CommPhitting: A Fitting Model for Investigating the Kinetics and Cross-Feeding of Microbial Communities

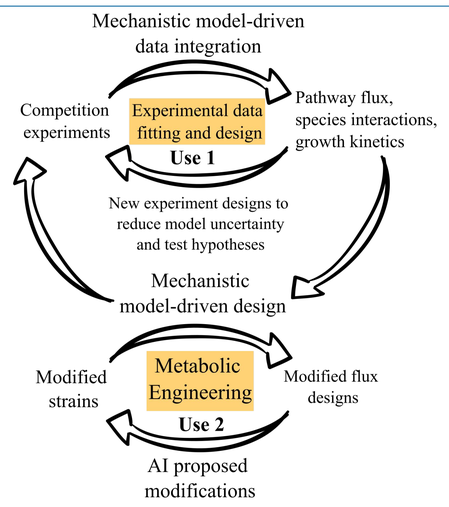
Andrew Freiburger[1] (afreiburger@anl.gov), Jeffrey Dewey,[2] Fatima Foflonker,[1] Gyorgy Babnigg,[2] Dionysios Antonopoulos[2], **Christopher Henry**[1] (chenry@anl.gov)

[1]Data Science and Learning Division, Argonne National Laboratory, Lemont, IL ; [2]Biosciences Division, Argonne National Laboratory, Lemont, IL

COMMUNITY MODELING

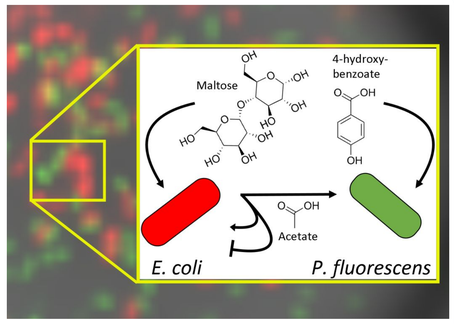
Microbial communities are ubiquitous in the environment as the default mode of prokaryotic existence. These consortia consequently influence medicine, biogeochemical cycles, and are piquing the interest of chemical industry as an efficient source of bioproduction. Community dynamics are therefore fundamentally important to society, yet these exponentially complex interactions remain elusive to experimental and computational methods.

We therefore developed a fitting model (CommPhitting) that integrates experimental growth and -omics data with genome-scale metabolic models to create a parameterized dynamic model of communities that particularly examines the involvement of phenotypes in community behaviors. Our model consists of two use cases: investigation and exploration. The first use case of resolves interaction mechanisms amongst the various metabolic phenotypes of each member, including determining growth kinetics parameters. The second use case applies the fitted model in various conditions to design and predict outcomes from rational engineering of the community and its members. These inter-related, yet distinct, use cases are depicted at right.



EXPERIMENTAL METHODS

An experimentally convenient 2-member community of *P. fluorescens* SBW25 and *E. coli* (minimally modified) MG1655 was assmebled to exemplify our model, for which no studies have been reported. This community exhibits interesting competitive dynamics that include cross-feeding of Acetate from *E. coli* to *P. fluorescens* in Maltose meda (depicted at right). The community strains were purchased from ATC and were transformed with a plasmid to constitutively express either mNeongreen or mRuby2 fluorescent proteins (GFP and RFP) for *P. fluorescens* and *E. coli*, respectively. The transformed cells were prepped and grown on their appropriate carbon source for examination via optical density (OD 590 nm) and fluorescence (544 excitation & 590 emission for *E. coli* and 485 excitation & 535 emission for *P. fluorescens*).



MODELING METHODS

DIMENSIONS

s	A species in the examined community. j.
k	A growth phenotype of species s.
t	An experimental time point.
j	An experimental trial.
i	Extracellular metabolite.

