Details of the Model of Symbiosis among Randomly Generated Networks

To numerically examine how complicated forms of symbiosis can emerge because of cell-level optimization for each cell's growth, the following algorithm is employed (Fig. A). To solve the ODE, the 4th order Runge-Kutta method is implemented using C++. Here, note that the diffusion coefficient of the nutrient, $D_0^{(\alpha)}$, is fixed to 1.0 for all cell species α throughout the simulation.

Step 1. Prepare a "species pool" that contains $N = 50$ cell species with different, randomly generated reaction networks. (see the next subsection for details).

Step 2. Select the cell species 0 whose growth rate in isolation, $\mu_{\text{iso}}^{(0)}$, is highest among the N species. First, only cell species 0 with the fastest growth in isolation is placed in an environment with fixed V_{env}, S_{env}, and $D_S^{(\text{env})} = 20$. Then, the diffusion coefficient of each non-nutrient metabolite i is chosen from $D_i^{(0)} = 0.001, 0.002, 0.004, \cdots, 1.024$ to optimize the growth rate of the cell species. If the optimal diffusion coefficients of all non-nutrient metabolites, $D_i^{(0)}$, equal zero (i.e., if species 0 has no leak-advantage chemicals), the species pool is re-generated (otherwise, species 0 dominates and coexistence of multiple specie does not occur).

Step 3. Cell-level adaptation of the diffusion coefficient for the invading species. Each cell species with the other $N - 1$ networks is examined, one by one, to determine whether they can invade (i.e., grow and coexist) or not. This new cell species β initially constitutes a proportion $p_\beta = 0.0025$ of the population, and the diffusion coefficient of each non-nutrient metabolite i is selected from $D_i^{(\beta)} = 0.001, 0.002, 0.004, \cdots, 1.024$ so that the growth rate is maximized in the environmental condition with pre-existing cell species (initially only the fastest growing species, given at Step 2, exists). (In other words, at the onset of the invasion, the invading species' population is small enough not to affect the environmental conditions.) Next, the population dynamics of $\dot{p}_\alpha = (\mu_\alpha - \bar{\mu})p_\alpha$ are computed for a period $T = 1.8 \times 10^6$ that is long enough for populations to reach a stationary distribution with fixed $\mathbf{D}^{(\alpha)}$. Here, if the population ratio p_α of a cell species α is below $p_{\min} \equiv 0.00025$, the cell species α becomes extinct.

Step 4. Cell-level adaptation of the diffusion coefficient for all coexisting species after invasion. Each of the surviving species independently changes its diffusion coefficient by $\Delta D_{\text{Leak}} = 0.001$ to optimize its growth over a sufficiently long period T (see the following subsection and Fig. B for details). After that, the population dynamics are calculated over a sufficiently long period T without changing their diffusion coefficients $\mathbf{D}^{(\alpha)}$.

Step 5. Repeat steps 3 and 4 (i.e., invasion and adaptation) until the population of coexisting species converges.

In this model, it is assumed that the timescales of the following three procedures are distinct: "the dynamical change of the environmental concentration and population" $\ll$ "the adaptation of the diffusion coefficients of each cell species" $\ll$ "the invasion of new species." Assuming this timescale separation, the above algorithm is justifiable as a model of the invasion of the microbial ecological system by additional species.
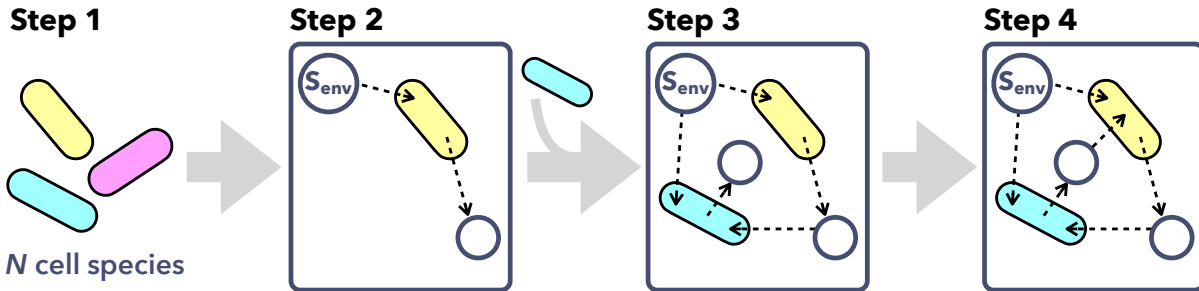

Fig. A. Schematic illustration of the model of symbiosis among randomly generated networks.

