Transition to diversification by competition for resources in catalytic reaction networks

Atsushi Kamimura and Kunihiro Kaneko

Department of Basic Science, The University of Tokyo, 3-8-1, Komaba, Meguro-ku, Tokyo 153-8902, Japan
(Dated: August 12, 2014)

All life, including cells and artificial protocells, must integrate diverse molecules into a single unit in order to reproduce. Despite expected pressure to evolve a simple system with the fastest replication speed, the mechanism by which the use of the great variety of components, and the coexistence of diverse cell-types with different compositions are achieved is as yet unknown. Here we show that coexistence of such diverse compositions and cell-types is the result of competitions for a variety of limited resources. We find that a transition to diversity occurs both in chemical compositions and in protocol types, as the resource supply is decreased, when the maximum inflow and consumption of resources are balanced.

PACS numbers: 87.17.-d, 82.40.Qt, 89.75.Fb

Cells, even in their primitive forms, must integrate diverse molecules into a single unit so that they keep reproducing. The use of diverse chemical components is sustained from the origins of life, while biodiversity in ecosystems is emphasized as salient characteristics. How diversity originated in cellular reproduction is an important open question. One may naively expect that a simple replicator consisting of only a few molecular species would evolutionarily triumph as it can reproduce faster than a complex system using diverse components. Additionally, as replicators with the fastest division speeds would dominate, coexistence of diverse replicators would not be expected. In fact, several in silico artificial life models show dominance of such simple replicators.

In nature, however, diverse components exist within a cell and diversity of cell types(replicators) is sustained. Are both “compositional” diversity at an individual level and “phenotypic” diversity at the population level present at the level of “protocols”, a primitive early stage of cellular evolution, in spite of the selective pressure for survival?

Chemical protocol replicators have been investigated using the hypercycle model, in which different molecular species mutually catalyze the replication of each other. Protocells encapsulating such hypercycles can exhibit robust reproduction even in the presence of parasitic molecules that may destroy the mutual catalytic reactions. In these studies, it is often assumed that chemical resources for synthesis of biopolymers are supplied sufficiently. Thus, a protocol consisting of a simple hypercycle with few components replicates faster, and quickly dominates the pool.

As the catalytic activities of molecules and cell populations increase, however, nutrient depletion or resource competition is inevitable. When only a single resource is provided, competition for the limited resource leads to survival of only the fittest protocol type. When multiple resources are competed for, then, do protocells diversify into distinct types specialized for the use of different resources, and is coexistence of diverse cell types possible? Here, we show through numerical simulation of a model of interacting protocells consisting of hypercycle reaction networks that such a transition to increased diversity occurs when available resources are limited.

For this purpose, we consider a protocol model in which each molecule $(X_j; j = 1,..., K_M)$ replicates with the catalytic aid of $X_i$, predetermined by a random network, by consuming a corresponding resource $(S_j; j = 1,..., K_R = K_M)$. The outline of the model is as follows(Fig. 1). There exist $M_{tot}$ protocols, consisting of $K_M$ species of replicating molecules where some possibly have null population. Molecules of each species $X_j$ are replicated with the aid of some other catalytic molecule $X_i$, determined by a random catalytic reaction network, by consuming a predetermined resource $S_j$, one of the supplied resource chemicals $S_k (k = 1,..., K_R)$, as follows:

$$X_j + X_i + S_j \xrightarrow{c_i} 2X_j + X_i.$$

For this reaction to replicate $X_j$, one resource molecule is needed, and the replication reaction does not occur $S_j < 1$. When $K_M = K_R$, each resource species $S_j$ corresponds to each molecule species $X_j$, while for $K_M > K_R$, a common resource molecule species is used for replication of multiple molecule species $X_{j_1}, X_{j_2},...$, as discussed in the supplement. The reaction coefficient is given by the catalytic activity $c_i \in [0,1]$ of the molecule species $X_i$. With each replication, error occurs with probability $\mu$ as shown below.

The resources $(S_i)$ diffuse with diffusion constant $D$ through a common medium for a population of protocells from external reservoirs of concentrations $S_i^0$. $D$ controls the degree of the resource competition, as the resource supply is limited with decreasing $D$.

