

5076 (errors:5) words

File: main.tex  
Encoding: utf8  
Sum count: 5076  
Words in text: 4358  
Words in headers: 49  
Words outside text (captions, etc.): 567  
Number of headers: 19  
Number of floats/tables/figures: 8  
Number of math inlines: 85  
Number of math displayed: 17  
Subcounts:  
text+headers+captions (#headers/#floats/#inlines/#displayed)  
25+9+0 (1/0/0/0) \_top\_  
352+1+0 (1/0/0/0) Section: Abstract  
695+1+81 (1/1/1/0) Section: Introduction  
74+1+12 (1/1/0/0) Section: Methods} \label{methods  
323+2+0 (1/0/0/0) Subsection: Metabolic phenotypes} \label{phenotype\_methods  
565+6+0 (4/0/48/13) Subsection: Constraints} \label{constraints  
49+1+0 (1/0/1/1) Subsection: Objective} \label{objective  
1010+13+43 (4/1/27/3) Subsection: E. coli-Pseudomonas competitive community  
0+3+0 (1/0/0/0) Section: Results and Discussion  
143+8+171 (1/1/0/0) Subsection: New data fitting method for iterative experimental design  
209+2+0 (1/0/3/0) Subsection: Simulation insights  
475+1+260 (1/4/5/0) Subsection: Validation  
438+1+0 (1/0/0/0) Section: Conclusion

File: output.bbl  
Encoding: utf8  
Sum count: 0  
Words in text: 0  
Words in headers: 0  
Words outside text (captions, etc.): 0  
Number of headers: 0  
Number of floats/tables/figures: 0  
Number of math inlines: 0  
Number of math displayed: 0

Total  
Sum count: 5076  
Words in text: 4358  
Words in headers: 49  
Words outside text (captions, etc.): 567  
Number of headers: 19  
Number of floats/tables/figures: 8  
Number of math inlines: 85  
Number of math displayed: 17  
Files: 2  
Subcounts:  
text+headers+captions (#headers/#floats/#inlines/#displayed)  
4358+49+567 (19/8/85/17) File(s) total: main.tex  
0+0+0 (0/0/0/0) File(s) total: output.bbl

(errors:5)

# A microbial community growth model for dynamic phenotype predictions

Andrew P. Freiburger<sup>ID 1</sup>, Jeffrey A. Dewey<sup>ID 2</sup>, Fatima Foflonker<sup>ID 1</sup>, Gyorgy Babnigg<sup>ID 2</sup>, Dionysios A. Antonopoulos<sup>ID 2</sup>, and Christopher Henry<sup>ID 1\*</sup>

<sup>1</sup>Argonne National Laboratory, Computing, Environment, and Life Sciences Division, Lemont, IL 60148 USA

<sup>2</sup>Argonne National Laboratory, Biosciences Division, Lemont, IL 60148 USA  
\*e-mail: cshenry@anl.gov

May 23, 2023

## 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] and pathogens [4, 5], industrial assets [6] and antagonists [7], and ecological agents of biogeochemical cycling [8, 9, 10]. Microbial communities are therefore essential to understand in fields as diverse as medicine [11, 12] and climatology [13, 14, 15], but these systems are remain understudied [16]. Microbial communities embody peculiar ecology, where intimate co-habitation invites competition [17] yet it remains biologically favorable according to several advantages. Firstly, diversity in a) genetics, 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, the competitive community environment triggers diversification of members into ecological niches and specializations that consequentially augments production efficiency [22] and biomass growth rates [23]. These community dynamics are predicated on an economy of metabolic cross-feeding (syntrophy) [24, 25], where auxotrophs are satiated

by excreted byproducts. Communities can be engineered [26] for greater bioproduction efficiency than monocultures [6] or sustainable biotechnology by coupling phototrophs and heterotrophs [27]; however, ignorance of syntrophy hinders progress in synthetic community design.

Experimental methods have several notable problems measuring syntropy. One problem is that rapidly consumed compounds may never accumulate in the media and therefore remain invisible to most experimental methods, save fluxomics isotope labeling [28]. Another, more formidable, problem is that syntrophic exchanges are combinatorial with the size of a community, so detecting and tracing these exchanges quickly becomes untenable for even communities with a handful of members. A third problem is that member phenotypes dynamically change with environmental conditions, even in clonal monocultures [29], which introduces considerable time-variability to syntrophic exchanges.

Computational biology may be a necessary supplement to experimental methods for resolving syntrophic exchanges. Flux Balance Analysis (FBA) [30] is a prominent framework for simulating cellular metabolism that does not require kinetics or -omics data [31], yet has been expanded to accommodate metabolomics [32, 33], proteomics [34], transcriptomics [35, 36], and multi-omics [37] data. FBA is limited by an under-determined problem of more variables than equations, which compromises precision and reproducibility, and moreover cannot natively capture phenotypic variability, notwithstanding Dynamic FBA (dFBA) [38] to predict dynamic species growth. Machine-learning algorithms (ML) [39, 40] in contrast can arbitrarily resolve phenotype abundances from experimental data, but blackbox predictions from ML do not advance basic understanding that can facilitate design principles. Differential equation models are inversely mechanistic but are too stiff [41] to accommodate variability in phenotype abundances or diverse experimental datasets. Data fitting through optimization attractively balances the strengths and weaknesses of the aforementioned methods by flexibly accommodating diverse experimental data [42, 43, 44] while offering mechanistic resolution that advances basic understanding. Data fitting methods have been accordingly applied to deduce kinetics coefficients [45] and phenotype expression [46] from experimental -omics data, but these fitting methods miss the additional dimension of dynamic phenotype abundances, which is essential to precisely engineer microbial communities.

We therefore developed a global, data fitting, optimization model of community phenotypes (CommPhitting) that yields predictions of a) dynamic biomass for each community member and member phenotype; b) substrate media concentrations; c) update kinetics for each member phenotype [47]; and d) conversion factors from each experimental signal to  $\frac{g}{L}$  biomass. CommPhitting, as a multi-species dynamic model, defines metabolite flux profiles for each phenotype of each member – derived from a rigorous sequence of optimizations that channel metabolism based on the specified consumption for each given phenotype at a defined minimal growth rate – that are selectively expressed by the optimization over time to best recapitulate the experimental data. CommPhitting solves all data points simultaneously in a linear problem [48] to ascertain the global optimum [49], and thereby circumvent deceptive local minima.

