prediction that the *E. coli* acetate phenotype is minimally expressed. We purchased mutant knock-outs from the *E. coli* Genetic Stock Center [67] for the four proteins – PoxB, ACS, Pta, and AckA – that contribute directly to acetate metabolism in MG1655. Assessing mutant growth in an acetate mono- or co-culture revealed that only the Δ Pta mutant terminated all acetate consumption; however, because this mutant excretes lactate instead of acetate [68] during sugar metabolism, *Pseudomonas* growth was accentuated by the coculture with this mutant since lactate is a more energetically favorable substrate than acetate. This study affirms that acetate is the primary mediating compound, where *Pseudomonas* responded profoundly to the one mutant whose acetate metabolism was disrupted, and inspires additional simulations of this community with new potential metabolic growth phenotypes, such as lactate, that explore other potential syntrophic interactions.

The peak acetate concentration in Figure 7 is furthermore validated with metabolomics data. The peak concentration in our simulation of 2.4mM corroborates with the peak concentration from the metabolomics data of 1.6mM at the same OD of 1.1. The experimental concentration at this OD is extrapolated between the enclosing metabolomics points at OD's of 0.34 and 1.64; hence, the interpolated concentration embodies some error that may encompass the 50% excess of the predicted value, and it is possible that the predicted and experimental concentrations exhibit an even better match.

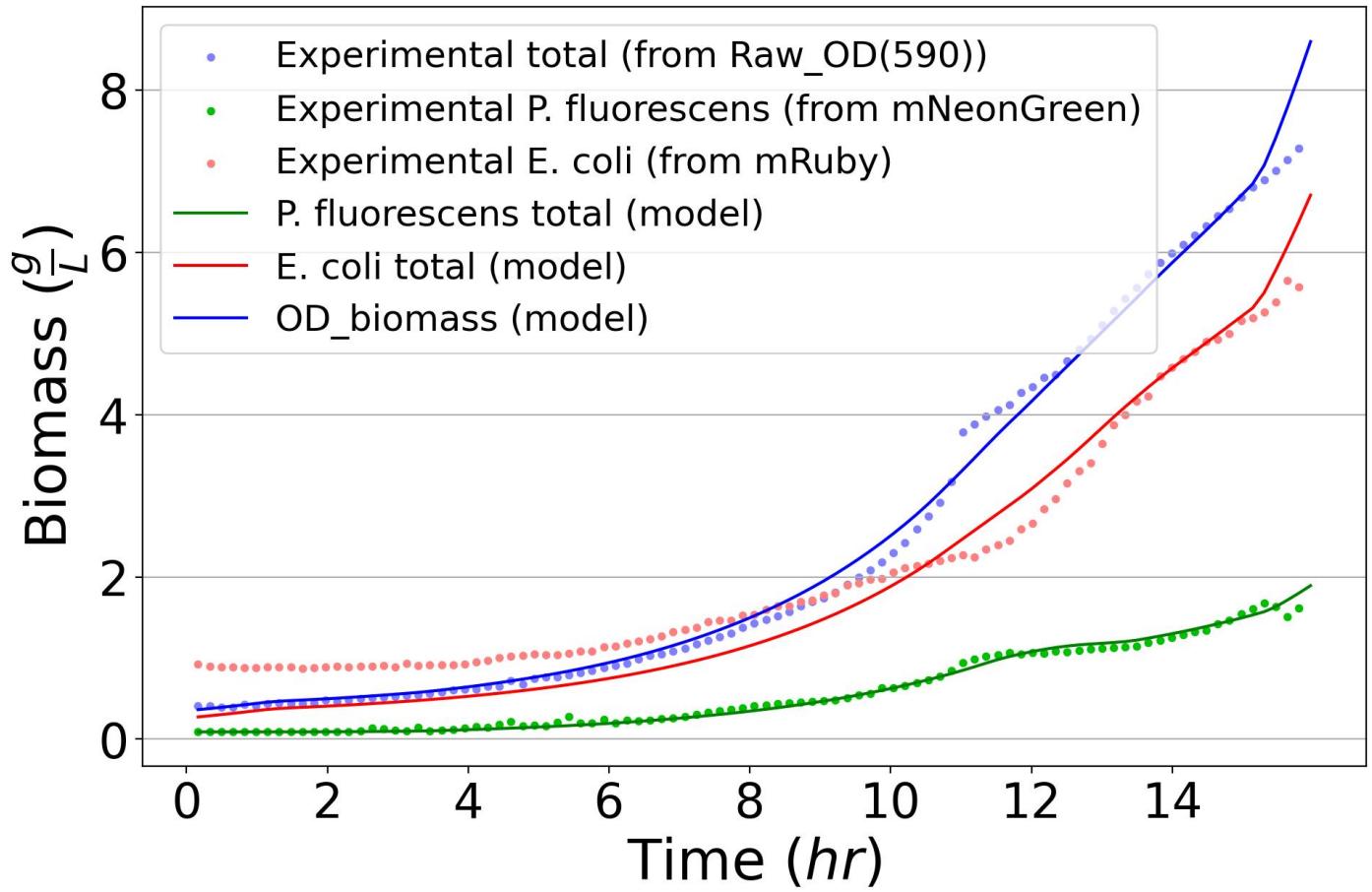


Figure 5: The converted experimental and predicted biomasses of the total community (OD), *E. coli* (RFP), and *P. fluorescens* (GFP) for the coculture experiment on maltose media (Table S1). Tight agreement between the experimental and predicted biomass values improves confidence in the predicted community behaviors and underlying chemical parameters.