For each molecule species, the density for the path of the catalytic reaction is given by $\rho$ (which is fixed at 0.1) so that each species has $\rho K_M$ catalysts on average. Once chosen, the reaction network is fixed throughout each run of simulations. Autocatalytic reactions in which $X_i$
catalyzes the replication of itself are excluded from the random catalytic network, and direct mutual connections are excluded so that \(X_j\) does not work as a catalyst for \(X_i\) if \(X_i\) is the catalyst for \(X_j\).

Structural changes may occur during replication to alter monomer sequences of polymers and catalytic properties of the molecule. In the present model, this alteration is included as a random change to other molecular species during the replication process. When replication of \(X_j\) occurs, it is replaced by another molecule \(X_i (i \neq j)\) with a probability \(\mu\). For simplicity, we assume that this error leads to all other molecule species with equal probability, \(\mu/\left(K_M - 1\right)\) where \(K_M\) is the number of molecule species.

When the total number of molecules in each cell exceeds a given threshold \(N\), the cell divides into two and randomly partitions molecules, and one randomly chosen cell is removed from the system in order to fix the total number of cells at \(M_{\text{tot}}\).

We simulated the model by changing the speed of resource supply \(D\). When the resources are supplied sufficiently fast (e.g., for \(D = 1\)), a recursively growing state is established with a few molecular species, where the composition is robust against noise and perturbations by the division process. In this state, the (typically) three primary components form a three-component hypercycle.

Other molecular species in a cell are typically catalyzed by the division process. In this state, the (typically) three primary components form a three-component hypercycle. Other molecular species in a cell are typically catalyzed by a member of the hypercycle. All the dividing cells adopt this three-component hypercycle, thus, there is neither compositional nor phenotypic diversity (Fig. 21).

To check cell reproduction fidelity, we introduced similarity between cells as follows: as each cellular state is characterized by the number of molecules of each chemical species \(\vec{N}_i = (n_1, n_2, \ldots, n_{K_M})\), similarity is defined as the inner product of these composition vectors between two cell division events, i.e., \(H_{ij} = \vec{N}_i \cdot \vec{N}_j / (|\vec{N}_i| \cdot |\vec{N}_j|)\) between the \(i\)-th and \(j\)-th division events. In the above case, the similarity between mother and daughter cells is close to one, implying high-fidelity reproduction.

As \(D\) decreases below 0.1, phenotypic diversity starts to increase. For example, two cell types (II-A, B) coexist in Fig. 2II) and consist of three-component hypercycles differing by one component. Both types divide with approximately equal speed and coexist over \(10^2\) generations. In 200 successive division events (Fig. 2II)(iii)), one type has a similarity near unity (red), and the similarity of the other ranges between 0.6 and 0.7 (yellow), implying that the two types mostly reproduce themselves with a small probability (approximately 0.01/division) of switching types. Over much longer generations, replication errors can produce different types capable of replacing the existing cell types.

As \(D\) decreases further, both the phenotypic and compositional diversity continue to increase. For \(D = 0.01\) (Fig. 2III)), six cell types (A–F) appear. Each type forms a distinct hypercycle network in which the species catalyze replication of each other. Here, some types (III–A and B, III–D and E) share some common molecular species, while the others do not. Similarieties are approximately equal to unity (red) for some cells, while cell types that have the common chemical components (yellow to light blue) as well as cell types with completely orthogonal composition (dark blue) appear from time to time (Fig. 2III)(iii)). Also, the number of replicating chemical species in each cell is slightly increased (see Fig. 2III(ii)). As \(D\) decreases to the order of 0.001, more cells with lower similarity appear, and compositional and phenotypic diversity further increase.

We statistically studied the quantitative dependence of diversity upon the parameter \(D\) using a variety of networks. As \(D\) decreases below \(\sim 0.1\), the compositional diversity of each protocol and the phenotypic diversity at the population level increase (Fig. 3A)). With the increase in diversity, reproductive fidelity decreases both at an individual level (Fig. 3B(i)) and at a population level, i.e., over all pairs of \(10^4\) division events in 30 runs (Fig. 3B(h)). Protocells with different hypercycles start to appear below \(D \sim 0.1\) (Fig. 2).