We exemplify CommPhitting with an idealized 2-member community of model organisms (*Escherichia coli* and *Pseudomonas fluorescens*) who exhibit complementary carbon preferences [50] but have not been investigated as a coculture. This benchtop community [51] demonstrated pivotal acetate cross-feeding, validated by numerous metabolomic experiments, where *E. coli* nourishes *P. fluorescens* in an otherwise insufficient growth media (Figure 1). CommPhitting achieves a remarkable fit to experimental data, and offers numerous testable hypotheses of metabolic concentrations, species update kinetics, and phenotype abundances. CommPhitting is available as an open-source Python module in the ModelSEEDpy, which will further augment the ability of this light-weight yet robust fitting model to illuminate nuances of community interactions and ultimately foster the rational design of microbial communities for myriad fundamental and industrial applications.

## 3 Methods

CommPhitting essentially transforms the ordinary differential equations that define microbial growth and chemical activity into a mixed integer linear optimization problem (MILP) with the variables, constraints, and parameters that are described in Table 1. The MILP objective is to minimize error between the predicted and experimental values – either biomass growth or species abundance – while avoiding overfitting by minimizing stationary phenotypes for each member. The MILP constraints and objective expression are detailed in the following sections.

### 3.1 Metabolic phenotypes

CommPhitting first defines the metabolic phenotype profiles for each member from specified exchanges and their genome-scale metabolic models (GEMs) that are reconstructed via the ModelSEED pipeline [52], which ensures the compatibility of exchange ID's between all members models to adequately simulate syntropy. The GEMs are then manipulated into phenotype profiles, for all possible carbon-sources by default, through the following sequence of constraints, objectives, and optimizations.

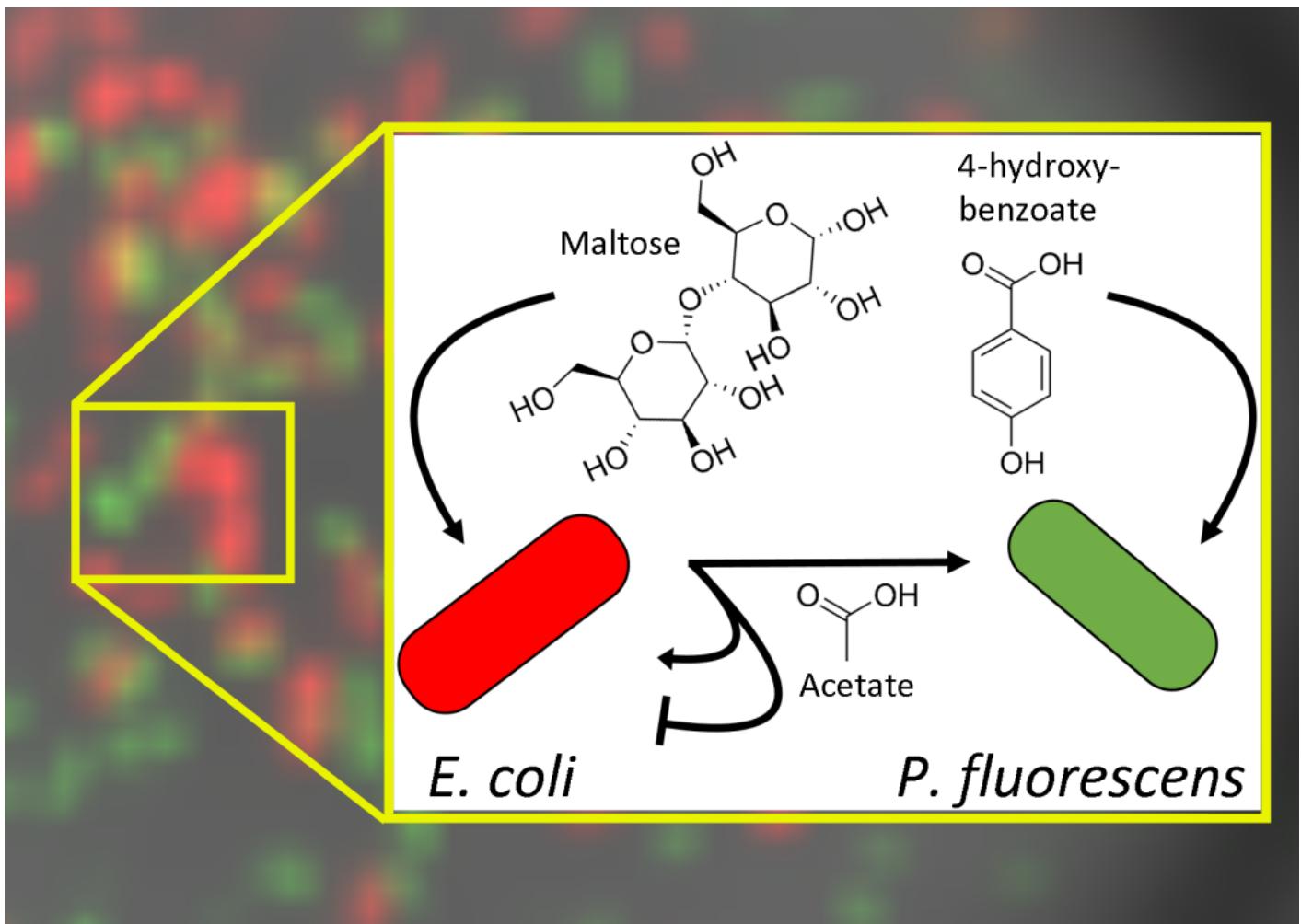


Figure 1: The primary metabolic exchanges that were experimentally elucidated and computationally modeled. The acetate byproduct of *E. coli* is the pivotal exchange of metabolic interest, where this source subsists *P. fluorescens*'s existence in maltose and exhibits interesting dual effects upon *E. 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*.

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

Term	Type	Description
<b>DIMENSIONS</b>		
$s$	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.
<b>PARAMETERS</b>		
$E_{s,t}$	Float	The experimental growth signal for a species at instant $t$ .
$es_{s,k}$	Boolean	A designation of truth for $k \in s$
$\Delta t$	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 $k$ .
<b>VARIABLES</b>		
$EC_k$	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.