Preparing a "species pool" of randomly generated networks in Step 1. We randomly generated catalytic reaction networks with n kinds of chemical components as shown below. Subsequently, we selected those networks that provide stable growth, namely, those whose growth rate μ is higher than 5×10^{-5}.

Each catalytic network consists of n chemicals and ρn two-body catalytic reactions with $\rho = 2$. Chemicals 1 to n_{enzyme} are defined as enzymes and they can be a catalyst or product of catalytic reactions but cannot be a substrate of reactions, while the rest of the chemicals (chemicals $n_{\text{enzyme}} + 1$ to $n - 1$ and nutrient chemical 0) are defined as metabolites that can be substrates or products but cannot catalyze any reactions.

The cellular growth rate is determined by the synthesis rate of a single chemical that contributes to biomass, where metabolite $n - 1$ and enzyme 1 are defined as the precursor for biomass synthesis and enzyme for biomass synthesis, respectively. The cellular growth rate $\mu(\mathbf{x})$ is thus given by the rate of biomass synthesis, $\mu(\mathbf{x}) = k_{\text{BM}} x_{n-1} x_1$.

Then, the time course of the concentration of chemical i $(i = 1, 2, \cdots, n-2)$ is expressed as

$$\frac{d}{dt}x_i = \underbrace{\sum_{j,l=0}^{n-1} P(j,i,l)x_j x_l - \sum_{j,l=0}^{n-1} P(i,j,l)x_i x_l}_{F_i(\mathbf{x})} + \underbrace{D_i(x_i^{(\text{env})} - x_i)}_{f_i(\mathbf{x}; D_i, x_i^{(\text{env})})} - \mu(\mathbf{x})x_i,$$

where $P(i,j,l)$ takes a positive value (that equals the reaction rate constant) if reaction $i + l \to j + l$ occurs; otherwise, it equals 0. For chemicals 0 (nutrient) and $n-1$ (biomass precursor), evolution can be written as

$$\frac{d}{dt}x_0 = -\sum_{j,l=0}^{n-1} P(0,j,l)x_0 x_l + D_S(x_0^{(\text{env})} - x_0) - \mu(\mathbf{x})x_0,$$

$$\frac{d}{dt}x_{n-1} = \sum_{j,l=0}^{n-1} P(j,n-1,l)x_j x_l - \sum_{j,l=0}^{n-1} P(n-1,j,l)x_{n-1}x_l - k_{\text{BM}}x_{n-1}x_1 + D_{n-1}(x_{n-1}^{(\text{env})} - x_{n-1}) - \mu(\mathbf{x})x_{n-1}.$$

Each reaction network consists of ρn catalytic reactions (with the rate constant 1), randomly chosen such that an enzyme l (i.e., chemical $1, 2, \cdots, n_{\text{enzyme}}$) catalyzes the conversion from a metabolite i (i.e., chemical $0, n_{\text{enzyme}}+1, n_{\text{enzyme}}+2, \cdots, n-1$) to another chemical j (a different enzyme or a different metabolite, which is not a nutrient 0; i.e., $j \neq 0, i, l$). Because of $j \neq l$, autocatalytic reactions are prohibited. The network can contain multiple identical reactions, and if so, the corresponding $P(i,j,l)$ takes the value 2 (or more) such that the sum of $P(i,j,l)$ is ρn, i.e., the corresponding reaction rate is doubled (multiplied). The number of non-zero elements in $P(i,j,l)$ is $\leq \rho n$.

Cell-level adaptation dynamics for diffusion coefficients in Step 4. To numerically compute the cell-level adaptation processes of diffusion coefficients $\mathbf{D}^{(\alpha)}$ (leakage and uptake) by each cell species α, we considered the following procedures where each cell species simultaneously increases or decreases the diffusion coefficient of one chemical, say $D_i^{(\alpha)}$, at regular time intervals τ such that its own growth rate increases.