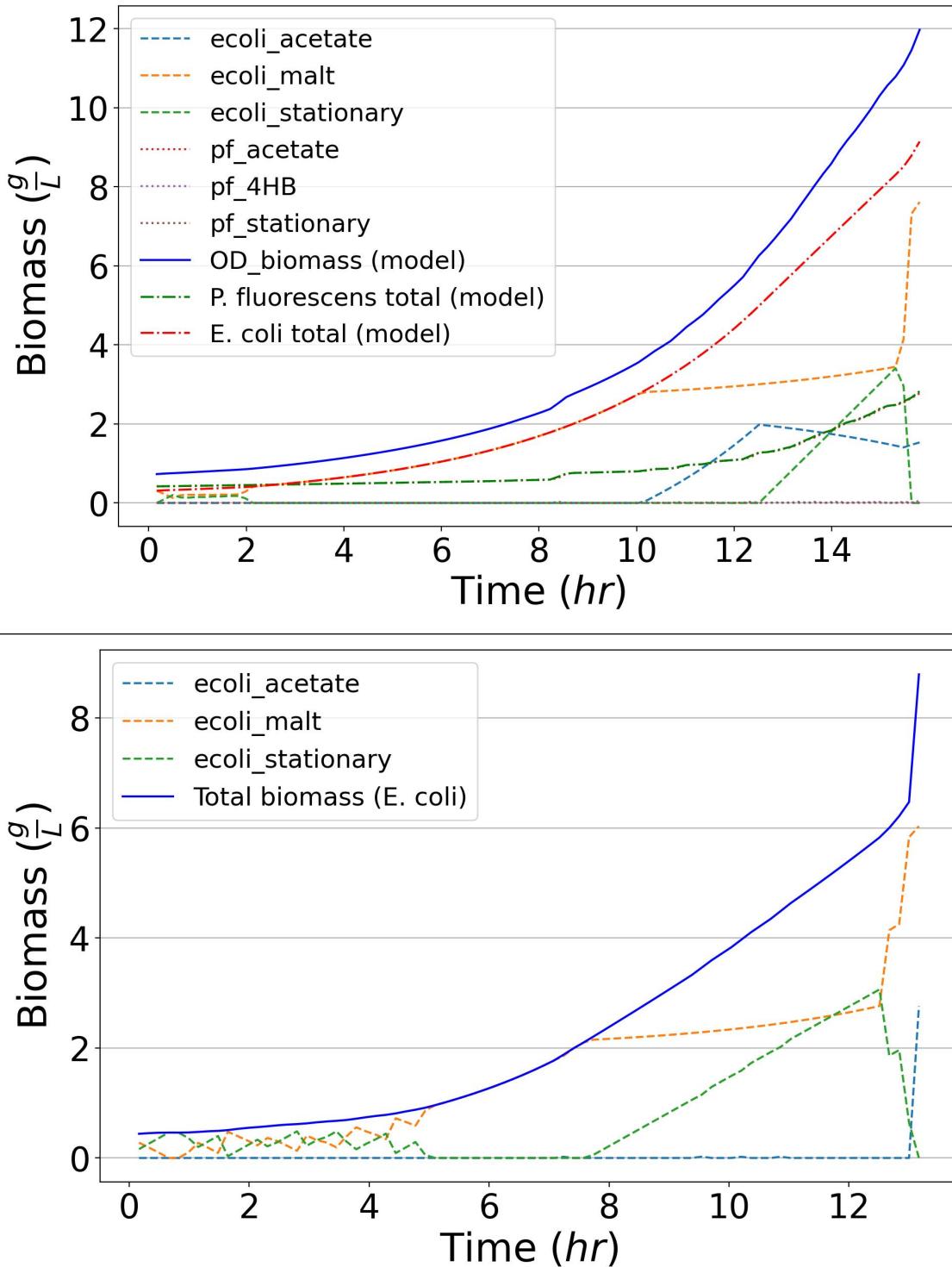


Figure 6: The phenotype abundances for an experiment with 5mM of maltose as the sole carbon sources. The top figure depicts the 1:1 coculture, where *E. coli* outcompeted *P. fluorescens*; however, *P. fluorescens* managed to subsist on the acetate excreta from *E. coli*. The bottom figure depicts a *E. coli* monoculture in maltose, where it reaches the same final biomass but consumes 50% more acetate, albeit very reluctantly. The monoculture interestingly also exhibited much more stationary phenotype during the lag phase than *E. coli* in the same conditions as a coculture. A *P. fluorescens* monoculture figure is not depicted since *P. fluorescens* exhibited no growth.