1. The GEM first is constrained by several cellular and environmental conditions. Hydrogen consumption is prohibited, which forces utilization of the specific carbon source. Second, oxygen consumption is limited to stoichiometric equivalency to the total consumption of phenotype carbon source(s), although, this threshold can be tailored for specific species biology and to properly regulate overflow metabolism. Finally, a minimal biomass growth is defined to prevent the subsequent steps from optimizing to extreme fluxes or biomass to zero.
2. The total influx of all carbon compounds and compounds with undefined formula are minimized, with a 1000x smaller coefficient on the defined phenotype carbon source(s) to incentivize its utilization. This minimization focuses the flux profile to the fluxes that are attributed phenotype carbon source(s) and also maximizes biomass yield for the fixed biomass growth. The simulated exchanges are constrained to the fluxes from this minimization and remain unchanged in subsequent optimizations.
3. Optionally, the flux of a specified phenotype excreta is maximized, which incorporates experimentally observed excretions from given carbon sources. The exchange fluxes for these specified excreta are constrained after this maximization.
4. The model is simulated, either with parsimonious FBA [53] as the minimization of all fluxes in the GEM where amenable with the specified model (presumably emulates efficiencies that evolution found in the metabolic network) or standard FBA otherwise.

The non-zero exchange fluxes following the above optimizations and constraints, become the metabolic phenotype that is simulated for the given community in CommPhitting. The above five-step sequence is repeated to create irreducible metabolic profiles for each phenotype of all community members.

## 3.2 Constraints

Various constraints of cellular chemistry or growth dynamics are necessary to capture community biology in CommPhitting, which are defined in the following sections.

### 3.2.1 Biomass abundances

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

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

The final  $EB_{s,t}$  is limited to be within 30% of the requisite biomass to consume 90% of the carbon source(s), according to the phenotype metabolic profile, which better anchors the predicted biomass to the experimental biomass and improves the accuracy of the  $EC_s$  conversion coefficient. Variance ( $EV_{s,t}$ ) between the converted experimental biomass and the predicted biomass  $b_{k,t}$  is determined for each phenotype  $k \in K$

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

where the binary  $es_s$  variable filters for only the phenotypes of species  $s$ . The predicted biomass  $b_{k,t}$  is further partitioned into five sub-groups that each exhibit a distinct  $kcat_z$  value and are weighted along a gradient

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

The expression of each biomass partition is controlled by its binary variable  $bin_k^z$

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

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 in each simulation

$$0 \leq \sum_{z,k}^{Z,K} (bin_k^z) \leq 4 . \quad (6)$$

The  $kcat_k$  value of each phenotype is calculated from the expression of  $kcat_z$  partitions over three sets of  $kcat_z$  values for each  $z \in Z$  partition ([100, 10, 1, 0.1, 0.01], then [4, 2, 1, 0.5, 0.25], and finally [1, 0.5, 0.25, 0.125, 0.0625])

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

where  $kcat_{k,new}$  becomes  $kcat_{k,old}$  in the next loop and the first  $kcat_{k,old}$  is an arbitrary guess. This approach approximates  $kcat_k$  values for each phenotype with increasing refinement over the iterations while maintaining linearity of the growth constraint in eq. (10) for our MILP. The above sequence for determining  $kcat_k$  is repeated in each of 5 time range partitions – [0-10]%, [10,25]%, [25,45]%, [45,70]%, [70,100]% – which permits variable growth kinetics at different substrate concentrations and thereby emulates the logarithmic Michaelis-Menten curve.

### 3.2.2 Phenotype transitions

The biomass change over time is calculated for non-stationary (growing) phenotypes via Heun's integration method [54, 55], which – for an arbitrary function  $y$ , its derivative  $y'$ , and a timestep  $\Delta t$  – is defined as

$$y_{t+1} = y_t + \frac{1}{2} * \Delta t * (y'_t + y'_{t+1}) . \quad (8)$$

Heun's method is a 2<sup>nd</sup>-order Runge-Kutta formulation [56] that captures dynamic changes with high numerical accuracy [57] while maintaining a linear formulation that is amenable with our MILP. Our application of Heun's method

$$b_{t,k} + \frac{\Delta t}{2}(g_{t,k} + g_{t+1,k}) + cvf_{t,k} - cvt_{t,k} = b_{t+1,k} . \quad (9)$$

equates biomass in the next timestep ( $b_{k,t+1}$ ) to the current biomass ( $b_{k,t}$ ) plus the midpoint of biomass growth over the timestep (with the biomass growth rate  $g_{k,t}$  as the derivative of biomass:  $\frac{\Delta \text{biomass}}{\Delta t}$ ) and plus the net transition of biomass to growth phenotypes  $cvf_{k,t} - cvt_{k,t}$ . The growth rate for phenotype  $k$  is constrained

$$kcat_k * \sum_z^Z (b_{k,t}^z) = g_{k,t} , \quad (10)$$

as the product of the current biomass  $b_{k,t}$  and the growth rate constant  $kcat_k$ , which reflects 1<sup>st</sup>-order kinetics with respect to biomass abundance and zero-order kinetics with respect to substrate concentration. Every species is defined with a stationary (not growing) phenotype that mediates, and subtly delays, phenotype transitions to reflect lagging expression changes. Further, the stationary phenotype allows each species to cease growth once carbon source(s) are exhausted. Future biomass for the stationary phenotype is analogous to eq. (9),

$$b_{k,t} - \sum_s^S (es_s * (cvf_{k,t} - cvt_{k,t})) = b_{k,t+1} \quad (11)$$

except that a)  $g_{k,t} = 0$  by definition; b) the net transition to growth phenotypes is negative, to reflect to opposite direction; and c) the many possible transitions to/from the stationary phenotype are summed. The fraction of biomass that can transition is constrained

$$cvt_{k,t} \leq bcv * b_{k,t} + cvmin \quad (12)$$

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

### 3.2.3 Concentrations

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

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

as another application of Heun's method (eq. (8)), where the future concentration ( $c_{t+1,i}$ ) equates the current concentration ( $c_{t,i}$ ) plus the inner product of the flux of metabolite  $i$  ( $n_{i,k}$ ) from the phenotype profile (section 3.1) and the biomass growth over the timestep midpoint.

### 3.3 Objective

The objective expression

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

minimizes the sum of variance ( $EV$ ) from eq. (2) and the sum total amount of biomass in the stationary phenotypes to best fit the data while avoiding overfitting by means of the non-growth phenotype and excessively transitioning phenotypes, where the stationary phenotype mediates phenotype transitions in eq. (9).