The transition to increased diversity generally occurs for sufficient resource diversity, i.e., for large \(K_R\), independent of reaction network choice. The phenotypic diversity increases as \(\sim K_R\), but is bounded by the finite number of interacting cells, \(M_{\text{tot}}\) (Supplementary Information). As \(M_{\text{tot}}\) increases, the number of coexisting cell types increases, while the compositional diversity, i.e., the number of components in each cell type, decreases (see Fig. 3C)). This trade-off between compositional and phenotypic diversity suggests that each cell-type is specialized for fewer chemical components as the number of cell types increases.

Altogether, the data show transition behavior around \(D = D_c \sim 0.1\). Below \(D_c\), the division speed decreases, while above \(D_c\) it is approximately constant (Fig. 2D)). This suggests that at the transition point the maximum inflow and consumption rates of resources are balanced. The maximum inflow rate is estimated as \(DS^0_j\), where \(S^0_j\) is a typical reservoir concentration. The intrinsic consumption rate by all cells is estimated to be \(M_{\text{tot}}c_j\rho_j^0\), where \(c_j\) is a typical catalytic activity, and \(\rho_j^0\) is the probability of picking up a pair that replicates the molecule \(X_j\). When sufficient resources are available, the three-component hypercycle dominates, and \(\rho_j^0 \sim 1/9\). Therefore, the value of \(D\) at which the maximum inflow and consumption rates of resources are balanced is estimated as \(D = M_{\text{tot}}c_j/9S^0_j\). Since \(c_j\) is distributed homogeneously in \([0, 1]\), its average value is 0.5; however the remaining components are typically biased to have higher catalytic activities, and \(c\) therefore typically ranges from 0.7 to 0.8. On the other hand, \(S^0_j\) is distributed homogeneously in \([0, M_{\text{tot}}]\), so the simple average of \(S^0_j\) is 50 with \(M_{\text{tot}} = 100\), but the remaining components are biased with have \(S^0_j \sim 70\). Thus the critical value of \(D\) is approximately 0.11 – 0.13. It is noteworthy that below \(D_c\), the division speed decreases more slowly than the inflow rate, as indicated by the line proportional to \(D\) (Fig. 3D)), suggesting cells can utilize more diverse resources for growth by increasing the available number of resource
species.

For $D > 0.2$, sufficient resources are available, thus, the intrinsic reaction rate of the three-component hypercycle is the main determinant of the division speed. The probability, $p_{ij}$, for picking up a pair between which a catalytic reaction exists is $\sim 1/3$. Using the above typical catalytic activity $\tilde{c}$, the division speed is estimated as $2\tilde{c}/3N$, which is approximately 0.0005 for $N = 1000$(See Fig. 2D).

Why does the transition to diversity occur with the decrease in resources? A simple case illustrates this diversity transition from dominance of a single type to coexistence of various types. Consider two types of cells that compete for resources. One type ($A$) consists of molecules $X$ and $Y$, while the other type ($B$) consists of molecules $X$ and $Z$. Each of molecule $X$, $Y$, and $Z$ is replicated by mutually catalytic reactions utilizing resources, $S_X$, $S_Y$, and $S_Z$, respectively. When supplied sufficient resources, the population of each cell type grows exponentially, where the growth rate is proportional to the number $n_{A(B)}$ of each cell type $A$ and $B$, i.e., $\frac{dn_{A(B)}}{dt} \propto \gamma_{A(B)}n_{A(B)}$, and the proportionality coefficient, $\gamma_{A(B)}$, is proportional to resource abundances $S_{Y(Z)}$. In this case, Darwinian selection works so that the type with a larger $\gamma$ dominates. This selection process works as long as the resources are sufficient. However, when all the resources are limited, competition for the available resources $S_{Y(Z)}$ among $n_{A(B)}$ cells decreases $\gamma_{A(B)}$, so that it is inversely proportional to $n_{A(B)}$, respectively. The population dynamics are, therefore, represented by $\frac{dn_{A(B)}}{dt} \propto c_{A(B)}$, with constant $c_{A(B)}$ when the maximum inflow is decreased to balance with the rate of resource consumption. In this form, the solution with two coexisting types is known to be stable $[16, 22]$ (Supplemental Information). The transition to diversity in composition and phenotypes in the original model is based on this change from exponential to linear growth due to resource limitation. On the other hand, the above argument on the diversity transition supports the generality of our result, as the number growth in chemical components by catalytic reactions changes from exponential to linear with the decrease in resources.