PARAMETERS

$E_{s,t,j}$	The experimental growth signal for a species at instant t in trial j.
$es_{s,k}$	A boolean description of $k \in s$
Δt	The seconds per timestep, which determines the amount of biomass growth per timestep.
$n_{k,i}$	The exchange flux of each metabolite i in each strain k.
$v_{k,t,j}$	The rate constant for growth of strain k at instant t in trial j. This parameter may be either a global value or a Michaelis-Menten flux such as $\frac{v_{max,\ k}}{k_{m,\ k}+c_{t,j,i}}$ that considers the concentration c of i.
$cvct$ & $cvcf$	Conversion coefficients of phenotype biomass to and from the stationary phase, respectively.
bcv_k	The greatest fraction of biomass ($0 < bcv < 1$) of strain k that can transition phenotypes in a timestep.
$cvmin$	The minimal value of variable $cvt_{k,t,j}$.

VARIABLES

EC_k	The conversion coefficient ($0 < EC < 1000$) from parameter $E_{s,t,j}$ into biomass, which is unique for each strain k.
$EB_{s,t,j}$	The computed biomass from each experimental datum, as the product of EC_k & $E_{s,t,j}$.
$b_{k,t,j}$	The predicted biomass from the fitting model.
$EV_{s,t,j}$	The variance between the computed experimental biomass $EB_{s,t,j}$ and the predicted biomass $b_{k,t,j}$.
$c_{t,j,i}$	The concentration of metabolite i at an experimental datum.
$g_{k,t,j}$	The predicted growth rate for each strain at each datum.
$cvt_{k,t,j}$ & $cvf_{k,t,j}$	The quantity of strain k biomass that transitions to and from the stationary phase, respectively, at an experimental datum.

CommPhitting captures biological objects through the dimensions, variables, and parameters that are depicted above and at right.

The first use case of CommPhitting which investigates a community by fitting experimental data, minimizes variance between simulated biomass and derived biomass from the experimental data

$$\sum_{s}^{S}\left(\sum_{t,j}^{T,J}\left(EV_{s,t,j}^{2}\right)\right)-\sum_{t,j,k}^{T,J,K}\left(cvct*cvt_{t,j,k}\right)-\sum_{t,j,k}^{T,J,K}\left(cvcf*cvf_{t,j,k}\right)$$

while minimizing the quantity of phenotypic transitions to mitigate overfitting. The biological processes of the community system are captured through the following system of linear constraints:

$$b_{t,j,k}+\frac{\Delta t}{2}\left(g_{t,j,k}+g_{t+1,j,k}\right)+cvf_{t,j,k}-cvt_{t,j,k}=b_{t+1,j,k} \quad \textbf{a}$$

$$b_{t,j,k}-\sum_{k,s}^{K,S}\left(es_{k,s}*cvf_{t,j,k}\right)+\sum_{k,s}^{K,S}\left(es_{k,s}*cvt_{t,j,k}\right)=b_{t+1,j,k} \quad \textbf{b}$$

$$c_{t,j,i}+\frac{\Delta t}{2}\sum_{k}^{K}\left(n_{i,k}\left(g_{t,j,k}+g_{t+1,j,k}\right)\right)=c_{t+1,j,i} \quad \textbf{c}$$

$$cvt_{t,j,k}\leq bcv*b_{t,j,k}+cvmin \quad \textbf{d}$$

$$EB_{s,t,j}-\sum_{k}^{K}\left(es_{k}*b_{t,j,k}\right)=EV_{s,t,j} \quad \textbf{e}$$