## 3.4 *E. coli*-*Pseudomonas* 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  $-80^{\circ}\text{C}$  before preparation for electrocompetency, and were transformed with a plasmid to constitutively express either mNeongreen or mRuby2 fluorescent proteins (GFP and RFP), respectively [58]. Transformed cells were freshly streaked on a LB agar plate with appropriate antibiotics from  $-80^{\circ}\text{C}$  glycerol stocks, and were incubated overnight at  $30^{\circ}\text{C}$ . A single colony from the plate was picked, placed into liquid LB (Lennox) broth with the antibiotics, and shaken @ 250 RPM overnight at  $30^{\circ}\text{C}$ . 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  $\mu\text{L}$  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  $\mu\text{L}$  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  $30^{\circ}\text{C}$ . 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:  $12^{\circ}\text{C}$  and  $37^{\circ}\text{C}$  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  $30^{\circ}\text{C}$ , 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 60h at  $12^{\circ}\text{C}$  but does not grow at  $37^{\circ}\text{C}$ , *E. coli* shows the opposite growth trend, and both organisms grow robustly overnight at  $30^{\circ}\text{C}$ . 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.

cells were prepared and analyzed as described above in the 'Plate-based Coculture Growth', 3.4.1.x section. Subsequently, the cell supernatant was separated via centrifugation at 13000xg for 10 minutes prior to metabolomics analysis.

### 3.4.2 Metabolomics

Various metabolomics measurements were acquired for our community system. The cells were prepared and analyzed as described above in the 'Plate-based Coculture Growth', 3.4.1.x section. Subsequently, the cell supernatant was separated via centrifugation at 13000xg for 10 minutes prior to metabolomics analysis.

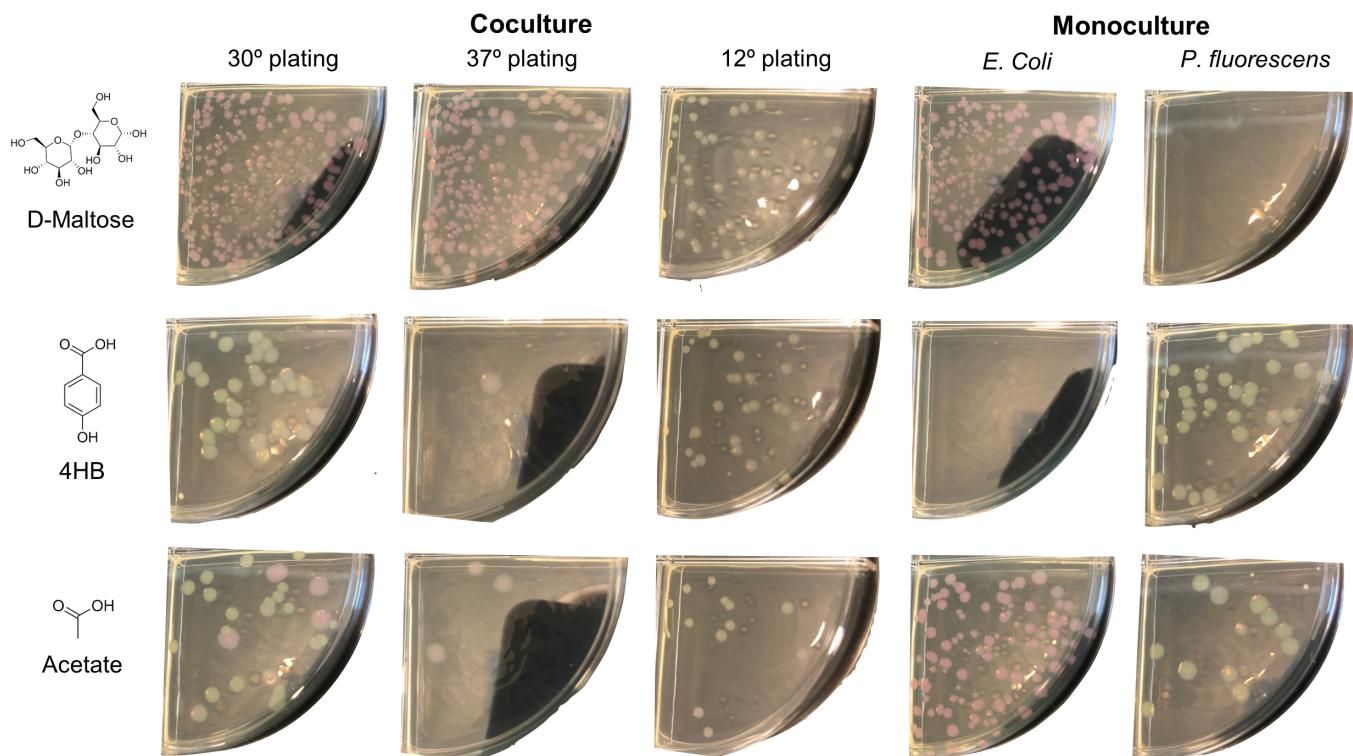


Figure 2: Organism abundance after co- and monoculture experiments in M9 media containing different carbon sources is visualized via plating at select temperatures. Coculture growth patterns replicate those observed in 96-well plate reader assays, confirming that fluorescence reporters accurately represent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

The University of Chicago metabolomics lab incubated the samples at -80 °C for at least one hour, or up to overnight. Extraction solvent (4 volumes of 100% methanol spiked with internal standards and stored at -80 °C) was added to the liquid sample (1 volume) in a microcentrifuge tube. Tubes were then centrifuged at -10 °C, 20,000 x g, for 15 min, and the supernatant was used for metabolomic analysis. The targeted data was acquired via an Agilent 8890/5977B and 7890B/5977B GCMS with chemical ionization (CI). The untargeted data was acquired via a Positive electron impact - gas chromatography - mass spectrometry (+EI-GC-MS, Agilent, 7890B).

### 3.4.3 Adapting the formulation for growth data

The general formulation in Sections 3.1-3.3 was tailored to this particular 2-member community through the following derivations. First, signal-to-biomass conversion factors ( $RFPC$ ,  $GFPc$ , and  $ODC$ ) were defined from eq. (1) for each signal and multiplied by each respective experimental signal ( $RFP_t$ ,  $GFP_t$ , and  $OD_t$ ) to acquire the corresponding biomass ( $RFPB_t$ ,  $GFPB_t$ , and  $ODB_t$ )

$$\begin{aligned} RFP_t * RFPC &= RFPB_t \\ GFP_t * GFPc &= GFPB_t \\ OD_t * ODC &= ODB_t . \end{aligned} \quad (15)$$

Second, the variance constraint in eq. (2) is adapted for each experimental signal

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

where  $pf_k$  and  $ec_k$  are binary variables that filter for *Psuedomonas* and *E. coli*, respectively. Third, the objective function of eq. (14) is applied to 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, data at and beyond plateau of the OD signal are removed, which focuses on the lag and growth phases and mitigates over-production of fluorescent proteins once cells enter their stationary phase. These data could also be filtered by phenotype-specific conversions of the fluorescence signals into biomass abundance, instead of species-specific conversion.