When life originated, a set of diverse, self-replicating catalytic polymers(replicators) would emerge from the primordial chemical mixture, to make a reproducing protocell. Although the importance of molecular and protocellular diversity has been noted by Dyson $[4]$, their origins are not well addressed, especially as compared with diversity in ecological systems. Our results indicate that competition for a variety of limiting resources can be a strong driving force to diversify intracellular dynamics of a catalytic reaction network and to develop diverse protocell types in a primitive stage of life. Indeed, it is natural that diverse chemical resources are available in the environment, and competition for resources increases as the protocells reproduce and more cells compete. Thus, diversification in composition and protocell types is an inevitable outcome.

According to our results, the diversification is understood as a kind of phase transition in population dynamics with decreased resource. In ecology, the niche dimension hypothesis in plant communities was proposed, in which the number of coexisting species increases with the number of limiting factors increases while greater resource abundance decreases diversity $[20, 22]$. Our results suggest that such population dynamics in ecology are possible with primitive mixtures of catalytic molecules competing for a variety of resources $[24]$, even without the need for genetic changes as in speciation.

This work is supported by the Japan Society for the Promotion of Science. This work is also supported in part by the Platform for Dynamic Approaches to Living System from the Ministry of Education, Culture, Sports, Science, and Technology of Japan, and the Dynamical Micro-scale Reaction Environment Project of the Japan Science and Technology Agency.

[1] Protocells: Bridging Nonliving and Living Matter. Ed. S. Rasmussen, et al. (The MIT Press, Cambridge, 2008)
[2] J. C. Blain, and J. W. Szostak, Annu. Rev. Biochem. 83:1.1-11.26 (2014).
[3] V. Noireaux, Y. T. Maeda, A. Libchaber, Proc. Natl. Acad. Sci. U. S. A. 108, 3473-80 (2011).
[4] F. Dyson, Origins of Life (Cambridge Univ. Press, Cambridge, 1985).
[5] R. M. May, Nature 238, 413-414 (1972).
[6] K. S. McCann, Nature 405, 228-233 (2000).
[7] M. Loreau et al., Science 294, 804-808 (2001).
[8] A. R. Hughes, B. D. Inouye, M. T. J. Johnson, N. Underwood, M. Vellend, Ecology Letters 11: 609-623 (2008).
[9] T. G. Whitham, J. K. Bailey, J. A. Schweitzer, S. M. Shuster, R. K. Bangert, et al. Nature Reviews Genetics 7, 510-523 (2006).
[10] W. Fontana, and L. W. Buss, Bull. Math. Biol. 56, 1-64 (1994).
[11] R. S. Ray, Artificial Life II pp 371-408. Ed. C. G. Langton. (Addison Wesley, Redwood City, 1991).
[12] R. Albert, and A. L. Barabasi, Rev. Mod. Phys 74, 47-97 (2002).
[13] S. Jain, and S. Krishna, Proc. Natl. Acad. Sci. U. S. A. 99, 2055-2060 (2002).
[14] R. Mehrotra, V. Soni, and S. Jain, J. R. Soc. Interface 6, 793-799 (2009).
[15] M. Eigen, and P. Schuster, The Hypercycle (Springer, New York, 1977).
[16] M. Eigen, Steps Towards Life (Oxford University Press, Oxford, 1992).
[17] J. M. Smith, Nature 280, 445-446 (1979).
[18] E. Szathmary, and L. Demeter, J. Theor. Biol. 128, 463-486 (1987).
[19] S. Altmeyer, and J. S. McCaskill, Phys. Rev. Lett. 86, 5819-5822 (2001).
FIG. 1: Schematic representation of our model. The system is composed of $M_{tot}$ cells, each of which contains molecule species $X_j$ ($j = 1, ..., M$) that form a catalytic reaction network to replicate each $X_j$. The cells share each resource $S_k$, which is consumed by the replication of molecule $X_k$. The resources flow into a common bath for the cells from an external reservoir (environment) via diffusion $-D(S_k - S_0^k)(k = 1, ..., K_R)$, where $S_0^k$ is a randomly-fixed constant $S_0^k \in [0, M_{tot}]$ and $D$ is the diffusion constant.
FIG. 2: Compositions and types of reproducing protocells for (I) \( D = 1 \), (II) \( D = 0.1 \), and (III) \( D = 0.01 \). Compositions of molecule species are shown in successive 200 division events in the system with \( K_M = K_R \). (i) At successive division events (abscissa), molecule species indices with more than 20 copy numbers in the dividing cell are marked. Cells are categorized into a few, or, several types, indicated by different colors, according to the majority set of molecule species. For each type, the indices of majority species are shown to the right of the figures. (ii) The catalytic network formed by the majority molecule species is shown for each cell type. Each numbered node corresponds to a molecule species, and the arrow from index \( i \) to \( j \) represents catalysis of \( X_j \) replication by \( X_i \). (iii) Similarities \( H_{ij} \) between cells \( i \) and \( j \), as defined in the text, are plotted over 200 successive division events, using a color code from \( H_{ij} \sim 1 \) (red) to \( H_{ij} \sim 0 \) (dark blue). If the similarity is close to one, the composition is almost preserved by cell division, while zero similarity indicates that a reproduced cell has completely different composition. Intermediate values of similarity indicate the overlap in some molecule species between the cells, as in Type II-A and -B, Type III-A and -B, and Type III-D and -E. Parameters are \( K_M = K_R = 200 \), \( M_{tot} = 100 \), \( N = 1000 \), and \( \mu = 0.001 \).
FIG. 3: (A) Compositional and phenotypic diversity plotted as a function of $D$. For compositional and phenotypic diversity, the numbers of chemical species included in each cell (left ordinate), and in more than 10 cells out of $M_{\text{tot}}$ cells (right ordinate) are shown. (B)(a) Average similarity of mother and daughter cells and (b) average similarity between cells in $10^4$ successive division events in 30 different networks, plotted against $D$. For (b), the average over cells with positive similarity (red) is shown in addition to the average for all cells (green). (C) Dependence of compositional (left ordinate) and phenotypic (right ordinate) diversity on the number of cells, $M_{\text{tot}}$, with fixed $D = 0.01$. (D) Average division speed as a function of $D$. For reference, a linear line with $D$ is plotted. For $D > 0.2$, an estimate of the division speed ($\sim 0.0005$) is also shown with a dotted line (see main text). Unless otherwise indicated, the data are obtained as the average over $10^5$ division events in 30 different networks. The parameters are $K_M = K_R = 200$, $N = 1000$, and $\mu = 0.001$. 
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 $X$ and $Y$ for $i = A, B$,