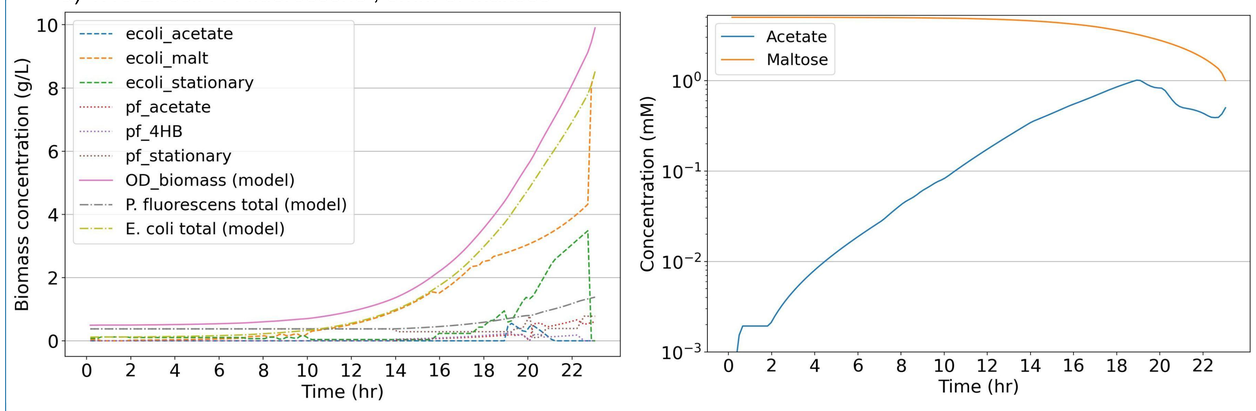
$$E_{s,t,j}*EC=EB_{s,t,j} \quad \textbf{f}$$

These constraints represent: biomass change in the a) non-stationary and b) stationary phenotypes; c) concentration change; d) the minimum biomass fraction that transitions its phenotype; e) biomass variance; and f) the model prediction of biomass, respectively. Constraints a) & c) utilize Heun's method, which is a 2nd-order Runge-Kutta integration method.