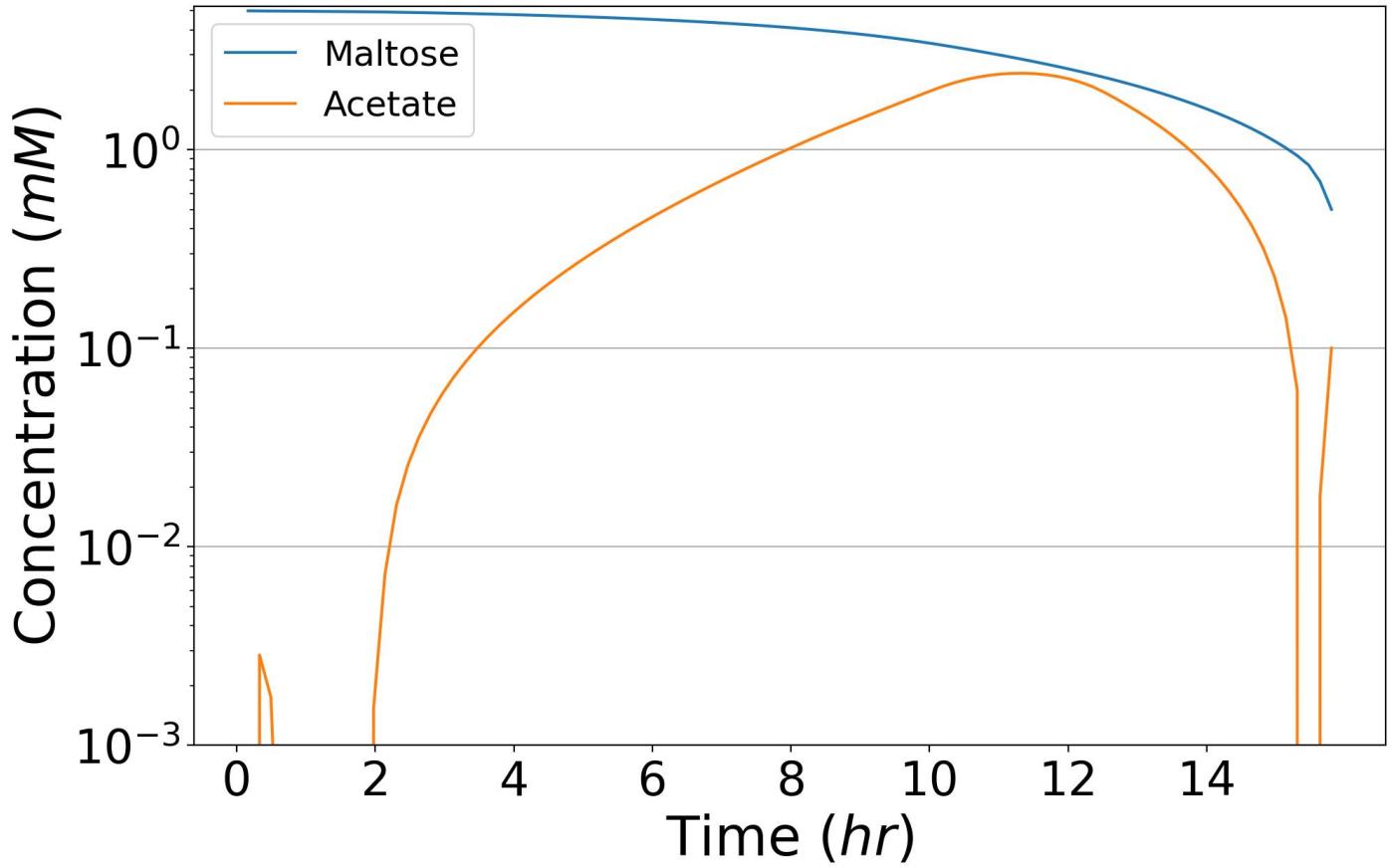


Figure 7: The concentrations of primary carbon sources and the predominate cross-feeding agent acetate from the coculture trial of Figure 6. The dynamic production and subsequent consumption of acetate elicits the unique information that can be derived from CommPhitting simulations. The peak acetate concentration at 2.4mM corroborates with the peak magnitude from the metabolomics data at the same OD value, which validates the concentration mechanisms of CommPhitting.

5 Conclusion

The ubiquity and uniqueness of microbial communities offer untold potential for basic understanding and technological advancement in diverse fields. This potential, however, remains hindered by knowledge gaps of phenotype dynamics and the interspecies exchanges that emanate from the combinatorial complexity of community interactions, which becomes experimentally intractable. Computational efforts to illuminate these gaps often improperly balance mechanistic resolution with data flexibility, and therefore fail to provide new knowledge from experimental datasets. Our Community Phenotype Fitting method (CommPhitting), in contrast, balances resolution and flexibility by adjusting biologically meaningful variables and parameters while best recapitulating the provided dataset of the system. CommPhitting only requires two inputs – member metabolic models and dynamic growth data – and yields time-resolved predictions of metabolite concentrations, predicted abundances for each species and their expressed growth phenotypes, and linear growth rate constants for each expressed phenotype. These outputs offer broad value for a) understanding community biology and b) designing qualitative and quantitative metabolic hypotheses with mechanistic support. The contribution to hypothesis development specifically facilitates the judicious allocation of limited resources to experiments with computational evidence, thereby accelerating experimental and theoretical progress towards understanding microbiome dynamics. Dissonance between predictions and data can further be a source of actionable reflection that leads to experimental improvements, e.g. misaligned concentrations may elicit that the set of parameterized phenotypes are insufficient to explain the experimental data, which improves understanding of the community system.

We envision that CommPhitting can extend from the elucidation phase that we introduce herein to an exploration phase that is introduced in Step 4 of Figure 3. This exploration phase can test engineering hypotheses of community behaviors after member mutations and environmental perturbations, such as the presence of toxins/antibiotics. This exploration phase would fix the parameter and variable values that were derived from fitting experimental data in the elucidation phase and then would exchange the optimization objective to reflect ecological principles that guide the behavior of these systems (e.g. maximizing total community growth). The exploration phase may then come full-circle by looping into the elucidation phase as a design/build/test/learn cycle [69] where the mechanisms of an engineered community that was informed by the exploration phase are revealed.

The aforementioned unique attributes of CommPhitting enable it to resolve the abundances and linear kinetics of the growth phenotypes for each member of a microbial community more robustly than extant methods. This tool will be an invaluable resource for expanding basic knowledge of community dynamics and ultimately engineering these communities for applications in fields as diverse as bioproduction, medical therapies, and national security.

