Metabolite mediated modeling of microbial community dynamics captures emergent behavior more effectively than species-species modeling: Appendices

J.D. Brunner[1,2] and N. Chia[1,2]

[1]Division of Surgical Research, Department of Surgery, Mayo Clinic

[2]Microbiome Program, Center for Individualized Medicine, Mayo Clinic

A The Positive Steady State of the generalized Lotka-Volterra model

A.1 Asymptotic stability analysis

We can re-scale the generalized Lotka-Volterra model with two species to

$$\begin{aligned}\frac{dx_1}{dt} &= r_1 x_1(1 - x_1 + \alpha_{12}x_2) \\ \frac{dx_2}{dt} &= r_2 x_2(1 - x_2 + \alpha_{21}x_1)\end{aligned} \tag{1}$$

in order to simplify notation, and analyze asymptotic behavior of this model by performing straightforward stability analysis on the equilibrium[1]. We see that eq. (1) has equilibrium at $(0,0)$, $(1,0)$, $(0,1)$ and

$$\boldsymbol{x}^* = \left(\frac{1+\alpha_{12}}{1-\alpha_{12}\alpha_{21}}, \frac{1+\alpha_{21}}{1-\alpha_{12}\alpha_{21}}\right)$$

Furthermore, linearization about each of those points reveals that $(0,0)$ is never stable, $(1,0)$ is stable if $\alpha_{21} < -1$, $(0,1)$ is stable if $\alpha_{12} < -1$. Lastly, $\boldsymbol{x}^* \in \mathbb{R}^2_{>0}$ if and only if $\{\alpha_{12} < -1, \alpha_{21} < -1\}$ or $\{-1 < \alpha_{12}, -1 < \alpha_{21}, \alpha_{12}\alpha_{21} < 1\}$ and if $\{\alpha_{12} < -1, \alpha_{21} < -1\}$, then $\boldsymbol{x}^*$ is unstable, and is in fact a saddle point. If $\{-1 < \alpha_{12}, -1 < \alpha_{21}, \alpha_{12}\alpha_{21} < 1\}$, then $\boldsymbol{x}^*$ is stable. All of this can be done through symbolic analysis of the Jacobian matrix evaluated at $\boldsymbol{x}^*$.

We can now characterize the outcomes observed in the paper using the parameters α_{12} and α_{21}:

(a) Coexistence: this is stability of the positive state, and so requires $\{-1 < \alpha_{12}, -1 < \alpha_{21}, \alpha_{12}\alpha_{21} < 1\}$.

(b) Invasion of one species regardless of initial condition: this is stability of one boundary state *and instability of the other*. This requires $\alpha_{12} < -1$, $\alpha_{21} > -1$ or the opposite. If $\alpha_{12} < -1$ then 2 invades 1.

(c) Bi-stability: This is stability of both boundary states, and requires $\{\alpha_{12} < -1, \alpha_{21} < -1\}$.

Interestingly, case (c) is not observed in the data of [2].

The three species model is

$$\frac{dx_1}{dt} = r_1 x_1(1 - x_1 + \alpha_{12}x_2 + \alpha_{13}x_3) \tag{2}$$

$$\frac{dx_2}{dt} = r_2 x_2(1 - x_2 + \alpha_{21}x_1 + \alpha_{23}x_3) \tag{3}$$

$$\frac{dx_3}{dt} = r_3 x_3(1 - x_3 + \alpha_{31}x_1 + \alpha_{32}x_2) \tag{4}$$

and here again we can compute model equilibrium states and stability. There are 8 equilibrium points, corresponding to each qualitative possibility of survival & extinction. Again, the $(0,0,0)$ equilibrium is never stable. There exist simple conditions on the parameters for local stability of all equilibrium points except for the state which represents coexistence of all three microbes. Stability of this last state can, however, be easily evaluated for any given parameters.

We can compute the Jacobian determinant to see that the stability conditions for the double extinction equilibrium points are

- $(1,0,0):\ \alpha_{21} < -1,\ \alpha_{31} < -1$
- $(0,1,0):\ \alpha_{12} < -1,\ \alpha_{32} < -1$
- $(0,0,1):\ \alpha_{13} < -1,\ \alpha_{23} < -1$

Taking advantage of symmetry, we investigate only one of the three single extinction equilibrium, which have the form $\left(\frac{1+\alpha_{12}}{1-\alpha_{12}\alpha_{21}}, \frac{1+\alpha_{21}}{1-\alpha_{12}\alpha_{21}}, 0\right)$. The first two eigenvalues of the Jacobian matrix at these points will follow the two dimensional case, so we have the necessary conditions for stability $\{-1 < \alpha_{12}, -1 < \alpha_{21}, \alpha_{12}\alpha_{21} < 1\}$. This is simply because after extinction of species k, the model is identical to the pair model. While unsurprising, this fact does imply that not all hypothetical combinations of existence and extinction outcomes for pair and trio experiments can be simultaneously explained by the parameters of the generalized Lotka-Volterra model. However, there were no instances in the trio experiments being considered in which such a "smoking gun" scenario was observed.

