I. A SIMPLE CASE - MUTUALLY CATALYTIC MOLECULES -

A. Reactions

Let us consider two types of cells. One type, denoted by $A$, consists of molecule species $X$ and $Y$. The other type, $B$, consists of molecule species $X$ and $Z$. The molecule species mutually catalyze the replication of each other to form a minimal hypercycle as follows:

$$X + Y + S_X \rightarrow 2X + Y, \quad Y + X + S_Y \rightarrow 2Y + X$$

and

$$X + Z + S_X \rightarrow 2X + Z, \quad Z + X + S_Z \rightarrow 2Z + X.$$ 

We denote the intrinsic catalytic activities of $S$ and $X$ for cell type $A$. The rates of replicating $X$ is the concentration and $i$ is the number of molecule species $i$. The rates of replicating $Y$ and $Z$ are, respectively, given by

$$F_A^X = V_A f_A^X s_X, \quad F_B^X = V_B f_B^X s_X,$$

respectively, for cell type $A$ and $B$. Here, $V_A$ and $V_B$ are the volumes of cell types $A$ and $B$, respectively, and $f_A^X = c_Y x_A^X x_Y^A, f_B^X = c_Z x_B^X x_Z^B$, where $x_i^I = N_i^I / V_i (i = X, Y, Z; I = A, B)$ and $N_i^I$ is the number of molecule species $i$. The rates of replicating $Y$ and $Z$ are, respectively, given by

$$F_A^Y = V_A f_A^Y s_Y, \quad F_B^Z = V_B f_B^Z s_Z.$$ 

Here, $f_A^X = c_Y x_A^X x_Y^A, f_B^B = c_Z x_B^X x_Z^B$. $s_i = S_i / S_i^0 (i = X, Y, Z)$ is the normalized concentration of the resource where $S_i$ is the concentration and $S_i^0$ is introduced to normalize $S_i$ to one when $S_i = S_i^0$.

The rate equations of cell-types $A$ and $B$ for molecule species $X$ are written as,

$$\frac{dN_A^X}{dt} = F_A^X = V_A f_A^X s_X, \quad \frac{dN_B^X}{dt} = F_B^X = V_B f_B^X s_X,$$

and, for $Y$ and $Z$,

$$\frac{dN_A^Y}{dt} = F_A^Y = V_A f_A^Y s_Y, \quad \frac{dN_B^Z}{dt} = F_B^Z = V_B f_B^Z s_Z.$$ 

The dynamics of resources $S_X$, $S_Y$, and $S_Z$ are respectively written as

$$\frac{dS_X}{dt} = -(V_A f_A^X + V_B f_B^X) s_X + D S_X^0 (1 - s_X),$$

$$\frac{dS_Y}{dt} = -V_A f_A^Y s_Y + D S_Y^0 (1 - s_Y),$$

$$\frac{dS_Z}{dt} = -V_B f_B^Z s_Z + D S_Z^0 (1 - s_Z).$$ 

In the steady state, the value of $s_X$, $s_X$, is written as,

$$s_X = \frac{DS_X^0}{(V_A f_A^X + V_B f_B^X) + DS_X^0}.$$
In Fig. 1, we show \( s_X \) as a function of \( D \) with a set of parameters. For large \( D \), \( s_X \rightarrow 1 \). As \( D \) is decreased, \( s_X \) starts to decrease and deviates from one. For smaller \( D \), \( s_X \) decreases linearly as \( \alpha D \).

Similarly, for resources \( S_Y \) and \( S_Z \) the steady state values are written as

\[
\begin{align*}
\tilde{s}_Y &= \frac{DS_Y^0}{V_A f^A_1 + DS_Y^0}, \\
\tilde{s}_Z &= \frac{DS_Z^0}{V_B f^B_1 + DS_Z^0}.
\end{align*}
\]

(7)

The dynamics of each molecule species changes with the diffusion constant \( D \). When \( D \) is sufficiently large, \( s_X \) is approximately one. As \( D \) decreases and the resource is limited, \( s_X \) deviates from one, and the maximum inflow rate of the resource is given by \( J_X = DS_X^0\left(1 - \tilde{s}_X\right) \). In each range of \( D \), we refer to \( (X - i) : S_X \) is sufficiently supplied, and \( (X - ii) : S_X \) is limited.