If the growth rate of cell species α increases after time τ has passed since $D_i^{(\alpha)}$ is increased (decreased), then $D_i^{(\alpha)}$ is increased (decreased) again; while if the growth rate of cell species α decreases after time τ has passed since $D_i^{(\alpha)}$ is increased (decreased), then $D_i^{(\alpha)}$ is decreased (increased) (see also Fig. B):

$$D_i^{(\alpha)}(t+\tau+dt) = D_i^{(\alpha)}(t+\tau) + \Delta D_i^{(\alpha)}(t+\tau)$$
$$= D_i^{(\alpha)}(t+\tau) + \Delta D_{\text{Leak}}\text{Sgn}\left[\mu^{(\alpha)}(t+\tau) - \mu^{(\alpha)}(t)\right]\text{Sgn}\left[\Delta D_i^{(\alpha)}(t)\right],$$

where the initial value is set to $\Delta D_i^{(\alpha)}(t=0) = \Delta D_{\text{Leak}} = 0.001$, and τ is chosen as $\sim 10/\mu$.

As the diffusion coefficients of all metabolites change adaptively, the index i of the selected chemical is randomly switched to another chemical i' at a constant probability $p_{\text{change}} = 0.1$, if the diffusion coefficient $D_i^{(\alpha)}$ of chemical i has already been changed more than twice. If $D_i^{(\alpha)}(t+\tau) = 0$ and $\Delta D_i^{(\alpha)}(t+\tau) < 0$, then $D_i^{(\alpha)}(t+\tau+dt)$ is maintained at 0, and the index i of the selected chemical is randomly switched to another chemical i'.
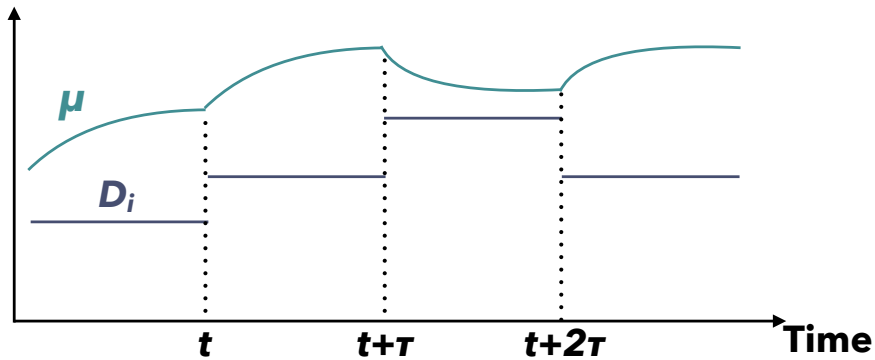

Fig. B. Schematic illustration of cell-level adaptation dynamics of diffusion coefficients. In this example of the time series of diffusion coefficient $D_i^{(\alpha)}$ (dark blue) and $\mu^{(\alpha)}$ (green), $D_i^{(\alpha)}$ is increased at time $t+\tau$ as $D_i^{(\alpha)}(t+\tau+dt) = D_i^{(\alpha)}(t+\tau) + \Delta D_{\text{Leak}}$ because the increase in $D_i^{(\alpha)}$ at time t increases the growth rate as $\mu^{(\alpha)}(t+\tau) > \mu^{(\alpha)}(t)$. Likewise, $D_i^{(\alpha)}$ is decreased at time $t+2\tau$ because the increase in $D_i^{(\alpha)}$ at time $t+\tau$ decreases the growth rate as $\mu^{(\alpha)}(t+2\tau) < \mu^{(\alpha)}(t+\tau)$.