$$f_X^A = f_Y^A = f_Z^B = f_X^B,$$ 

respectively, for cell type $A$ and $B$. Here, $V_A$, and $V_B$ are the volumes of cell types $A$ and $B$, respectively, and $f_X^A = c_Y x_X^A x_Y^A$, $f_Z^B = c_Z x_Z^B x_Z^B$, where $x_i^1 = N_i^1 / V_i (i = X, Y, Z; I = A, B)$ and $N_i^1$ is the number of molecule species $i$. The rates of replicating $Y$ and $Z$ are, respectively, given by

$$F_Y^A = V_A f_Y^A s_Y, \quad F_Z^B = V_B f_Z^B s_Z.$$ 

Here, $F_Y^A = c_Y x_X^A x_Y^A$, $f_Z^B = c_Z x_Z^B x_Z^B$. $s_i = S_i / S_i^0 (i = X, Y, Z)$ 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,

$$\frac{dN_A}{dt} = F_X^A = V_A f_X^A s_X, \quad \frac{dN_B}{dt} = F_X^B = V_B f_X^B s_X,$$ 

and, for $Y$ and $Z$,

$$\frac{dN_A}{dt} = F_Y^A = V_A f_Y^A s_Y, \quad \frac{dN_B}{dt} = F_Z^B = V_B f_Z^B s_Z.$$ 

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

$$\frac{dS_X}{dt} = - (V_A f_X^A + V_B f_X^B) s_X + DS_X^0 (1 - s_X),$$ 

$$\frac{dS_Y}{dt} = - V_A f_Y^A s_Y + DS_Y^0 (1 - s_Y),$$ 

$$\frac{dS_Z}{dt} = - V_B f_Z^B s_Z + DS_Z^0 (1 - s_Z).$$ 

In the steady state, the value of $s_X$, $\bar{s}_X$, is written as,

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

Similarly, for resources $S_Y$ and $S_Z$ the steady state values are written as

$$\bar{s}_Y = \frac{DS_Y^0}{V_A J_Y + DS_Y^0}, \quad \bar{s}_Z = \frac{DS_Z^0}{V_B J_Z + DS_Z^0}. \tag{7}$$