References

- [1] Stephen Nayfach, Simon Roux, Rekha Seshadri, Daniel Udwyar, Neha Varghese, Frederik Schulz, Dongying Wu, David Paez-Espino, I. Min Chen, Marcel Huntemann, Krishna Palaniappan, Joshua Ladau, Supratim Mukherjee, T. B.K. Reddy, Torben Nielsen, Edward Kirton, José P. Faria, Janaka N. Edirisinghe, Christopher S. Henry, Sean P. Jungbluth, Dylan Chivian, Paramvir Dehal, Elisha M. Wood-Charlson, Adam P. Arkin, Susannah G. Tringe, Axel Visel, Helena Abreu, Silvia G. Acinas, Eric Allen, Michelle A. Allen, Lauren V. Alteio, Gary Andersen, Alexandre M. Anesio, Graeme Attwood, Viridiana Avila-Magaña, Yacine Badis, Jake Bailey, Brett Baker, Petr Baldrian, Hazel A. Barton, David A.C. Beck, Eric D. Becraft, Harry R. Beller, J. Michael Beman, Rizlan Bernier-Latmani, Timothy D. Berry, Anthony Bertagnoli, Stefan Bertilsson, Jennifer M. Bhatnagar, Jordan T. Bird, Jeffrey L. Blanchard, Sara E. Blumer-Schuette, Brendan Bohannan, Mikayla A. Borton, Allyson Brady, Susan H. Brawley, Juliet Brodie, Steven Brown, Jennifer R. Brum, Andreas Brune, Donald A. Bryant, Alison Buchan, Daniel H. Buckley, Joy Buongiorno, Hinsby Cadillo-Quiroz, Sean M. Caffrey, Ashley N. Campbell, Barbara Campbell, Stephanie Carr, Jo Lynn Carroll, S. Craig Cary, Anna M. Cates, Rose Ann Cattolico, Ricardo Cavicchioli, Ludmila Chistoserdova, Maureen L. Coleman, Philippe Constant, Jonathan M. Conway, Walter P. Mac Cormack, Sean Crowe, Byron Crump, Cameron Currie, Rebecca Daly, Kristen M. DeAngelis, Vincent Denef, Stuart E. Denman, Adey Desta, Hebe Dionisi, Jeremy Dodsworth, Nina Dombrowski, Timothy Donohue, Mark Dopson, Timothy Driscoll, Peter Dunfield, Christopher L. Dupont, Katherine A. Dynarski, Virginia Edgcomb, Elizabeth A. Edwards, Mostafa S. Elshahed, Israel Figueroa, Beverly Flood, Nathaniel Fortney, Caroline S. Fortunato, Christopher Francis, Claire M.M. Gachon, Sarahi L. Garcia, Maria C. Gazitua, Terry Gentry, Lena Gerwick, Javad Gharechahi, Peter Girguis, John Gladden, Mary Gradoville, Stephen E. Grasby, Kelly Gravuer, Christen L. Grettenberger, Robert J. Gruninger, Jiarong Guo, Mussie Y. Habteselassie, Steven J. Hallam, Roland Hatzenpichler, Bela Hausmann, Terry C. Hazen, Brian Hedlund, Cynthia Henny, Lydie Herfort, Maria Hernandez, Olivia S. Hershey, Matthias Hess, Emily B. Hollister, Laura A. Hug, Dana Hunt, Janet Jansson, Jessica Jarett, Vitaly V. Kadnikov, Charlene Kelly, Robert Kelly, William Kelly, Cheryl A. Kerfeld, Jeff Kimbrel, Jonathan L. Klassen, Konstantinos T. Konstantinidis, Laura L. Lee, Wen Jun Li, Andrew J.

Loder, Alexander Loy, Mariana Lozada, Barbara MacGregor, Cara Magnabosco, Aline Maria da Silva, R. Michael McKay, Katherine McMahon, Chris S. McSweeney, Mónica Medina, Laura Meredith, Jessica Mizzi, Thomas Mock, Lily Momper, Mary Ann Moran, Connor Morgan-Lang, Duane Moser, Gerard Muyzer, David Myrold, Maisie Nash, Camilla L. Nesbø, Anthony P. Neumann, Rebecca B. Neumann, Daniel Noguera, Trent Northen, Jeanette Norton, Brent Nowinski, Klaus Nüsslein, Michelle A. O’Malley, Rafael S. Oliveira, Valeria Maia de Oliveira, Tullis Onstott, Jay Osvatic, Yang Ouyang, Maria Pachiadaki, Jacob Parnell, Laila P. Partida-Martinez, Kabir G. Peay, Dale Pelletier, Xuefeng Peng, Michael Pester, Jennifer Pett-Ridge, Sari Peura, Petra Pjevac, Alvaro M. Plominsky, Anja Poehlein, Phillip B. Pope, Nikolai Ravin, Molly C. Redmond, Rebecca Reiss, Virginia Rich, Christian Rinke, Jorge L. Mazza Rodrigues, William Rodriguez-Reillo, Karen Rossmassler, Joshua Sackett, Ghasem Hosseini Salekdeh, Scott Saleska, Matthew Scarborough, Daniel Schachtman, Christopher W. Schadt, Matthew Schrenk, Alexander Sczyrba, Aditi Sengupta, Joao C. Setubal, Ashley Shade, Christine Sharp, David H. Sherman, Olga V. Shubenkova, Isabel Natalia Sierra-Garcia, Rachel Simister, Holly Simon, Sara Sjöling, Joan Slonczewski, Rafael Soares Correa de Souza, John R. Spear, James C. Stegen, Ramunas Stepanauskas, Frank Stewart, Garret Suen, Matthew Sullivan, Dawn Sumner, Brandon K. Swan, Wesley Swingley, Jonathan Tarn, Gordon T. Taylor, Hanno Teeling, Memory Tekere, Andreas Teske, Torsten Thomas, Cameron Thrash, James Tiedje, Claire S. Ting, Benjamin Tully, Gene Tyson, Osvaldo Ulloa, David L. Valentine, Marc W. Van Goethem, Jean VanderGheynst, Tobin J. Verbeke, John Vollmers, Aurèle Vuillemin, Nicholas B. Waldo, David A. Walsh, Bart C. Weimer, Thea Whitman, Paul van der Wielen, Michael Wilkins, Timothy J. Williams, Ben Woodcroft, Jamie Wootton, Kelly Wrighton, Jun Ye, Erica B. Young, Noha H. Youssef, Feiqiao Brian Yu, Tamara I. Zemskaya, Ryan Ziels, Tanja Woyke, Nigel J. Mouncey, Natalia N. Ivanova, Nikos C. Kyriidis, and Emiley A. Eloe-Fadrosh. A genomic catalog of Earth’s microbiomes. *Nature Biotechnology*, 39(4):499–509, 2021. ISSN 15461696. doi: 10.1038/s41587-020-0718-6.