For each range, the right-hand-sides of the rate equations (1) can be written as

\[
(F_X^A, F_X^B) = \begin{cases} 
(V_A f_X^A, V_B f_X^B) & \text{for } (X - i) \\
(r_A j_X, r_B j_X) & \text{for } (X - ii)
\end{cases}
\]

where \( r_A \) and \( r_B = 1 - r_A \) are the ratios of the resource being distributed to cell types \( A \) and \( B \).

For the resources \( S_Y \) and \( S_Z \), we assume the range changes simultaneously for both of \( Y \) and \( Z \) because, otherwise, either type with a limited resource will vanish; we refer to \( (YZ - i) : S_Y \text{ and } S_Z \text{ are sufficiently supplied, and } (YZ - ii) : S_Y \text{ and } S_Z \text{ are limited. For each range, the right-hand-sides of the rate equations (2) can be written as}

\[
F_Y^i = \begin{cases} 
V_A f_Y^i & \text{for } (YZ - i) \\
J_Y & \text{for } (YZ - ii)
\end{cases}
\]

where \( J_Y = DS_Y^0(1 - \tilde{s}_Y) \) and \( J_Z = DS_Z^0(1 - \tilde{s}_Z) \). As will be shown in subsequent subsections, we investigate four cases for the conditions \( (X, Y, Z) = (i, i), (i, ii), (ii, i) \) and \( (ii, ii) \).

The volume \( V_A, V_B \) of each cell increases in proportional to the numbers of molecules as

\[
V_A = \frac{1}{R} (N_X^A + N_Y^A), \quad V_B = \frac{1}{R} (N_X^B + N_Z^B),
\]

where \( R \) is a constant.

Then the increase in volumes is given by

\[
\frac{dV_A}{dt} = \frac{1}{R} (F_X^A + F_Y^A), \quad \frac{dV_B}{dt} = \frac{1}{R} (F_X^B + F_Z^B),
\]
Let us suppose that a cell divides when its volume exceeds a certain critical value and that the total volume of all cells is restricted by some constant \( \sum V_i = V_T = \text{const} \). Then the volume fraction of each type \( A \) and \( B \) follows

\[
\frac{dv_A}{dt} = \frac{1}{RV_T} (F_A^I + F_B^I) - \frac{v_A}{R} \sum F_I^T
\]
\[
\frac{dv_B}{dt} = \frac{1}{RV_T} (F_B^I + F_Z^I) - \frac{v_B}{R} \sum F_I^T
\]

where \( v_I = V_i/V_T (I = A, B) \), \( F_1 = \sum F_I^T \).

**B. The case in which all the resources are sufficiently available: condition \( (X, YZ) = (i, i) \)**

For the condition \( (X, YZ) = (i, i) \), when all resources are sufficiently available, the increases in molecule species are written as

\[
F_A^I = V_A f_A^I, \quad F_Y^I = V_A f_Y^I, \quad F_X^I = V_B f_X^I, \quad F_Z^I = V_B f_Z^I
\]

where

\[
f_A^I = c_Y x_A x_Y, \quad f_X^I = c_Z x_X x_Z x_Y, \quad f_Y^I = c_X x_A x_Y, \quad f_Z^I = c_X x_X x_Y x_Z
\]

and \( x_i^I = N_i^I / V_I (I = A, B, i = X, Y, Z) \). Thus,

\[
\frac{dv_A}{dt} = \frac{f_A}{R} v_A - \frac{v_A}{R} \sum f_I^T
\]
\[
\frac{dv_B}{dt} = \frac{f_B}{R} v_B - \frac{v_B}{R} \sum f_I^T
\]

where \( f_A = f_X^I + f_Y^I, f_B = f_Z^I + f_Z^I \). The stationary states give

\[
v_A f_A^I = v_A (v_A f_A^I + v_B f_B^I), \quad v_B f_B^I = v_B (v_A f_A^I + v_B f_B^I).
\]

The solutions to these equations are \( (v_A, v_B) = (1, 0) \) or \( (0, 1) \).