The third eigenvalue is the value of $r_3(1 - 2x_3 + \alpha_{31}x_1 + \alpha_{32}x_2)$ evaluated at this point, which is

$$\lambda_3 = r_3\left(1 + \alpha_{31}\frac{1+\alpha_{12}}{1-\alpha_{12}\alpha_{21}} + \alpha_{32}\frac{1+\alpha_{21}}{1-\alpha_{12}\alpha_{21}}\right) \tag{5}$$

Clearly if α_{31} and α_{32} are both positive, this state is unstable. The condition for linear stability is

$$\alpha_{31}\frac{1+\alpha_{12}}{1-\alpha_{12}\alpha_{21}} + \alpha_{32}\frac{1+\alpha_{21}}{1-\alpha_{12}\alpha_{21}} < -1. \tag{6}$$

A.2 Lack of limit cycles of the two-species gLV model

We can rule out closed orbits in the two species gLV model using Dulac's criterion. Letting

$$g(x_1, x_2) = \frac{1}{x_1 x_2} \tag{7}$$

we compute

$$\nabla \cdot \left((\frac{dx_1}{dt}, \frac{dx_2}{dt})g(x_1, x_2)\right) = -\frac{r_1}{x_2} - \frac{r_2}{x_1} < 0 \tag{8}$$

for all $x_1, x_2 > 0$. This implies that there are no solution to eq. (1) is a closed orbit in the positive quadrant.

Note the that the standard predator-prey "Lotka-Volterra" model does allow closed orbits. This is because that model does not include the quadratic terms $-r_1x_1^2$ and $-r_2x_2^2$ that appear in eq. (1), and because that model does not enforce the assumption that $r_i > 0$. This can be interpreted as an assumption of infinite carrying capacity of the prey species and exponential decay of the predator.

A.3 Pseudo-Genetic Algorithm

We search for a parameter set to fit qualitative growth behavior by performing a pseudo-genetic algorithm which attempts to minimize

$$M = \sum_{trios}\left(\sum_{\Lambda_i}\lambda^* + p_i\right) \tag{9}$$

where Λ_i is the set of eigenvalues which corresponded to the equilibrium point which matches the experimental outcome of trio i, and $p_i = 0$ if the three pairs of trio i match experimental outcome, and $p_i = 1000$

otherwise. The chromosomes of the genetic search are taken to be the parameter sets, represented as a matrix whose i, j entry contained α_{ij}. We use the rows of this matrix as genes, and so the mating procedure is to choose for each row of the child the row of one or the other parent with even probability.

We describe this as a "pseudo-genetic" algorithm because we are searching over a continuous parameter space. In order to account for this, random mutation of parameters is done by perturbation with a continuous random variable. First, to determine if mutation occurred, we draw a uniform random variable in $(0, 1)$ and mutate if this variable is less than a thresh-hold of 0.2. If mutation occurs, a random matrix whose entries are generated uniformly in $[-0.05, 0.05]$ is added to the matrix representing the parameter set.

A.4 The stochastic generalized Lotka-Volterra model

The model is as follows:

$$X_i(t) = X_i(0) + Y_i^1\left(r_i \int_0^t X_i(s)ds\right) - Y_i^2\left(\hat{r}_i \int_0^t X_i(s)(X_i(s)-1)ds\right) + \sum_{i\neq j} Y_{ij}\left(\hat{\alpha}_{ij}\int_0^t X_i(s)X_j(s)ds\right). \quad (10)$$

Here, $Y(p(t))$ are non-homogeneous Poisson (jump) processes with time-varying propensity $p(t) = \int_0^t f(s)ds$. The new parameters $\hat{r}_i$ and $\hat{\alpha}_{ij}$ depend on the "volume" of the experiment, i.e. the population size scale. Precisely, with a volume N we take α_{ij} as fitted to pair growth experiments and let

$$\hat{r_i} = \frac{r_i}{N} \quad \hat{\alpha}_{ij} = \frac{\alpha_{ij}}{N}$$

Then, as $N \to \infty$, realizations of the stochastic model $\frac{X}{N}$ approach trajectories of the deterministic model [3].

B Stability of equilibrium of QSMI model.