- [2] Saeed Shoae, Pouyan Ghaffari, Petia Kovatcheva-Datchary, Adil Mardinoglu, Partho Sen, Estelle Pujos-Guillot, Tomas De Wouters, Catherine Juste, Salwa Rizkalla, Julien Chilloux, Lesley Hoyle, Jeremy K. Nicholson, Joel Dore, Marc E. Dumas, Karine Clement, Fredrik Bäckhed, and Jens Nielsen. Quantifying Diet-Induced Metabolic Changes of the Human Gut Microbiome. *Cell Metabolism*, 22(2):320–331, 2015. ISSN 19327420. doi: 10.1016/j.cmet.2015.07.001.
- [3] Manish Kumar, Boyang Ji, Karsten Zengler, and Jens Nielsen. Modelling approaches for studying the microbiome. *Nature Microbiology*, 4(8):1253–1267, 2019. ISSN 20585276. doi: 10.1038/s41564-019-0491-9. URL <http://dx.doi.org/10.1038/s41564-019-0491-9>.
- [4] Saeed Shoae, Fredrik Karlsson, Adil Mardinoglu, Intawat Nookaew, Sergio Bordel, and Jens Nielsen. Understanding the interactions between bacteria in the human gut through metabolic modeling. *Scientific Reports*, 3:1–10, 2013. ISSN 20452322. doi: 10.1038/srep02532.
- [5] Luke McGuigan and Máire Callaghan. The evolving dynamics of the microbial community in the cystic fibrosis lung. *Environmental Microbiology*, 17(1):16–28, 2015. ISSN 14622920. doi: 10.1111/1462-2920.12504.
- [6] Howard F. Jenkinson and Richard J. Lamont. Oral microbial communities in sickness and in health. *Trends in Microbiology*, 13(12):589–595, 2005. ISSN 0966842X. doi: 10.1016/j.tim.2005.09.006.
- [7] Robin K. Pettit. Mixed fermentation for natural product drug discovery. *Applied Microbiology and Biotechnology*, 83(1):19–25, 2009. ISSN 01757598. doi: 10.1007/s00253-009-1916-9.
- [8] Ethan T. Hillman, Louis Edwards Caceres-Martinez, Gozdem Kilaz, and Kevin V. Solomon. Top-down enrichment of oil field microbiomes to limit souring and control oil composition during extraction operations. *AIChE Journal*, (September):1–13, 2022. ISSN 15475905. doi: 10.1002/aic.17927.
- [9] Bryan S. Griffiths and Laurent Philippot. Insights into the resistance and resilience of the soil microbial community. *FEMS Microbiology Reviews*, 37(2):112–129, 2013. ISSN 01686445. doi: 10.1111/j.1574-6976.2012.00343.x.
- [10] William B. Whitman, David C. Coleman, and William J. Wiebe. Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences of the United States of America*, 95(12):6578–6583, 1998. ISSN 00278424. doi: 10.1073/pnas.95.12.6578.
- [11] Anne E. Dukas, Rachel S. Poretsky, and Victoria J. Orphan. Deep-Sea Archaea Fix and Share Nitrogen in Methane-Consuming Microbial Consortia. *Science*, 326(October):422–427, 2009.
- [12] Xuesong He, Jeffrey S. McLean, Lihong Guo, Renate Lux, and Wenyuan Shi. The social structure of microbial community involved in colonization resistance. *ISME Journal*, 8(3):564–574, 2014. ISSN 17517370. doi: 10.1038/ismej.2013.172.

- [13] Daniel van der Lelie, Akihiko Oka, Safiyh Taghavi, Junji Umeno, Ting Jia Fan, Katherine E. Merrell, Sarah D. Watson, Lisa Ouellette, Bo Liu, Muyiwa Awoniyi, Yunjia Lai, Liang Chi, Kun Lu, Christopher S. Henry, and R. Balfour Sartor. Rationally designed bacterial consortia to treat chronic immune-mediated colitis and restore intestinal homeostasis. *Nature Communications*, 12(1):1–17, 2021. ISSN 20411723. doi: 10.1038/s41467-021-23460-x.
- [14] Ricardo Cavicchioli, William J. Ripple, Kenneth N. Timmis, Farooq Azam, Lars R. Bakken, Matthew Baylis, Michael J. Behrenfeld, Antje Boetius, Philip W. Boyd, Aimée T. Classen, Thomas W. Crowther, Roberto Danovaro, Christine M. Foreman, Jef Huisman, David A. Hutchins, Janet K. Jansson, David M. Karl, Britt Koskella, David B. Mark Welch, Jennifer B.H. Martiny, Mary Ann Moran, Victoria J. Orphan, David S. Reay, Justin V. Remais, Virginia I. Rich, Brajesh K. Singh, Lisa Y. Stein, Frank J. Stewart, Matthew B. Sullivan, Madeleine J.H. van Oppen, Scott C. Weaver, Eric A. Webb, and Nicole S. Webster. Scientists’ warning to humanity: microorganisms and climate change. *Nature Reviews Microbiology*, 17(9):569–586, 2019. ISSN 17401534. doi: 10.1038/s41579-019-0222-5. URL <http://dx.doi.org/10.1038/s41579-019-0222-5>.
- [15] V. J. Orphan, C. H. House, K. U. Hinrichs, K. D. McKeegan, and E. F. DeLong. Methane-consuming archaea revealed by directly coupled isotopic and phylogenetic analysis. *Science*, 293(5529):484–487, 2001. ISSN 00368075. doi: 10.1126/science.1061338.
- [16] Jennifer B. Glass and Victoria J. Orphan. Trace metal requirements for microbial enzymes involved in the production and consumption of methane and nitrous oxide. *Frontiers in Microbiology*, 3(FEB):1–20, 2012. ISSN 1664302X. doi: 10.3389/fmicb.2012.00061.
- [17] Jacob D. Palmer and Kevin R. Foster. Bacterial species rarely work together. *Science*, 376(6593):581–582, 2022. ISSN 10959203. doi: 10.1126/science.abn5093.
- [18] Cristiane Aparecida Pereira, Rogério Lima Romeiro, Anna Carolina Borges Pereira Costa, Ana Karina Silva MacHado, Juliana Campos Junqueira, and Antonio Olavo Cardoso Jorge. Susceptibility of *Candida albicans*, *Staphylococcus aureus*, and *Streptococcus mutans* biofilms to photodynamic inactivation: An in vitro study. *Lasers in Medical Science*, 26(3):341–348, 2011. ISSN 02688921. doi: 10.1007/s10103-010-0852-3.
- [19] Thien Fah C. Mah and George A. O’Toole. Mechanisms of biofilm resistance to antimicrobial agents. *Trends in Microbiology*, 9(1):34–39, 2001. ISSN 0966842X. doi: 10.1016/S0966-842X(00)01913-2.
- [20] K. Sauer, M. C. Cullen, A. H. Rickard, L. A.H. Zeef, D. G. Davies, and P. Gilbert. Characterization of nutrient-induced dispersion in *Pseudomonas aeruginosa* PAO1 biofilm. *Journal of Bacteriology*, 186(21):7312–7326, 2004. ISSN 00219193. doi: 10.1128/JB.186.21.7312-7326.2004.
- [21] Prashanth Suntharalingam and Dennis G. Cvitkovitch. Quorum sensing in streptococcal biofilm formation. *Trends in Microbiology*, 13(1):3–6, 2005. ISSN 0966842X. doi: 10.1016/j.tim.2004.11.009.
- [22] Michael T. Mee, James J. Collins, George M. Church, and Harris H. Wang. Syntrophic exchange in synthetic microbial communities. *Proceedings of the National Academy of Sciences of the United States of America*, 111(20), 2014. ISSN 10916490. doi: 10.1073/pnas.1405641111.
- [23] Samay Pande, Holger Merker, Katrin Bohl, Michael Reichelt, Stefan Schuster, Luís F. De Figueiredo, Christoph Kaleta, and Christian Kost. Fitness and stability of obligate cross-feeding interactions that emerge upon gene loss in bacteria. *ISME Journal*, 8(5):953–962, 2014. ISSN 17517370. doi: 10.1038/ismej.2013.211.
- [24] Erica C. Seth and Michiko E. Taga. Nutrient cross-feeding in the microbial world. *Frontiers in Microbiology*, 5 (JULY):1–6, 2014. ISSN 1664302X. doi: 10.3389/fmicb.2014.00350.
- [25] Edwin H. Wintermute and Pamela A. Silver. Emergent cooperation in microbial metabolism. *Molecular Systems Biology*, 6(407):1–7, 2010. ISSN 17444292. doi: 10.1038/msb.2010.66. URL <http://dx.doi.org/10.1038/msb.2010.66>.
- [26] Jack A. Connolly, William R. Harcombe, Michael J. Smanski, Linda L. Kinkel, Eriko Takano, and Rainer Breitling. Harnessing intercellular signals to engineer the soil microbiome. *Natural Product Reports*, 39(2):311–324, 2022. ISSN 14604752. doi: 10.1039/d1np00034a.
- [27] Cristal Zuñiga, Tingting Li, Michael T. Guarnieri, Jackson P. Jenkins, Chien Ting Li, Kerem Bingol, Young Mo Kim, Michael J. Betenbaugh, and Karsten Zengler. Synthetic microbial communities of heterotrophs and phototrophs facilitate sustainable growth. *Nature Communications*, 11(1), 2020. ISSN 20411723. doi: 10.1038/s41467-020-17612-8. URL <http://dx.doi.org/10.1038/s41467-020-17612-8>.