The *P. fluorescens* and *E. coli* GEMs were constructed in KBase Narrative 93465 by leveraging RAST [59] and the ModelSEED pipeline [52]. 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* [60] and *P. fluorescens* [61], however, these models were a) not sufficiently responsive to our phenotyping methods (3.1) and b) contained incompatibilities that would hinder accuracy in the community simulation.

## 4 Results and Discussion

### 4.1 New data fitting method for iterative experimental design

CommPhitting is an original method to resolve extracellular concentrations and phenotype biomass and growth kinetics in microbial communities from simple growth data. The workflow of this model is illustrated in Figure 3. First, the experimental data of a trial is parsed into a standardized format and parameterized into specified growth phenotypes for all provided community members. Second, CommPhitting globally determines the variable and parameter values – namely phenotype abundances, metabolite concentrations, and growth kinetics – that best recapitulate the experimental data while minimizing overfitting. Finally, these predicted values are processed through robust, built-in, functions that consolidate the high-dimensional simulation results into exportable figures, HDF5 files, and spreadsheets, which streamlines biological discovery and improves accessibility to investigators without data science expertise. CommPhitting is available as an open-source Python package in the ModelSEEDpy library ModelSEEDpy ReadTheDocs, and leverages the Optlang optimization module [62] to construct and execute the method.

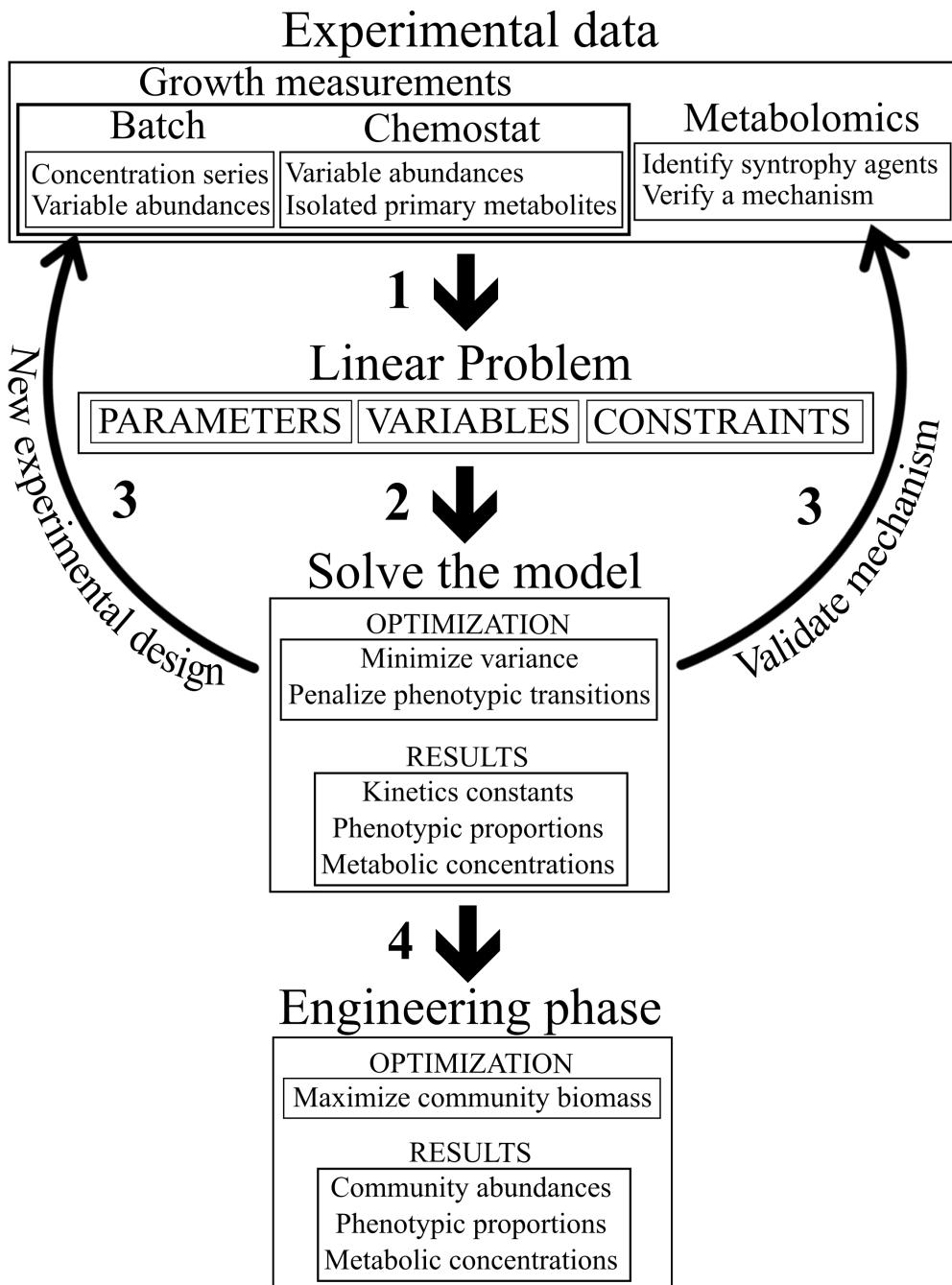


Figure 3: A workflow of the fitting model. **Step 1:** experimental data – from growth and possibly metabolomics measurements – is parsed into a MILP problem that consists of the parameters and variables that are detailed in Table 1 and the constraints that are explained in Section 3. **Step 2:** the 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 gleaning mechanistic insights of microbial communities and then immediately using these insights to rationally design a community with desirable activity.

## 4.2 Simulation insights

CommPhitting recapitulated the growth experimental data of our 2-member community in Figure 4 remarkably well, and even accommodated data anomalies like those in the maltose + 4HB trials of Figure 5, which demonstrates the precision and flexibility of the method and visualizes areas for further experimental improvement. The maltose simulations further demonstrate that the simulation only expresses biologically plausible phenotypes, where *Pseudomonas* expresses only the acetate phenotype and not 4HB in a simulation without this substrate. Both *E. coli* phenotypes of maltose and acetate, in contrast, are dynamically expressed until maltose is exhausted, when it transitions to solely consuming acetate. The timing of this transition, however, differs considerably from *E. coli* as monoculture in Figure 6, which illustrates the effect of environmental competitors on an organism's metabolism.