Consider the model for n microbes

$$\frac{d}{dt}x_i = \kappa_i x_i y - d_i x_i \qquad i = 1, .., n \quad (11)$$

$$\frac{d}{dt}y = f_y - d_y y - \sum_{i=1}^{n} \kappa_i x_i y \quad (12)$$

This has equilibrium at $x_i = 0, y = \frac{f_y}{d_y}$, and at $y = \frac{d_i}{\kappa_i}$ for each i, with $x_j > 0$ if and only if $\frac{d_j}{\kappa_j} = \frac{d_i}{\kappa_i}$. The general form of the characteristic equation of the Jacobian matrix about any steady state for this system is

$$\det(J - \lambda I) = (-d_y - \sum_{i=1}^{n} \kappa_i x_i - \lambda)\prod_{i=1}^{n}(\kappa_i y - d_i - \lambda) + \sum_{i=1}^{n} \kappa_i^2 y x_i \prod_{j\neq i}(\kappa_j y - d_j - \lambda) \quad (13)$$

Solving at the extinction steady state $x_i = 0, y = \frac{f_y}{d_y}$, we have eigenvalues $\lambda_{n+1} = -d_y$, $\lambda_j = \frac{\kappa_j f_y}{d_y} - d_j$. Therefore, this state is linearly stable if and only if

$$d_i d_y > k_i f_y \; \forall i \quad (14)$$

Next, for each i let $\mathcal{Q}_i = \left\{ j | \frac{d_j}{\kappa_j} = \frac{d_i}{\kappa_i} \right\}$. Then for each i we have the set of equilibrium defined by

$$\sum_{j \in \mathcal{Q}_i} \kappa_j x_j = \frac{\kappa_i}{d_i} \left(f_y - d_y \frac{d_i}{\kappa_i} \right) \tag{15}$$

and $x_l = 0$ if $l \notin \mathcal{Q}_i$. The characteristic equation becomes

$$\det(J - \lambda I) = - \left[\lambda^2 + \lambda d_y + \lambda \frac{\kappa_i}{d_i} \left(f_y - d_y \frac{d_i}{\kappa_i} \right) + \sum_{j \in \mathcal{Q}_i} \kappa_j^2 x_j \right] \lambda^{|\mathcal{Q}_i|-1} \prod_{l \notin \mathcal{Q}_i} \left(\kappa_l \frac{d_i}{\kappa_i} - d_l - \lambda \right) \tag{16}$$

First, we see that if $\frac{d_i}{\kappa_i} \neq \min_{l=1,\ldots,n} \left\{ \frac{d_l}{\kappa_l} \right\}$, then this is unstable. If we do have the minimum $\frac{d_i}{\kappa_i}$, then the remaining nonzero eigenvalues are

$$\lambda_{\pm} = \frac{1}{2} \left[-B \pm \left(B^2 + 4A \right)^{1/2} \right] \tag{17}$$

where $B = d_y + \frac{\kappa_i}{d_i} \left(f_y - d_y \frac{d_i}{\kappa_i} \right) > 0$ and $A = \sum_{j \in \mathcal{Q}_i} \kappa_j^2 x_j > 0$. These both then have negative real part, implying that the hyperplane of solutions is attracting (note that if $|\mathcal{Q}_i| = 1$, this implies a linearly stable equilibrium point).

Next, we consider the two species cross-feeding or cross-poisoning model:

$$\frac{d}{dt} x_1 = \kappa_{11} x_1 y_1 - d_1 x_1 + \psi_{12} x_1 y_2 \tag{18}$$

$$\frac{d}{dt} x_2 = \kappa_{21} x_2 y_1 - d_2 x_2 \tag{19}$$

$$\frac{d}{dt} y_1 = f_1 - d_1^* y_1 - \kappa_{11} x_1 y_1 - \kappa_{21} x_2 y_1 \tag{20}$$

$$\frac{d}{dt} y_2 = \kappa_{21} x_2 y_1 - d_2^* y_2 - \kappa_{12} x_1 y_2 \tag{21}$$

Here, conditions for stability of the double extinction state are the same as above. Suppose $d_2 \kappa_{11} > d_1 \kappa_{21}$, so that if $\psi_{12} = 0$, this model behaves as the single metabolite model, and the state with $x_2 = 0$, $x_1 > 0$ is stable. We are interested in causing the opposite extinction. That steady state is

$$(x_1, x_2, y_1, y_2) = \left(0, \frac{1}{d_2} \left(f_1 - d_1^* \frac{d_2}{\kappa_{21}} \right), \frac{d_2}{\kappa_{21}}, \frac{1}{d_2^*} \left(f_1 - d_1^* \frac{d_2}{\kappa_{21}} \right) \right) \tag{22}$$

and the eigenvalues of the Jacobian matrix at this state can be computed symbolically, and the relavant eigenvalue is

$$\lambda = \kappa_{11} \frac{d_2}{\kappa_{21}} - d_1 + \psi_{12} \left(f_1 - d_1^* \frac{d_2}{\kappa_{21}} \right) \tag{23}$$

giving a condition for stability on ψ_{12} that can be achieved.