- [28] Nikodimos A. Gebreselassie and Maciek R. Antoniewicz. ^{13}C -metabolic flux analysis of co-cultures: A novel approach. *Metabolic Engineering*, 31:132–139, 2015. ISSN 10967184. doi: 10.1016/j.ymben.2015.07.005. URL <http://dx.doi.org/10.1016/j.ymben.2015.07.005>.
- [29] R Frank Rosenzweig, R R Sharp, David S Treves, and Julian Adams. Microbial evolution in a simple unstructured environment: Genetic differentiation in *Escherichia coli*. *Genetics Society of America*, 137:903–917, 1994.
- [30] Jeffrey D. Orth, Ines Thiele, and Bernhard O. Palsson. What is flux balance analysis? *Nature Biotechnology*, 28(3): 245–248, 2010. ISSN 10870156. doi: 10.1038/nbt.1614.
- [31] Eugen Bauer, Johannes Zimmermann, Federico Baldini, Ines Thiele, and Christoph Kaleta. BacArena: Individual-based metabolic modeling of heterogeneous microbes in complex communities. *PLoS Computational Biology*, 13(5): 1–22, 2017. ISSN 15537358. doi: 10.1371/journal.pcbi.1005544.
- [32] Denny Popp and Florian Centler. μ BialSim: Constraint-Based Dynamic Simulation of Complex Microbiomes. *Frontiers in Bioengineering and Biotechnology*, 8(June), 2020. ISSN 22964185. doi: 10.3389/fbioe.2020.00574.
- [33] Ali R. Zomorodi, Mohammad Mazharul Islam, and Costas D. Maranas. D-OptCom: Dynamic Multi-level and Multi-objective Metabolic Modeling of Microbial Communities. *ACS Synthetic Biology*, 3(4):247–257, 2014. ISSN 21615063. doi: 10.1021/sb4001307.
- [34] Aleksej Zelezniak, Sergej Andrejev, Olga Ponomarova, Daniel R. Mende, Peer Bork, and Kiran Raosaheb Patil. Metabolic dependencies drive species co-occurrence in diverse microbial communities. *Proceedings of the National Academy of Sciences of the United States of America*, 112(20):6449–6454, 2015. ISSN 10916490. doi: 10.1073/pnas.1421834112.
- [35] Christian Diener, Sean M. Gibbons, and Osbaldo Resendis-Antonio. MICOM: Metagenome-Scale Modeling To Infer Metabolic Interactions in the Gut Microbiota. *mSystems*, 5(1), 2020. ISSN 2379-5077. doi: 10.1128/msystems.00606-19.
- [36] Daniel R. Hyduke, Nathan E. Lewis, and Bernhard O. Palsson. Analysis of omics data with genome-scale models of metabolism. *Molecular BioSystems*, 9(2):167–174, 2013. ISSN 1742206X. doi: 10.1039/c2mb25453k.
- [37] Robert A. Dromms and Mark P. Styczynski. Systematic applications of metabolomics in metabolic engineering. *Metabolites*, 2(4):1090–1122, 2012. ISSN 22181989. doi: 10.3390/metabo2041090.
- [38] Xuewen Chen, Ana P. Alonso, Doug K. Allen, Jennifer L. Reed, and Yair Shachar-Hill. Synergy between ^{13}C -metabolic flux analysis and flux balance analysis for understanding metabolic adaption to anaerobiosis in *E. coli*. *Metabolic Engineering*, 13(1):38–48, 2011. ISSN 10967176. doi: 10.1016/j.ymben.2010.11.004. URL <http://dx.doi.org/10.1016/j.ymben.2010.11.004>.
- [39] Keren Yizhak, Tomer Benyamin, Wolfram Liebermeister, Eytan Ruppin, and Tomer Shlomi. Integrating quantitative proteomics and metabolomics with a genome-scale metabolic network model. *Bioinformatics*, 26(12):255–260, 2010. ISSN 13674803. doi: 10.1093/bioinformatics/btq183.
- [40] Weihua Guo and Xueyang Feng. OM-FBA: Integrate transcriptomics data with flux balance analysis to decipher the cell metabolism. *PLoS ONE*, 11(4):1–20, 2016. ISSN 19326203. doi: 10.1371/journal.pone.0154188.
- [41] Rogier J.P. Van Berlo, Dick De Ridder, Jean Marc Daran, Pascale A.S. Daran-Lapujade, Bas Teusink, and Marcel J.T. Reinders. Predicting metabolic fluxes using gene expression differences as constraints. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 8(1):206–216, 2011. ISSN 15455963. doi: 10.1109/TCBB.2009.55.
- [42] Min Kyung Kim and Desmond S. Lun. Methods for integration of transcriptomic data in genome-scale metabolic models. *Computational and Structural Biotechnology Journal*, 11(18):59–65, 2014. ISSN 20010370. doi: 10.1016/j.csbj.2014.08.009. URL <http://dx.doi.org/10.1016/j.csbj.2014.08.009>.
- [43] Vikash Pandey, Daniel Hernandez Gardiol, Anush Chiappino-pepe, and Vassily Hatzimanikatis. TEX-FBA : A constraint-based method for integrating gene expression , thermodynamics , and metabolomics data into genome-scale metabolic models. *bioRxiv*, pages 1–30, 2019.
- [44] Davide Chicco. Ten quick tips for machine learning in computational biology. *BioData Mining*, 10(1):1–17, 2017. ISSN 17560381. doi: 10.1186/s13040-017-0155-3.