Several other differences in *E. coli*'s monocultural growth are noteworthy. Firstly, much less time is spent growing, perhaps because acetate accumulates to inhibitory concentrations without other organisms (e.g. *Pseudomonas*) to consume as it is produced. Secondly, following the first observation, *E. coli* more readily expresses its acetate phenotype to presumably detoxify its environment, although it ultimately consumes less acetate as a monoculture as its growth is substantially impeded. Thirdly, the predicted  $k_{cat,acetate}$  is far greater ( $8.89 \frac{1}{hour}$  versus  $0.23 \frac{1}{hour}$ ), which may reflect growth inhibition from the higher acetate concentrations.

## 4.3 Validation

The CommPhitting predictions were validated with targeted and untargeted metabolomics experiments, since CommPhitting predicts various concentrations and the participating phenotypes for all members are fundamentally defined as metabolic profiles. The extracellular concentrations of our archetypal 2-member community was experimentally acquired and compared with the predicted concentrations in Figure 7, where both datasets are converted to biomass over OD to permit an equitable comparison. The Acetate concentration is noteworthy to highlight because it is the critical cross-feeding agent for the community and its concentration is therefore critical to accurately simulate. We further conducted targeted analysis of other potential cross-feeding agents, such as Formate, that we suspect may be cross-feeding agents between the members. We finally conducted a broad, untargeted, analysis of hundreds of metabolites to mitigate potentially biases in analyzing conventional cross-feeding agents. These studies did not detect peripheral cross-feeding from metabolites besides Acetate, which experimentally confirms the simulated observation that Acetate is the primary cross-feeding phenotype.

We further explored engineering methods to validate that Acetate is the central cross-feeding agent. We explored various *E. coli* mutants that cannot absorb Acetate in order to corroborate: a) our other experimental and modeling evidence for Acetate cross-feeding and b) the prediction that the *E. coli* Acetate phenotype is minimally expressed during simulation. We purchased mutant knock-outs from the *E. coli* Genetic Stock Center [63] for knock-outs of each of the four proteins – PoxB, ACS, Pta, and AckA – that directly contribute to Acetate metabolism in MG1655. Assessing mutant growth, as both mono- and co-cultures, revealed that only the  $\Delta$ Pta mutant eliminated Acetate production, and, interestingly exacerbated the growth of *Pseudomonas* when it is grown as a co-culture since it excretes Lactate instead of Acetate [64] and Lactate is a more energetically favorable substrate than Acetate. This affirms that Acetate is the primary mediating compound of *Pseudomonas* growth in our co-culture, where the only mutant that perturbed Acetate excretion was the only mutant that manifestly changed *Pseudomonas*' growth. This augmented growth of *Pseudomonas* in Lactate versus Acetate was confirmed by a dedicated CommPhitting simulation.

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.

Our targeted metabolomics data confirms that Acetate is the primary syntrophic exchange. The other suggested exchange of 2-hydroxybutyric acid is undefined in either of our community members, and is therefore not able to be simulated with these published models without reconstruction in one of the members for a necessary pathway to this byproduct.