For coexistence, we will assume that the initial model with $\psi_{12} = 0$ has survival of x_2, so $d_2 \kappa_{11} < d_1 \kappa_{21}$. Then we simply repeat the argument above to *destabilize* the equilibrium point, causing $\lambda > 0$. Then, all three of the double extinction and both single extinction equilibrium are unstable. We can conclude coexistence.

C Algorithmically created metabolite mediated models - Creation and Complexity

To estimate the complexity of metabolite mediated models which recapitulate some experimental outcome, we randomly generate experimental outcomes and algorithmically build models to recapitulate these results. The algorithm which builds the model follows the following steps:

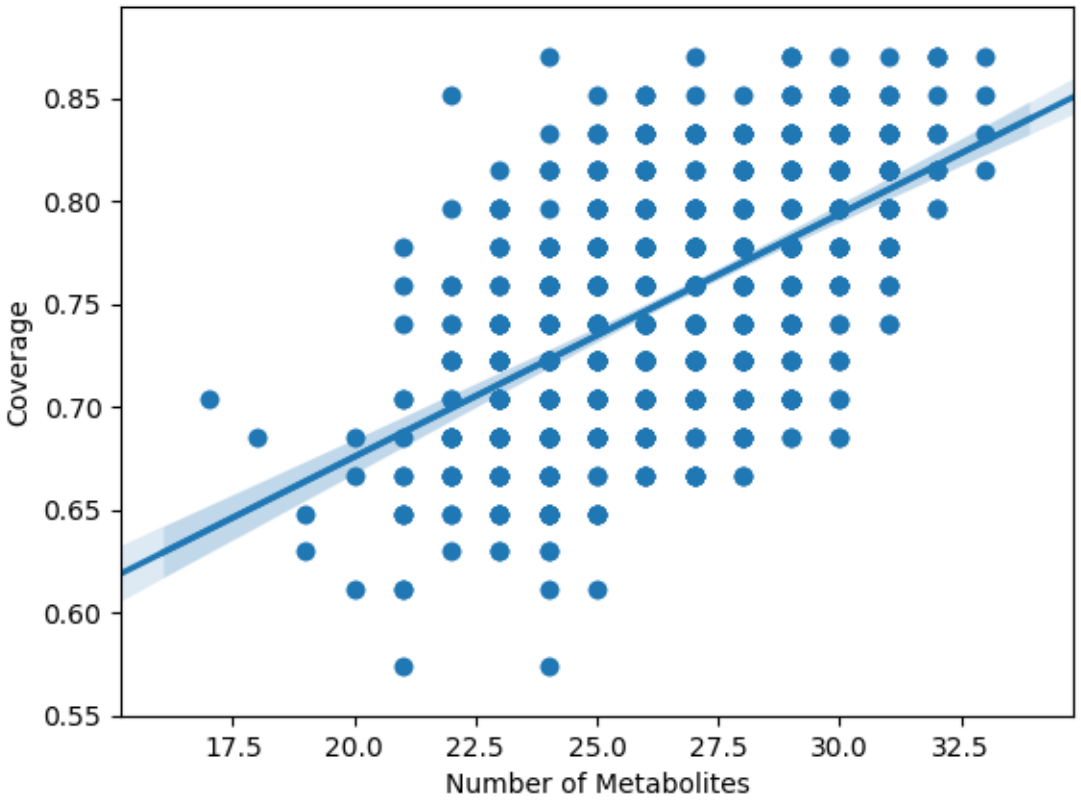

Figure 1: *Number of metabolites in model vs. coverage of randomly generated result set*

(1) Randomly choose growth parameters for a single-metabolite model.

(2) Add cross talk pathways so that pair experiments are recapitulated.

(3) Add cross talk pathways so that trio experiments are recapitulated.

Notice that step (1) causes some set of pair predictions which must be adjusted by step (2), and this in turns causes some set of trio predictions which must be corrected by step (3). Furthermore, a model is determined by the choice in step (1). Thus, optimization can be done over set of parameters chosen in step (1). A python script for generating random outcomes and corresponding models can be found in the supplemental code repository, labeled `random_metab.py`.

For any set of randomly generated outcomes, we record the coverage and number of metabolites in the best model we create. Histograms of the results are shown in the manuscript, in Figure 7. Below, in fig. 1, we show a scatter plot of number of molecules vs. coverage for these same models.

References

[1] Edelstein-Keshet L. 2005 *Mathematical Models in Biology* vol. 46 *Classics in Applied Mathematics*. SIAM.

[2] Friedman J, Higgins LM, Gore J. 2017 Community structure follows simple assembly rules in microbial microcosms. *Nature Ecology &Amp; Evolution* **1**, 0109 EP –.

[3] Kurtz TG. 1972 The Relationship between Stochastic and Determinisic Models for Chemical Reactions. *The Journal of Chemical Physics* **57**.