- [45] David T. Jones. Setting the standards for machine learning in biology. *Nature Reviews Molecular Cell Biology*, 20(11):659–660, 2019. ISSN 14710080. doi: 10.1038/s41580-019-0176-5. URL <http://dx.doi.org/10.1038/s41580-019-0176-5>.
- [46] Eva Maria Kapfer, Paul Stapor, and Jan Hasenauer. Challenges in the calibration of large-scale ordinary differential equation models. *IFAC-PapersOnLine*, 52(26):58–64, 2019. ISSN 24058963. doi: 10.1016/j.ifacol.2019.12.236. URL <https://doi.org/10.1016/j.ifacol.2019.12.236>.
- [47] Laura Lee, David Atkinson, Andrew G. Hirst, and Stephen J. Cornell. A new framework for growth curve fitting based on the von Bertalanffy Growth Function. *Scientific Reports*, 10(1):1–12, 2020. ISSN 20452322. doi: 10.1038/s41598-020-64839-y.
- [48] Jon E. Lindstrom, Ronald P. Barry, and Joan F. Braddock. Microbial community analysis: A kinetic approach to constructing potential C source utilization patterns. *Soil Biology and Biochemistry*, 30(2):231–239, 1997. ISSN 00380717. doi: 10.1016/S0038-0717(97)00113-2.
- [49] J L Garland. Potential and limitations of BIOLOG for microbial community analysis. *Microbial Biosystems: New Frontiers, Proceedings of the 8th International Symposium on Microbial Ecology*, pages 1–7, 1999. ISSN 1642395X.
- [50] Saratram Gopalakrishnan, Satyakam Dash, and Costas Maranas. K-FIT: An accelerated kinetic parameterization algorithm using steady-state fluxomic data. *Metabolic Engineering*, 61(January):197–205, 2020. ISSN 10967184. doi: 10.1016/j.ymben.2020.03.001.
- [51] Jared L Gearhart, Kristin L Adair, Richard J Detry, Justin D Durfee, Katherine A Jones, and Nathaniel Martin. Comparison of Open-Source Linear Programming Solvers. Technical Report October, Sandia National Laboratories, 2013. URL <http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online>.
- [52] A. Cassioli, D. Di Lorenzo, M. Locatelli, F. Schoen, and M. Sciandrone. Machine learning for global optimization. *Computational Optimization and Applications*, 51(1):279–303, 2012. ISSN 09266003. doi: 10.1007/s10589-010-9330-x.
- [53] Fernando Rojo. Carbon catabolite repression in Pseudomonas: Optimizing metabolic versatility and interactions with the environment. *FEMS Microbiology Reviews*, 34(5):658–684, 2010. ISSN 01686445. doi: 10.1111/j.1574-6976.2010.00218.x.
- [54] Hans C. Bernstein. Reconciling Ecological and Engineering Design Principles for Building Microbiomes. *mSystems*, 4(3):1–5, 2019. ISSN 23795077. doi: 10.1128/msystems.00106-19.
- [55] Tanya Barrett, Dennis B. Troup, Stephen E. Wilhite, Pierre Ledoux, Dmitry Rudnev, Carlos Evangelista, Irene F. Kim, Alexandra Soboleva, Maxim Tomashevsky, Kimberly A. Marshall, Katherine H. Phillippe, Patti M. Sherman, Rolf N. Muertter, and Ron Edgar. NCBI GEO: Archive for high-throughput functional genomic data. *Nucleic Acids Research*, 37(SUPPL. 1):885–890, 2009. ISSN 03051048. doi: 10.1093/nar/gkn764.
- [56] Adam P. Arkin, Robert W. Cottingham, Christopher S. Henry, Nomi L. Harris, Rick L. Stevens, Sergei Maslov, Paramvir Dehal, Doreen Ware, Fernando Perez, Shane Canon, Michael W. Sneddon, Matthew L. Henderson, William J. Riehl, Dan Murphy-Olson, Stephen Y. Chan, Roy T. Kamimura, Sunita Kumari, Meghan M. Drake, Thomas S. Brettin, Elizabeth M. Glass, Dylan Chivian, Dan Gunter, David J. Weston, Benjamin H. Allen, Jason Baumohl, Aaron A. Best, Ben Bowen, Steven E. Brenner, Christopher C. Bun, John Marc Chandonia, Jer Ming Chia, Ric Colasanti, Neal Conrad, James J. Davis, Brian H. Davison, Matthew Dejongh, Scott Devoid, Emily Dietrich, Inna Dubchak, Janaka N. Edirisinghe, Gang Fang, José P. Faria, Paul M. Frybarger, Wolfgang Gerlach, Mark Gerstein, Annette Greiner, James Gurtowski, Holly L. Haun, Fei He, Rashmi Jain, Marcin P. Joachimiak, Kevin P. Keegan, Shinnosuke Kondo, Vivek Kumar, Miriam L. Land, Folker Meyer, Marissa Mills, Pavel S. Novichkov, Taeyun Oh, Gary J. Olsen, Robert Olson, Bruce Parrello, Shiran Pasternak, Erik Pearson, Sarah S. Poon, Gavin A. Price, Srividya Ramakrishnan, Priya Ranjan, Pamela C. Ronald, Michael C. Schatz, Samuel M.D. Seaver, Maulik Shukla, Roman A. Sutormin, Mustafa H. Syed, James Thomason, Nathan L. Tintle, Daifeng Wang, Fangfang Xia, Hyunseung Yoo, Shinjae Yoo, and Dantong Yu. KBase: The United States department of energy systems biology knowledgebase. *Nature Biotechnology*, 36(7):566–569, 2018. ISSN 15461696. doi: 10.1038/nbt.4163.
- [57] Christopher S. Henry, Matthew Dejongh, Aaron A. Best, Paul M. Frybarger, Ben Lindsay, and Rick L. Stevens. High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nature Biotechnology*, 28(9):977–982, 2010. ISSN 10870156. doi: 10.1038/nbt.1672.