Note that one can consider another way to numerically simulate the cell-level adaptation of the diffusion coefficients. For instance, the growth change due to small additional leakage/uptake can be computed using the inverse matrix of the Jacobi matrix at a steady state [see Eq. [1] in Yamagishi *et al.* (2020) for details], under the assumption that such small changes in diffusion coefficients do not affect the environmental concentrations $\mathbf{x}^{(\text{env})}$. This "more rigorous" procedure for the cell-level adaptation, however, requires much more computing time, while it reproduces qualitatively identical results (Fig. C).
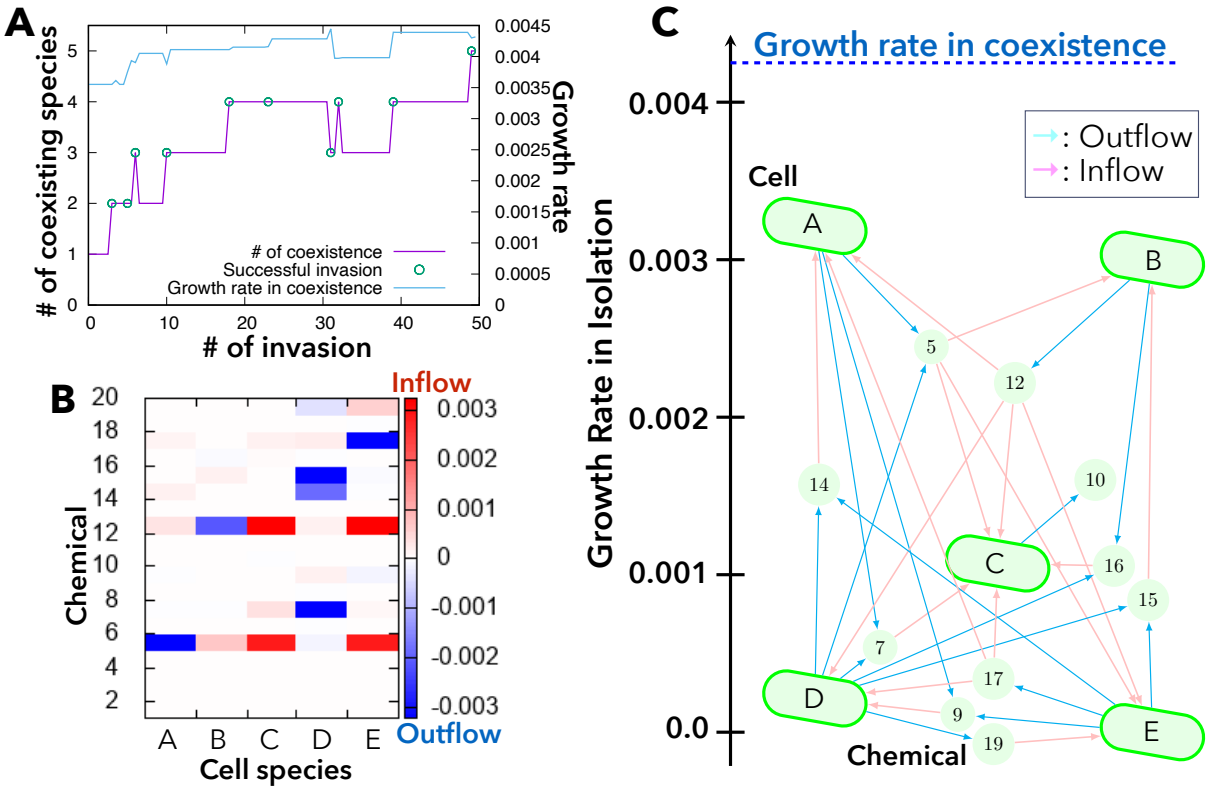
Jumpei F. Yamagishi, Nen Saito, Kunihiko Kaneko

Fig. C. An example of symbiosis when the adaptation dynamics with the Jacobi matrix is adopted. (A) Time series of the number of coexisting species by successful invasions of new species and the growth rate of cell species in coexistence. (B) Plot of leakage (blue) and uptake (red) fluxes of non-nutrient chemicals from each cell species, A-E. (C) Structure of metabolic exchange among five species that originally have different growth rates in isolation. Cyan and pink arrows indicate the leakage and uptake of each chemical component, respectively. Symbiosis among multiple cell species raises the growth rate to reach $\mu_{\mathrm{symbiosis}}$ (as indicated on the top), which is higher than the growth rate of each cell species in isolation, μ_{iso}. In the numerical simulation, the parameters were set to $n = 20$, $S_{\mathrm{env}} = 0.03$, $V_{\mathrm{env}} = 3$, $D_S^{(\mathrm{env})} = 20$, $D_S = 1$, $R_{\mathrm{deg}} = 5 \times 10^{-5}$, $n_{\mathrm{enzyme}} = n/5$.