By writing \( v_A = \delta \) and \( v_B = 1 - \delta \), linearization around \( (v_A, v_B) = (0, 1) \) gives

\[
\frac{d\delta}{dt} = \frac{f_A}{R} \delta - \frac{\delta}{R} (f_A^I + (1 - \delta) f_B^I)
\]
\[
= \frac{\delta}{R} (f_A - f_B).
\]

Thus, the fixed point \( (v_A, v_B) = (0, 1) \) is stable when \( f_B > f_A \). On the other hand, \( (v_A, v_B) = (1, 0) \) is stable when \( f_A > f_B \). Hence, the fittest cell type, i.e., that with larger \( f_i (i = A \text{ or } B) \) dominates the population.

**C. The case \( S_X \) is limited: the condition \( (X, YZ) = (ii, i) \)**

For the condition \( (X, YZ) = (ii, i) \), when \( S_X \) is limited,

\[
F_A^I = r_A J_X, \quad F_X^I = r_B J_X, \quad F_Y^I = V_A f_Y^I, \quad F_Z^I = V_B f_Z^I
\]

where \( J_X \) is the maximum inflow of \( S_X \) and \( r_A \) and \( r_B = (1 - r_A) \) are the ratios of \( S_X \) distributed into \( A \) and \( B \). Then,

\[
\frac{dv_A}{dt} = \frac{1}{R} (r_A J_X + f_Y^I v_A) - \frac{v_A}{R} (J_X + f_Y^I v_A + f_Z^I v_B),
\]
\[
\frac{dv_B}{dt} = \frac{1}{R} (r_B J_X + f_Z^I v_B) - \frac{v_B}{R} (J_X + f_Y^I v_A + f_Z^I v_B),
\]
where \( j_X = J_X/V_T \). The stationary condition is written as

\[
\begin{align*}
v_A \left( j_X + f^A_X (v_A - 1) + f^B_X v_B \right) &= r_A j_X, \\
v_B \left( j_X + f^B_X (v_B - 1) + f^A_X v_A \right) &= r_B j_X.
\end{align*}
\]

The stationary condition has solutions \((v_A, v_B) = (1,0), (0,1)\). Then the solution \((v_A, v_B) = (1,0)\) satisfies the condition with \( r_A = 1 \). Linearizing the rate equation around the fixed point (1,0) by writing \( v_A = 1 - \delta \), we get

\[
\frac{d\delta}{dt} = -\frac{1}{R} \left( f^A_X - f^B_X + j_X \right) \delta.
\]

When \( f^A_X \geq f^B_X \), the solution is stable. The solution \((v_A, v_B) = (0,1)\) satisfies the equation when \( r_A = 0 \). By linearizing the rate equation around the fixed point (0,1) and writing \( v_A = \delta \), we obtain

\[
\frac{d\delta}{dt} = -\frac{1}{R} \left( f^A_X - f^B_X + j_X \right) \delta.
\]

When \( f^B_X \geq f^A_X \), the solution is stable.

For the case \( f^A_X = f^B_X \), the equation is written as

\[
\frac{d\delta}{dt} = -\frac{j_X}{R} \delta,
\]

thus, both solutions are stable.

### D. The case \( S_Y \) and \( S_Z \) are limited: condition \((X, Y, Z) = (i, ii)\)

For the condition \((X, Y, Z) = (i, ii)\) when \( S_Y \) and \( S_Z \) are limited, increases of molecule species are written as

\[
F^A_X = V_A f^A_X, F^B_X = V_B f^B_X, F^A_Y = J_Y, F^B_Z = J_Z.
\]

Then,

\[
\frac{dv_A}{dt} = \frac{1}{R} \left( f^A_X v_A + j_Y \right) - \frac{v_A}{R} \left( f^A_X v_A + f^B_X v_B + j_Y + j_Z \right),
\]

\[
\frac{dv_B}{dt} = \frac{1}{R} \left( f^B_X v_B + j_Z \right) - \frac{v_B}{R} \left( f^A_X v_A + f^B_X v_B + j_Y + j_Z \right).
\]