- [58] Mark Lotkin. A Note on the Midpoint Method of Integration. *Journal of the ACM (JACM)*, 3(3):208–211, 1956. ISSN 1557735X. doi: 10.1145/320831.320840.
- [59] L. F. Shampine. Stability of the leapfrog/midpoint method. *Applied Mathematics and Computation*, 208(1):293–298, 2009. ISSN 00963003. doi: 10.1016/j.amc.2008.11.029. URL <http://dx.doi.org/10.1016/j.amc.2008.11.029>.
- [60] J. C. Butcher. A history of Runge-Kutta methods. *Applied Numerical Mathematics*, 20:247–260, 1996.
- [61] W. H. Witty. A New Method of Numerical Integration of Differential Equations. *Mathematics of Computation*, 18(87):497, 1964. ISSN 00255718. doi: 10.2307/2003774.
- [62] Rosemarie Wilton, Angela J. Ahrendt, Shalaka Shinde, Deirdre J. Sholto-Douglas, Jessica L. Johnson, Melissa B. Brennan, and Kenneth M. Kemner. A New Suite of Plasmid Vectors for Fluorescence-Based Imaging of Root Colonizing Pseudomonads. *Frontiers in Plant Science*, 8(February):1–15, 2018. ISSN 1664462X. doi: 10.3389/fpls.2017.02242.
- [63] Ramy K. Aziz, Daniela Bartels, Aaron Best, Matthew DeJongh, Terrence Disz, Robert A. Edwards, Kevin Formsma, Svetlana Gerdes, Elizabeth M. Glass, Michael Kubal, Folker Meyer, Gary J. Olsen, Robert Olson, Andrei L. Osterman, Ross A. Overbeek, Leslie K. McNeil, Daniel Paarmann, Tobias Paczian, Bruce Parrello, Gordon D. Pusch, Claudia Reich, Rick Stevens, Olga Vassieva, Veronika Vonstein, Andreas Wilke, and Olga Zagnitko. The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics*, 9:1–15, 2008. ISSN 14712164. doi: 10.1186/1471-2164-9-75.
- [64] Sven E.F. Borgos, Sergio Bordel, Håvard Sletta, Helga Ertesvåg, Øyvind Jakobsen, Per Bruheim, Trond E. Ellingsen, Jens Nielsen, and Svein Valla. Mapping global effects of the anti-sigma factor MucA in *Pseudomonas fluorescens* SBW25 through genome-scale metabolic modeling. *BMC Systems Biology*, 7:1–15, 2013. ISSN 17520509. doi: 10.1186/1752-0509-7-19.
- [65] Jonathan M Monk, Colton J Lloyd, Elizabeth Brunk, Nathan Mih, Anand Sastry, Zachary King, Rikiya Takeuchi, Wataru Nomura, Zhen Zhang, Hirotada Mori, Adam M Feist, and Bernhard O Palsson. iML1515 , a knowledgebase that computes *E. coli* traits. *Nature Biotechnology*, 35(10):8–12, 2017.
- [66] Kristian Jensen, Joao G.R. Cardoso, and Nikolaus Sonnenschein. Optlang: An algebraic modeling language for mathematical optimization. *The Journal of Open Source Software*, 2(9):139, 2017. doi: 10.21105/joss.00139.
- [67] Tomoya Baba, Takeshi Ara, Miki Hasegawa, Yuki Takai, Yoshiko Okumura, Miki Baba, Kirill A. Datsenko, Masaru Tomita, Barry L. Wanner, and Hirotada Mori. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: The Keio collection. *Molecular Systems Biology*, 2, 2006. ISSN 17444292. doi: 10.1038/msb4100050.
- [68] Sara Castaño-Cerezo, José M. Pastor, Sergio Renilla, Vicente Bernal, José L. Iborra, and Manuel Cánovas. An insight into the role of phosphotransacetylase (pta) and the acetate/acetyl-CoA node in *Escherichia coli*. *Microbial Cell Factories*, 8:1–19, 2009. ISSN 14752859. doi: 10.1186/1475-2859-8-54.
- [69] Christopher E. Lawson, William R. Harcombe, Roland Hatzenpichler, Stephen R. Lindemann, Frank E. Löffler, Michelle A. O’Malley, Héctor García Martín, Brian F. Pfleger, Lutgarde Raskin, Ophelia S. Venturelli, David G. Weissbrodt, Daniel R. Noguera, and Katherine D. McMahon. Common principles and best practices for engineering microbiomes. *Nature Reviews Microbiology*, 17(12):725–741, 2019. ISSN 17401534. doi: 10.1038/s41579-019-0255-9. URL <http://dx.doi.org/10.1038/s41579-019-0255-9>.