## 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

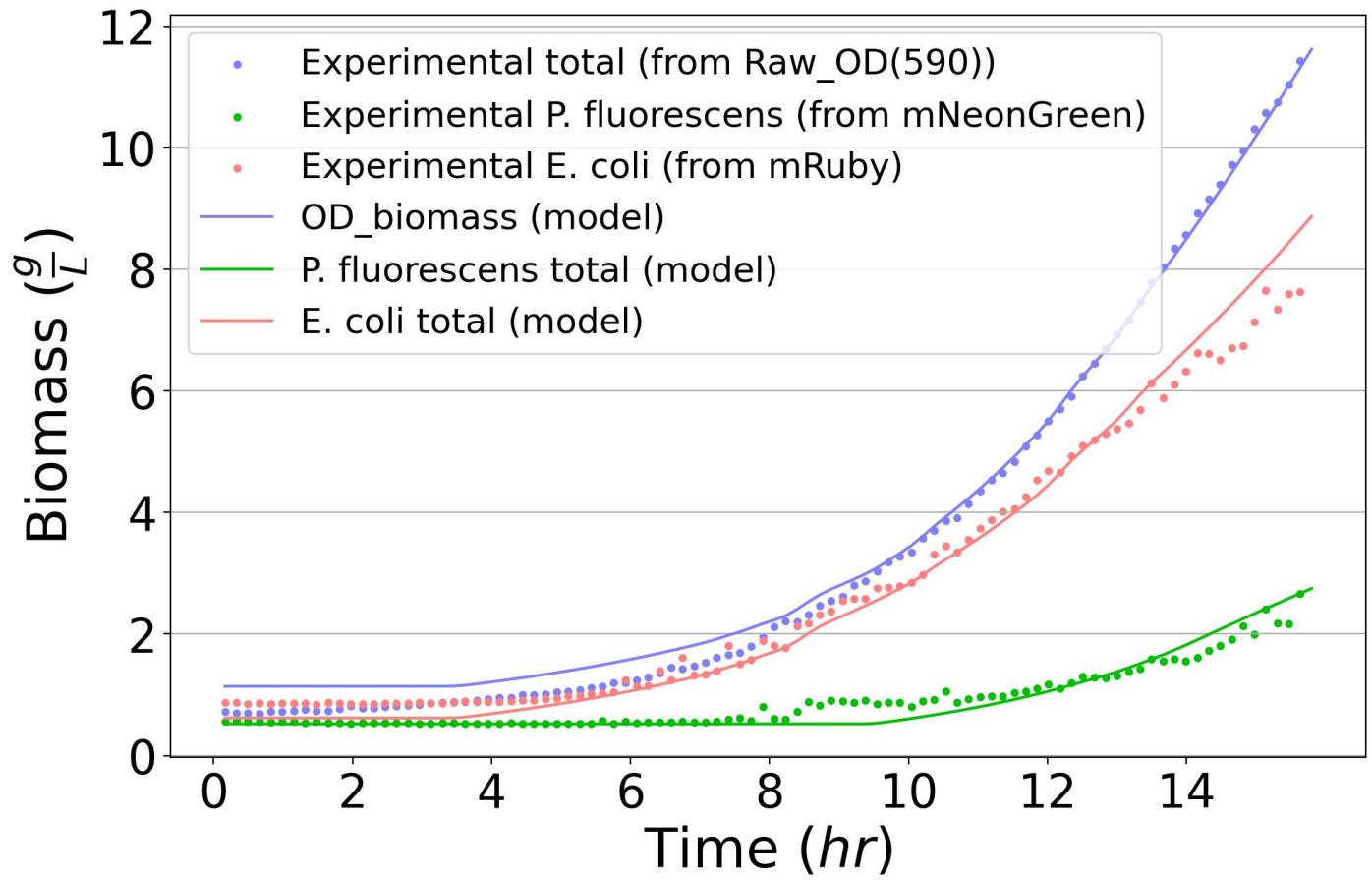


Figure 4: The converted experimental and predicted biomasses of the total community (OD), *E. coli* (RFP), and *P. fluorescens* (GFP) for the coculture experiment on maltose media (Table S1). Tight agreement between the experimental and predicted biomass values improves confidence in the predicted community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

images/maltose\_4HB/coculture\_fit.jpg

Figure 5: The converted experimental and predicted biomasses of the total community (OD), *E. coli* (RFP), and *P. flourescens* (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.

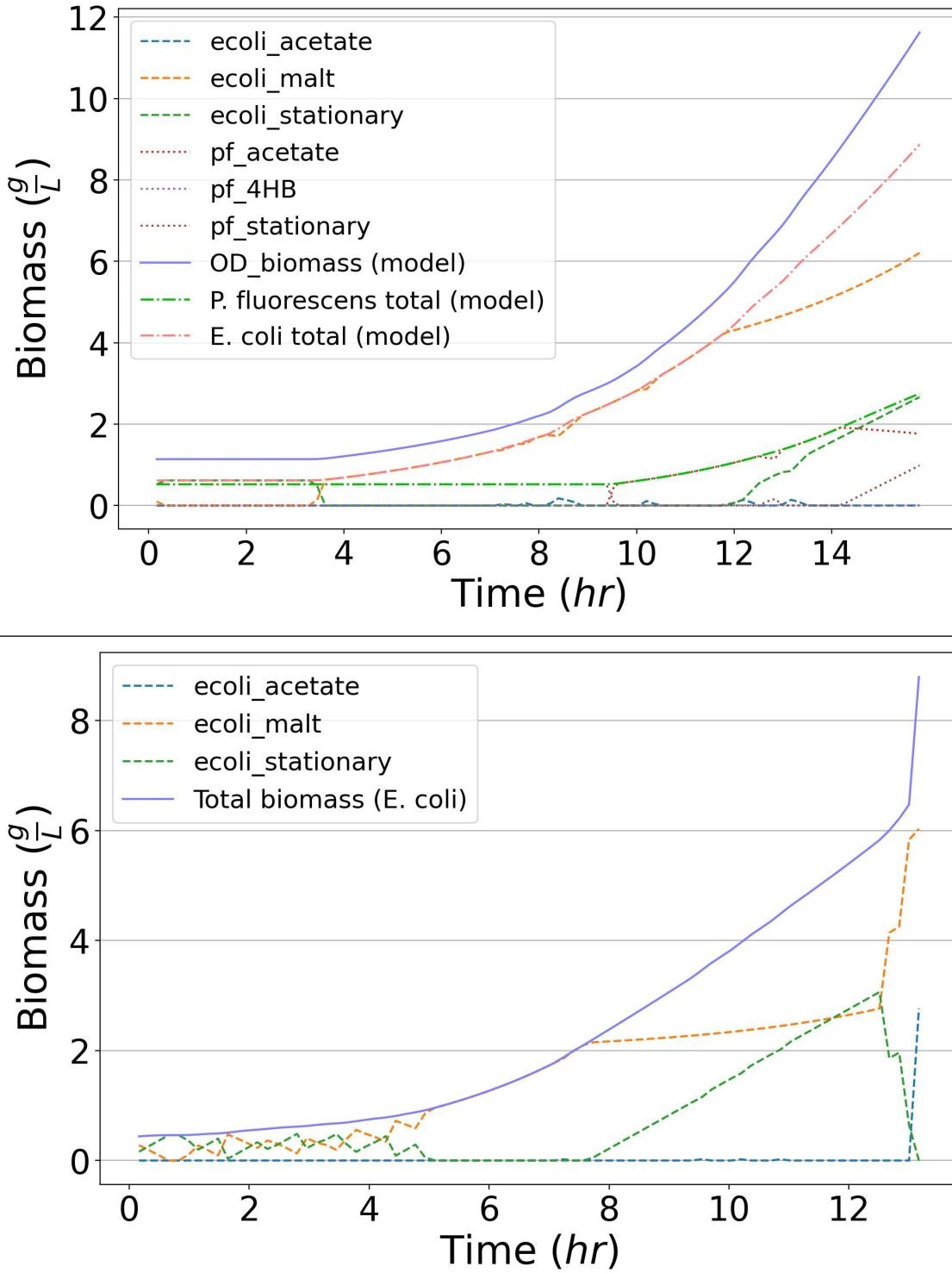


Figure 6: The phenotype abundances for an experiment with 5mM of maltose as the sole carbon sources. The top figure depicts the 1:1 coculture, where *E. coli* outcompeted *P. fluorescens*; however, *P. fluorescens* managed to subsist on the acetate excreta from *E. coli*. The bottom figure depicts a *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.