The stationary state gives

\[
(f^A_X - f^B_X) v_A^2 + (-f^A_X + f^B_X + j_Y + j_Z) v_A - j_Y = 0.
\]

When \( f^A_X = f^B_X \), this equation reduces to a solution

\[
v_A = v_A^0 = \frac{j_Y}{j_Y + j_Z}, v_B = v_B^0 = \frac{j_Z}{j_Y + j_Z}.
\]

By linearizing the dynamics around the fixed point as \( v_A = v_A^0 + \delta \), we get

\[
\frac{d\delta}{dt} = -\frac{1}{R} \left( j_Y + j_Z \right) \delta,
\]

and the solution is stable.

When \( f^A_X > f^B_X \), the equation has a fixed-point solution with coexistence of \( A \) and \( B \) as

\[
v_A = v_A^0 = \frac{1}{2 \left( f^A_X - f^B_X \right)} \left\{ f^B_X - f^A_X - j_Y - j_Z \pm \sqrt{\left( f^B_X - f^A_X - j_Y - j_Z \right)^2 + 4 j_Y (f^A_X - f^B_X)} \right\}.
\]

As will be shown below, \( 0 < v_A^0 < 1 \), so that the two types coexist: First, since \( f^A_X - f^B_X > 0 \), \( v_A^0 > 0 \). Then the difference in the numerator and denominator in the expression of \( v_A^0 \) is given by

\[
(numerator) - (denominator) = -(f^A_X - f^B_X + j_Y + j_Z) + \sqrt{(f^A_X - f^B_X - j_Y - j_Z)^2 + 4 j_Y (f^A_X - f^B_X)}.
\]
Here,

\[|f_X^A - f_X^B + j_Y + j_Z|^2 - \{(f_X^A - f_X^B - j_Y - j_Z)^2 + 4j_Y(f_X^A - f_X^B)\} = 4j_Z(f_X^A - f_X^B) > 0,\]

so that the denominator is greater than the numerator. Thus, \(v_A^0 < 1\). Therefore, \(0 < v_A^0 < 1\).

By writing \(v_A = v_A^0 + \delta\), and linearizing the dynamics around the fixed point, we obtain

\[\frac{d\delta}{dt} = -\{(f_X^A - f_X^B)(2v_A^0 - 1) + j_Y + j_Z\} \delta.\]

The solution is stable if \((f_X^A - f_X^B)(2v_A^0 - 1) + j_Y + j_Z > 0\). For example, if \(f_X^A > f_X^B\) and \(j_Y \geq j_Z\), the condition is satisfied.

When \(f_X^B > f_X^A\), the stationary condition is

\[(f_X^B - f_X^A)v_B^2 + (-f_X^B + f_X^A + j_Y + j_Z)v_B - j_Z = 0,\]

for \(v_B\). The fixed-point solution is obtained in the same way as in the case \(f_X^A > f_X^B\), by replacing \(A\) with \(B\), and \(Y\) with \(Z\).

E. All the \(S_X, S_Y\) and \(S_Z\) are limited: condition \((X,Y,Z) = (ii, ii)\)

For the condition \((X,Y,Z) = (ii, ii)\), when all resources are limited,

\[F_X^A = r_AJ_X, F_X^B = r_BJ_X, F_Y^A = J_Y, F_Y^B = J_Z.\]

Then,

\[\frac{dv_A}{dt} = \frac{1}{R}(r_AJ_X + j_Y) - \frac{v_A}{R} (j_X + j_Y + j_Z),\]

\[\frac{dv_B}{dt} = \frac{1}{R}(r_BJ_X + j_Z) - \frac{v_B}{R} (j_X + j_Y + j_Z).\]

The steady state gives

\[v_A^0 = \frac{r_AJ_X + j_Y}{j_X + j_Y + j_Z}, v_B^0 = \frac{r_BJ_X + j_Z}{j_X + j_Y + j_Z}.\]

Again by writing \(v_A = v_A^0 + \delta\),

\[\frac{d\delta}{dt} = -\frac{j_X + j_Y + j_Z}{R} \delta,\]

follows, and the coexisting state is stable.