The dynamics of each molecule species changes with the diffusion constant $D$. When $D$ is sufficiently large, $\bar{s}_X$ is approximately one. As $D$ decreases and the resource is limited, $\bar{s}_X$ deviates from one, and the maximum inflow rate of the resource is given by $J_X = DS_X^0(1 - \bar{s}_X)$. 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

$$\begin{align*}
(F_X^A, F_X^B) &= \begin{cases} 
(V_A J_X, V_B J_X) & \text{for (X-i)} \\
(r_A J_X, r_B J_X) & \text{for (X-II)}
\end{cases} \\
(F_Y^A, F_Y^B) &= \begin{cases} 
(V_A f_Y^A, V_B f_Y^B) & \text{for (Y-i)} \\
(r_A f_Y^B, r_B f_Y^B) & \text{for (Y-II)}
\end{cases}
\end{align*}$$

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): $S_Y$ and $S_Z$ are sufficiently supplied, and (YZ-II): $S_Y$ and $S_Z$ are limited. For each range, the right-hand-sides of the rate equations (2) can be written as

$$\begin{align*}
F_Y^A &= \begin{cases} 
V_A f_Y^A & \text{for (YZ-i)} \\
J_Y & \text{for (YZ-II)}
\end{cases} \\
F_Y^B &= \begin{cases} 
V_B f_Y^B & \text{for (YZ-i)} \\
J_Z & \text{for (YZ-II)}
\end{cases}
\end{align*}$$

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

$$\frac{dV_A}{dt} = \frac{1}{R}(F_X^A + F_Y^A), \quad \frac{dV_B}{dt} = \frac{1}{R}(F_X^B + F_Y^B).$$
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 \( \sum V_k = V_T = \text{const} \). Then the volume fraction of each type \( A \) and \( B \) follows

\[
\begin{align*}
\frac{dv_A}{dt} &= \frac{1}{RV_T} (F^A_X + F^A_Y) - \frac{v_A}{R} \sum F^I_i \frac{V_T}{V_T}, \\
\frac{dv_B}{dt} &= \frac{1}{RV_T} (F^B_X + F^B_Y) - \frac{v_B}{R} \sum F^I_i \frac{V_T}{V_T}
\end{align*}
\]

where \( v_I = V_I / V_T (I = A, B) \), \( F_I = \sum F^I_i \).

**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

\[
F^A_X = V_A f^A_X, F^A_Y = V_A f^A_Y, F^B_X = V_B f^B_X, F^B_Y = V_B f^B_Y,
\]

where

\[
f^A_X = c_x x^A x^A, f^B_X = c_x x^B x^B, f^A_Y = c_x x^A x^Y, f^B_Y = c_x x^B x^Y,
\]

and \( x^I_i = N^I_i / V_I (I = A, B, i = X, Y, Z) \). Thus,

\[
\frac{dv_A}{dt} = \frac{f^A}{R} v_A - \frac{v_A}{R} \sum_i v_I f^I_i,
\]

\[
\frac{dv_B}{dt} = \frac{f^B}{R} v_B - \frac{v_B}{R} \sum_i v_I f^I_i,
\]

where \( f^A = f^A_X + f^A_Y, f^B = f^B_X + f^B_Y \). The stationary states give

\[
v_A f^A = v_A (v_A f^A + v_B f^B),
\]

\[
v_B f^B = v_B (v_A f^A + v_B f^B).
\]

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

\[
\frac{d\delta}{dt} = \frac{f_A}{R} \delta - \frac{\delta}{R} (\delta f^A + (1 - \delta) f^B)
\]

\[
= \frac{\delta}{R} (f_A - f_B).
\]

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,

\[
F^A_X = r_A J_X, F^B_X = r_B J_X, F^A_Y = V_A f^A_Y, F^B_Y = V_B f^B_Y,
\]

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,

\[
\frac{dv_A}{dt} = \frac{1}{R} (r_A J_X + f^A_Y v_A) - \frac{v_A}{R} (j_X + f^A_Y v_A + f^B_Y v_B),
\]