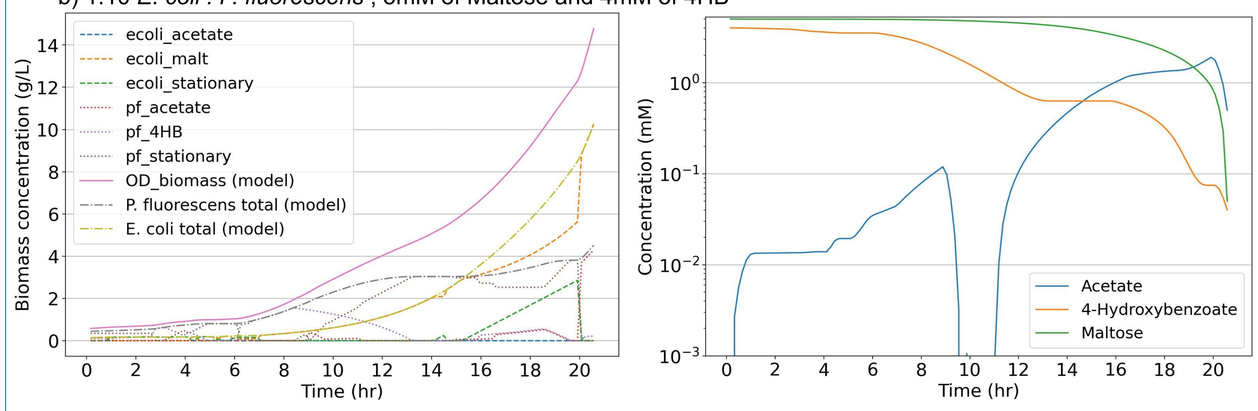
EXPERIMENTAL FIT

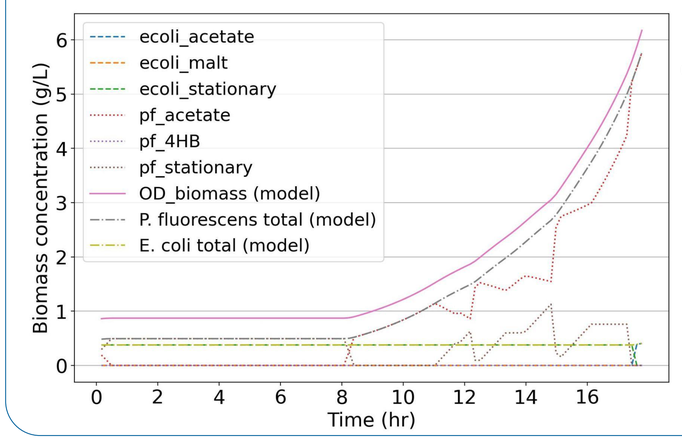
A few batch growth datasets of our 2-member community on **a)** Maltose, **b)** Maltose+4HB, or **c)** Acetate media, coupled with precise Acetate concentrations from metabolomics data, were fitted via CommPhitting to exemplify the first use case of investigating community dynamics from experimental data. A representative set of conditions for each media were simulated, from which results the phenotype biomass and media concentrations are depicted, in the following sets of figures.

a) 1:10 *E. coli : P. fluorescens* , 5mM of Maltose

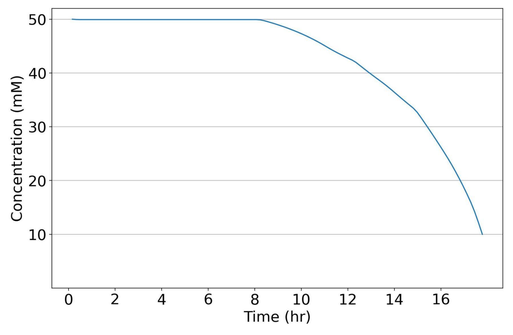


b) 1:10 *E. coli : P. fluorescens* , 5mM of Maltose and 4mM of 4HB

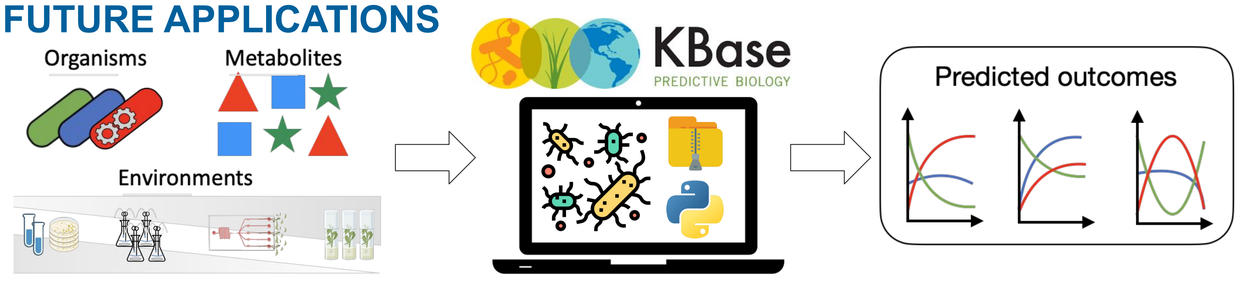




c) 1:1 *E. coli : P. fluorescens* , 50mM of Acetate



FUTURE APPLICATIONS



We are beginning to explore the second use case of our model for exploring community behaviors in various environments. A specific goal is to identify *E. coli* modifications – e.g. new metabolic pathways, gene knockouts, and altered expression profiles – that can augment its competitive fitness with *P. fluorescens* in various media and growth conditions (chemostat/batch/2D plate). These insights may then ultimately steer the community towards bioproduction capabilities, which will have generalizable value for different and larger community systems. We are further diversifying experimental inputs such as BIOLOG data, which will resolve kinetic information for an array of experimental conditions and thereby improve predictions.

We plan to begin exploring larger communities, including a 10-member Plant-Microbe Interface community from our collaborators at ORNL and gradually building towards complete microbiomes. Our data, models, and methods will finally be integrated into a KBase Narrative Application that should foster open-science and the rational design of microbial communities.

U.S. DEPARTMENT OF ENERGY
Argonne National Laboratory is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC.
BERKELEY LAB
JGI JOINT GENOME INSTITUTE
Northwestern University
THE UNIVERSITY OF CHICAGO
Argonne NATIONAL LABORATORY