F. Numerical simulations

In this subsection, we present results of numerical simulations of the rate equations for the continuous concentration variables to show whether competition of limited resources results in dominance or coexistence of cells of type \(A\) and \(B\). The results for stochastic simulations are shown in the main text.

We consider \(M_{tot}\) cells of type \(A\) or \(B\). The dynamics of the number \(N_i^I\) of molecule species \(i (i = X,Y,Z)\) in a cell-type \(I (I = A,B)\) is written as eqs. (1) and (2). The dynamics of each resource \(S_i (i = X,Y,Z)\) is given by eqs. (3)-(5).

As an initial distribution of cell types, types \(A\) and \(B\) are each represented at 50%. In each type, the numbers \(N_X, N_Y\) in type \(A\), and \(N_X, N_Z\) in type \(B\) are randomly assigned. A cell divides when \(V_K = \sum_i N_i^K\) exceeds a threshold \(V_{max}\). At the division, the number of each molecule species \(N_i^K\) is divided between the two daughter cells with a normal distribution with average \(N_i^K/2\) and variances \(N_i^K/4\). The normal distribution approximates the random partition of molecules into two daughter cells.

The average time required for the initial condition to reach a state in which either of the types is extinct is shown in Fig. 2. The parameters approximately correspond to those of Fig. 1. The time drastically increases around \(D^* = 100,\)
FIG. 2: Average time to a state in which either A or B dominates. Here, $M_{\text{tot}} = 100$, $V_{\text{max}} = 1000$, $S_X^0 = S_Y^0 = S_Z^0 = 1000$.

FIG. 3: Average number of $N_X$, $N_Y$, and $N_Z$ at division events. The initial state is given as $n_A = n_B = M_{\text{tot}}/2$, and in each type A and B, we randomly distribute molecules $X$ and $Y$, and $X$ and $Z$, respectively, with equal probabilities.

which is consistent with the point where the resource $S_X$ starts to be limited in Fig. 1. Correspondingly, the average numbers of $N_X$, $N_Y$ and $N_Z$ at division events start to deviate below that point (Fig. 3), and available resources show deviations between $S_X$ and $S_Y$ or $S_Z$ below that point (Fig. 4).

At the point $D = D^*$, consumption and inflow of resources are balanced and $S_X$ starts to compete. The condition for balance in $S_X$ is

$$\frac{M_{\text{tot}} V}{4} \frac{S_X}{S_X^0} = D^* (S_X^0 - S_X)$$

where $V$ denotes the average volume of cells. By substituting $V = 750$, $S_X = 850$, $S_X^0 = 1000$, $M_{\text{tot}} = 100$, $D^*$ is estimated as $\sim 100$.

For $S_Y$ and $S_Z$, competition starts roughly when the inflow is half that of $S_X$ because approximately half of $M_{\text{tot}}$ cells consume the respective resource. This suggests that below $D = 0.5 D^*$, all the resources are limited; thus, two types of cells coexist. This is consistent with our numerical observations that below $D = 40$, the types coexist.
II. WHEN DIFFERENT MOLECULE SPECIES CONSUME COMMON RESOURCES: THE CASE WITH $K_R < K_M$

In this section, we investigate the case with $K_R < K_M$ where plural resource species are commonly used for the replication of different molecules $X_i$.

A. Model

In the reaction

$$X_j + X_i + S_j \rightarrow 2X_j + X_i,$$

the correspondence between molecule $X_j$ and resource $S_j$ is randomly assigned and fixed throughout simulations. Hence, each resource species is used commonly for $K_M/K_R$ reactions on the average.

B. Results

We show the number of molecule species in each cells (compositional diversity) and that in 10 or more cells (phenotypic diversity), respectively, in Fig. 5(A) for $K_R = 1$, i.e., a single resource is consumed to replicate all the molecule species $X_i$ ($i = 1, \ldots, K_M$). In this case, neither compositional nor phenotypic diversity increases as the diffusion constant $D$ decreases.