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

\[
\begin{align*}
  v_A \left\{ j_X + f_A^A (v_A - 1) + f_A^B v_B \right\} &= r_A j_X, \\
  v_B \left\{ j_X + f_B^B (v_B - 1) + f_B^A v_A \right\} &= r_B j_X.
\end{align*}
\]

The stationary condition has solutions \((v_A, v_B) = (1, 0), (0, 1)\). 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

\[
\frac{d\delta}{dt} = -\frac{1}{R} (f_A^A - f_A^B + j_X) \delta.
\]

When \( f_A^A \geq f_A^B \), 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

\[
\frac{d\delta}{dt} = -\frac{1}{R} (f_A^B - f_A^A + j_X) \delta.
\]

When \( f_A^B \geq f_A^A \), the solution is stable.

For the case \( f_A^A = f_A^B \), 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

\[
F_A^X = V_A f_A^A, F_B^X = V_B f_B^B, F_Y^A = J_Y, F_Z^B = J_Z.
\]

Then,

\[
\begin{align*}
  \frac{dv_A}{dt} &= \frac{1}{R} \left( f_A^A v_A + j_Y \right) - \frac{v_A}{R} \left( f_A^A v_A + f_B^B v_B + j_Y + j_Z \right), \\
  \frac{dv_B}{dt} &= \frac{1}{R} \left( f_B^B v_B + j_Z \right) - \frac{v_B}{R} \left( f_A^A v_A + f_B^B v_B + j_Y + j_Z \right).
\end{align*}
\]

The stationary state gives

\[
(f_A^A - f_A^B) v_A^0 + (-f_A^A + f_B^A + j_Y + j_Z) v_A - j_Y = 0.
\]

When \( f_A^A = f_A^B \), this equation reduces to a solution

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

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

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

and the solution is stable.

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

\[
v_A = v_A^0 = \frac{1}{2 \left( f_A^A - f_B^B \right)} \left\{ f_A^A - f_B^B - j_Y - j_Z + \sqrt{(f_A^A - f_B^B - j_Y - j_Z)^2 + 4 j_Y (f_A^A - f_A^B)} \right\}.
\]

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

\[
\left( \text{numerator} \right) - \left( \text{denominator} \right) = -(f_A^A - f_B^A + j_Y + j_Z) + \sqrt{(f_A^A - f_B^A - j_Y - j_Z)^2 + 4 j_Y (f_A^A - f_B^A)}.
\]
Here,
\[ |f_X^A - f_X^B + jy + jz|^2 - \{(f_X^A - f_X^B - jy - jz)^2 + 4jy(f_X^A - f_X^B)\} = 4jz(f_X^A - f_X^B) > 0, \]
so that the denominator is greater than the numerator. Thus, \( v_A^0 < 1 \). Therefore, \( 0 < v_A^0 < 1 \).

By writing \( \nu = v_A^0 + \delta \), and linearizing the dynamics around the fixed point, we obtain
\[
\frac{d\delta}{dt} = -\{(f_X^A - f_X^B)(2v_A^0 - 1) + jy + jz\} \cdot \delta. 
\]
The solution is stable if \( (f_X^A - f_X^B)(2v_A^0 - 1) + jy + jz > 0 \). For example, if \( f_X^A > f_X^B \) and \( jy \geq jz \), the condition is satisfied.

When \( f_X^B > f_X^A \), the stationary condition is
\[
(f_X^B - f_X^A)v_B^2 + (-f_X^B + f_X^A + jy + jz)v_B - jz = 0, 
\]
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 \( Z \).

\[ E. \quad \text{All the} \ S_X, \ S_Y \text{and} \ S_Z \text{are limited: condition} \ (X,Y,Z) = (ii,ii) \]

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,
\[
\frac{dv_A}{dt} = \frac{1}{R} \left( r_AJ_X + jy \right) - \frac{v_A}{R} \left( jx + jy + jz \right), \\
\frac{dv_B}{dt} = \frac{1}{R} \left( r_BJ_X + jz \right) - \frac{v_B}{R} \left( jx + jy + jz \right). 
\]
The steady state gives
\[ v_A^0 = \frac{r_AJ_X + jy}{jx + jy + jz}, \quad v_B^0 = \frac{r_BJ_X + jz}{jx + jy + jz}. \]

Again by writing \( \nu = v_A^0 + \delta \),
\[
\frac{d\delta}{dt} = -\frac{jx + jy + jz}{R} \cdot \delta, 
\]
follows, and the coexisting state is stable.

\[ F. \quad \text{Numerical simulations} \]

In this subsection, we present results of numerical simulations to show whether competition of limited resources results in dominance or coexistence of cells of type \( A \) and \( B \).

\[ 1. \quad \text{Continuous simulation} \]

We have simulated the rate equations for the continuous concentration variables to confirm coexistence by resource competition.