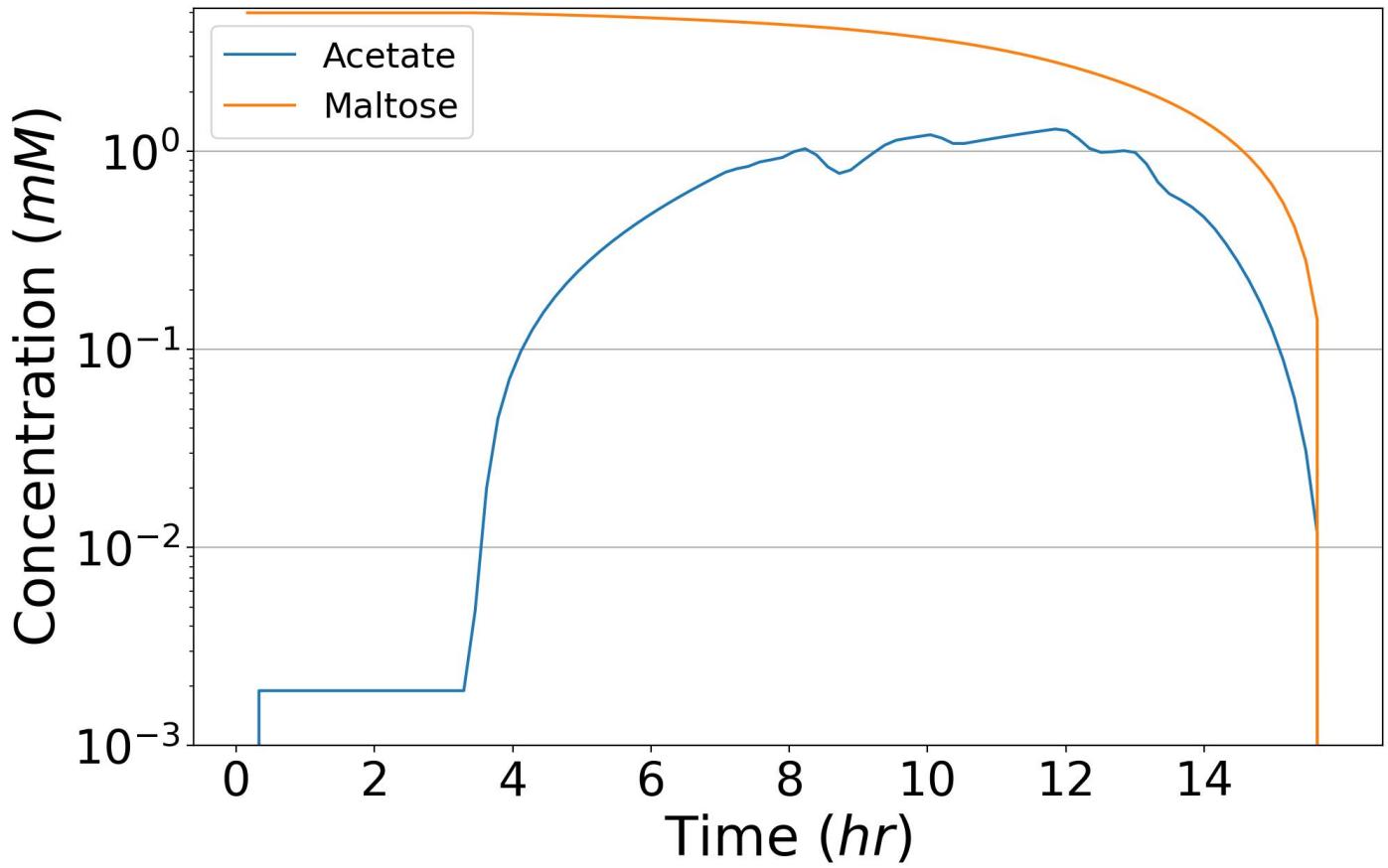


Figure 7: The concentrations of primary carbon sources and the predominate cross-feeding agent acetate from the coculture trial of Figure 6. The dynamic production and subsequent consumption of acetate elicits the unique information that can be derived from CommPhitting simulations. The peak acetate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

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 [65] 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.

community responses to media environments or understanding how engineered organisms with new phenotypes may change community dynamics.

## 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 Gradovalle, 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. Kyrpides, 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] 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>.
- [3] Saeed Shoaei, 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.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] 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>.
- [14] 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.
  - [15] 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.
  - [16] A F White, S A Billings, A M Rossi, K R Hubbert, S M Mudd, E T A Mitchard, W E Dietrich, K Singha, B J Minsley, W S Holbrook, A F Sheehan, M Leopold, S P Anderson, J Taylor Perron, S J Martel, K Singha, R S Anderson, S P Anderson, S P Anderson, G E Tucker, and S L Brantley. A unified initiative to harness Earth's microbiomes. *Science*, 315(6260):2015–2017, 2015.
  - [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. *13C-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] 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.
- [32] 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.
- [33] Xuewen Chen, Ana P. Alonso, Doug K. Allen, Jennifer L. Reed, and Yair Shachar-Hill. Synergy between 13C-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>.
- [34] Keren Yizhak, Tomer Benyaminini, 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.
- [35] 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.
- [36] 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>.
- [37] 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.
- [38] Radhakrishnan Mahadevan, Jeremy S. Edwards, and Francis J. Doyle. Dynamic Flux Balance Analysis of diauxic growth in *Escherichia coli*. *Biophysical Journal*, 83(3):1331–1340, 2002. ISSN 00063495. doi: 10.1016/S0006-3495(02)73903-9. URL [http://dx.doi.org/10.1016/S0006-3495\(02\)73903-9](http://dx.doi.org/10.1016/S0006-3495(02)73903-9).
- [39] 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.
- [40] 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>.
- [41] 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>.
- [42] 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.
- [43] 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.
- [44] 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.

- [45] 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.ymaben.2020.03.001.
- [46] 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.
- [47] Raphy Zarecki, Matthew A. Oberhardt, Keren Yizhak, Allon Wagner, Ella Shtifman Segal, Shiri Freilich, Christopher S. Henry, Uri Gophna, and Eytan Ruppin. Maximal sum of metabolic exchange fluxes outperforms biomass yield as a predictor of growth rate of microorganisms. *PLoS ONE*, 9(5), 5 2014. ISSN 19326203. doi: 10.1371/journal.pone.0098372.
- [48] 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>.
- [49] 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.
- [50] 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.
- [51] 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.
- [52] 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.
- [53] Nathan E. Lewis, Kim K. Hixson, Tom M. Conrad, Joshua A. Lerman, Pep Charusanti, Ashoka D. Polpitiya, Joshua N. Adkins, Gunnar Schramm, Samuel O. Purvine, Daniel Lopez-Ferrer, Karl K. Weitz, Roland Eils, Rainer König, Richard D. Smith, and Bernhard Palsson. Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. *Molecular Systems Biology*, 6, 2010. ISSN 17444292. doi: 10.1038/msb.2010.47.
- [54] 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.
- [55] 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>.
- [56] J. C. Butcher. A history of Runge-Kutta methods. *Applied Numerical Mathematics*, 20:247–260, 1996.
- [57] 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.
- [58] 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.
- [59] 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.
- [60] 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.

- [61] 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.
- [62] 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.
- [63] 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.
- [64] 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.
- [65] 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>.