When two resource species are available ($K_R = 2$; Fig. 5(B)), the phenotypic diversity increases as $D$ decreases, but a clear increase is not discernible in the compositional diversity. In the case of small $K_R$, the randomly determined reservoir concentrations, $S_i^0 \in [0, M_{\text{tot}}]$ ($i = 1, 2$), are also relevant parameters to determine the point at which the resources become limited, in addition to the diffusion constant $D$. Here, we also show the results for fixed $S_i^0 = M_{\text{tot}}$ ($i = 1, 2$). In contrast to the $K_M = K_R$ case, there is an increase in phenotypic diversity for $D < 0.1$, while even below the point, the number of remaining chemical species is approximately constant and does not increase as $D$ is decreased. This result indicates that the diversity is bounded by the number of resource species: coexistence of at most two cell types is possible.

For larger $K_R$ ($K_R = 10$ in Fig. 5(C), and $K_R = 100$ in Fig. 5(D)), both compositional and phenotypic diversity increase as $D$ is decreased because $K_R$ is sufficiently large so that it does not effectively restrict the number of cell types. The phenotypic diversity, i.e., the number of coexisting cell types increases as $\sim K_R$, but the increase is saturated for larger $K_R$ (see Fig. 5), which is also bounded by $M_{\text{tot}}$. Indeed, as $M_{\text{tot}}$ is increased, the number increases (see below).
FIG. 5: Compositional and phenotypic diversity plotted as a function of $D$ for (A) $K_R = 1$, (B) $K_R = 2$, (C) $K_R = 10$, and (D) $K_R = 100$. Compositional and phenotypic diversity are computed in the same way as in the main text, i.e., as the numbers of chemical species included in (a) each cell, and in (b) more than 10 cells out of $M_{tot}$ cells are shown, respectively. For (B) $K_R = 2$, the data for two cases are shown: $S_0(i = 1, 2)$ is determined randomly between 0 and $M_{tot}$, and both $S_0(i = 1, 2)$ are fixed to $M_{tot}$. All the data are obtained as the average over $10^5$ division events in 5 different networks with $M_{tot} = 100$, $K_M = 200$, $N = 1000$, and $\mu = 0.001$.

III. DEPENDENCE ON $M_{tot}$

We investigated the dependence of the diversity on the number of interacting cells, $M_{tot}$, in the case $K_M = K_R$. The number of molecule species in each cells (compositional diversity) and that in 10 or more cells (phenotypic diversity) for $M_{tot} = 100, 200, 300$ is shown in Fig. 6.

While both measures of diversity increase for each $M_{tot}$, as $D$ is decreased, the increment depends on $M_{tot}$. The increment in compositional diversity decreases as $M_{tot}$ is increased. On the other hand, the increment in phenotypic diversity with the decrease in $D$ increases as $M_{tot}$ is increased. As shown in Fig. 3(C) of the main text, the two types
of diversity clearly show opposite dependence on $M_{\text{tot}}$ for fixed $D$.

As discussed in the previous section, multiple cell types (up to $K_R$) can coexist; thus, more types of cells can be present and are not eliminated from the system as $M_{\text{tot}}$ is increased, which results in the increase in phenotypic diversity. On the other hand, as the number of cell types increases, a greater number of resource species are competitive because the cell population can consume more resource species. This leads each individual cell to be more specialized with fewer components, which results in suppression of the increase in compositional diversity.

IV. SUPPLEMENTARY FIGURES

Figs 7 and 8 show similarities $H_{ij}$ among a period of division events, where the cell indices given by the $x-$ and $y-$ axes are rearranged so that the same (similar) types are clustered. The data are identical to the similarities represented in Fig. 2(II)(iii) and (III)(iii) of the main text, but the plots are given after rearrangement of cell indices.

Types II-A and II-B are clustered as shown at the top of Fig. 7. The similarities between cells of the same type have values close to 1, while types A and B have similarities around 0.6.

The types from III-A to III-F are clustered at the top of Fig. 8. The similarities between cells of the same type have values close to 1, and those between types A and B, and between $D$ and $E$ have positive values. For other pairs, the values are almost orthogonal.
FIG. 7: The indices are arranged from Fig. 2(II)(iii) in the main text to categorize into each type A and B.

FIG. 8: The indices are arranged from Fig. 2(III)(iii) in the main text to categorize into each type A – F.