We consider \( M_{\text{tot}} \) 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. \[3\] - \[4\].

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
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.

$V_{\text{max}}$. At the division, the number of each molecule species $N^K_i$ is divided between the two daughter cells with a normal distribution with average $N^K_i/2$ and variances $N^K_i/4$. 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 $D^* = 100$, 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

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

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 $D = 0.5D^*$, 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.
2. Stochastic simulation

We next show the results of stochastic simulations. The setup is given in the same way as in the main text. The resources $S_X$, $S_Y$, and $S_Z$ flow from the outer reservoir with diffusion constant $D$. We consider $M_{tot}$ cells, each of which consists of either pair of $X, Y$ or $X, Z$, corresponding to type $A$ or $B$, respectively. For simplicity, we assume $c_X = c_Y = c_Z = 1$, $S_X^0 = S_Y^0 = S_Z^0 = M_{tot}$. The diffusion constants are identical for $X, Y$ and $Z$.

The evolution of the number of type $A$, $n_A$, for $D = 0.3, 0.2$ and $0.1$ is given in Fig. 5. For $D = 0.3$, either $A$ or $B$ is extinct after a relatively short time span. Around $D = 0.2$, the transient time before the extinction increases, and two types coexist over more than $10^5$ division events for $D = 0.1$.

The average number of division events to achieve the dominant state as a function of $D$ is given in Fig. 6. It is clear that the coexistence time starts to increase around $D^* = 0.25$, which is consistent with the point where the resources start to compete. In the stationary state without competition, $N^A_X = N^A_Y$ and $N^B_X = N^B_Z$. Thus, the point where $S_X$ is limited is given by

$$D^* S_X^0 = \frac{1}{4} M_{tot}$$

As $D$ is further decreased, the coexistence of types $A$ and $B$ is achieved when $S_Y$ and $S_Z$ are also limited and competed for by cells. In this case, the coexistence is stable as analyzed in the previous subsection. When all the resources are limited, the steady state approaches $N^A_X = 2N^A_Y$, $N^B_X = 2N^B_Z$. This gives

$$D^+ S^0 \sim \frac{M_{tot}}{2} \frac{1}{3} \frac{1}{3} = \frac{M_{tot}}{9}.$$  

For $S^0 = M_{tot}$, $D^+ \sim 0.11$. This is also consistent with the observation in Fig. 5.
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. 7(A) for $K_R = 1$, i.e., a single resource is consumed to replicate all the molecule species $X_i$ ($i = 1, ..., K_M$). 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. 7(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, $S_{i0} \in [0, M_{tot}]$ ($i = 1, 2$), are also relevant parameters to determine the point at which the resources become limited, in addition to the diffusion constant. Here, we also show the results for fixed $S_{i0} = M_{tot}$ ($i = 1, 2$). 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. 7(C), and $K_R = 100$ in Fig. 7(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 $\sim K_R$, while even below the point, the number of remaining chemical species is approximately constant and does not increase as $D$ is decreased. Indeed, as $M_{tot}$ is increased, the number increases (see below).

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. 8.

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 $M_{tot}$ 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 $M_{tot}$ 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 9 and 10 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. 9. 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. 10. 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: 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 $\mu = 0.001$. 
FIG. 8: Compositional and phenotypic diversity for $M_{\text{tot}} = 100, 200, \text{ and } 300$. The data are obtained as the average over $10^5$ division events in 30 different networks with $K_M = K_R = 200$, $N = 1000$, and $\mu = 0.001$.

FIG. 9: The indices are arranged from Fig. 2(II)(iii) in the main text to categorize into each type A and B.
FIG. 10: The indices are arranged from Fig. 2(III)(iii) in the main text to categorize into each type $A - F$. 