

# Evolution of division of labour in mutualistic symbiosis

Yu Uchiumi\*<sup>1</sup> and Akira Sasaki<sup>1,2</sup>

<sup>1</sup> Department of Evolutionary Studies of Biosystems, The Graduate University for Advanced Studies, SOKENDAI, Hayama, Kanagawa 240–0193, Japan

<sup>2</sup> Evolution and Ecology Program, International Institute for Applied Systems Analysis, Schlosplatz 1, A-2361, Laxenburg, Austria

\* Corresponding author.

E-mail: uchiumi.yu@gmail.com

Phone: +81-46-858-1580, Facsimile: +81-46-858-1544

This article is published in

*Proceedings of the Royal Society B: Biological Sciences* (2020) **287**: 20200669.

<https://doi.org/10.1098/rspb.2020.0669>

## Abstract

Mutualistic symbiosis can be regarded as interspecific division of labour, which can improve the productivity of metabolites and services but deteriorate the ability to live without partners. Interestingly, even in environmentally acquired symbiosis, involved species often rely exclusively on the partners despite the lethal risk of missing partners. To examine this paradoxical evolution, we explored the coevolutionary dynamics in symbiotic species for the amount of investment in producing their essential metabolites, which symbiotic species can share. Our study has shown that, even if obtaining partners is difficult, “perfect division of labour” can be maintained evolutionarily, where each species perfectly specializes in producing one of the essential metabolites so that every member entirely depends on the others for survival, i.e., in exchange for losing the ability of living alone. Moreover, the coevolutionary dynamics shows multistability with other states including a state without any specialization. It can cause evolutionary hysteresis: once perfect division of labour has been achieved evolutionarily when obtaining partners was relatively easy, it is not reverted even if obtaining partners becomes difficult later. Our study suggests that obligate mutualism with a high degree of mutual specialization can evolve and be maintained easier than previously thought.

## Introduction

Division of labour is thought to be developed with the demand for the increased total production in industrialized society [1]. The same idea may apply for differentiated cells and body parts in multicellular organisms [2,3], specialized behavioural groups (castes) in eusocial insects [4], and members of social groups in mammals [5]. Mutualistic symbiosis can also be regarded as interspecific division of labour: in mutualistic symbiosis, each species often specializes in absorbing different nutrients, to produce different metabolites, or to provide different services to complement each other [6]. For example, in a lichen, fungi specialize in providing a stable habitat and minerals to the symbiotic system, and algae specialize in supplying the photosynthetic products as nutrient resources [7].

Such specialization or division of labour can bring mutualistic benefit to both symbiotic partners efficiently but, at the same time, make it difficult for them to live alone when they fail to find the partner [8]. In fact, most species of fungi constituting lichen cannot live without the symbiotic algae [7]. Such obligate interactions are common in mutualistic symbiosis: for example, corals and symbiotic algae, fig and fig wasps, aphids and *Buchnera*, and land plants and arbuscular mycorrhizal fungi (e.g. [9]). The most extreme indispensable interactions are those between eukaryotic cells and mitochondria and plastids.

Interestingly, several species that rely exclusively on symbiotic interactions acquire the partners not from their parents but their surroundings despite the lethal risk of failing to find symbiotic partners. For example, lichen-constituting fungi usually disperse their spores which do not contain their algal partners during sexual reproduction, and the offspring have to find new free-living algae at the patches they colonized [7]. Similarly, more than 70% of species of corals also lay eggs that do not include symbiotic algae [10]. The suggested reason why corals do not inherit their symbionts from parents is that the offspring floating on the sea surface are damaged by strong light and high temperature if they have photosynthetic systems,

or symbiotic algae [8]. In this way, even in mutualistic symbiosis, acquisition of symbionts from the environment can be favoured under developmental or environmental constraints.

On the other hand, it is not clear why the specialization in symbiotic interaction evolves despite the potential risk of missing symbiotic partners in environmentally acquired symbiosis. One may suppose that such risk might be small by certain mechanisms in assembling symbionts from surroundings. For example, corals will attract their symbiotic algae by chemical compounds [11] and green fluorescence [11,12], enhancing the successful initiation of symbiosis. However, although possible mechanisms to obtain partners have been listed (e.g. [13]), it is still empirically unclear how securely they can. Moreover, even theoretically, it remains unresolved how secure should encountering partners be for the evolution of specialization, or division of labour, towards such obligate mutualism. An important question to ask is how and when division of labour evolves in environmentally acquired symbiosis.

Assembling all the symbiotic partners would become difficult as the number of species involved in symbiosis increases, as the potential risk of failing to find them from surroundings. However, symbiotic systems involving more than two species are becoming more common in the literature—“hidden partners” are found in many symbiotic systems, for example, symbiotic yeasts in lichen [14,15] and additional protozoan symbionts in coral [16]. If symbiotic systems with more than two simultaneously interacting symbiotic species are by no means rare, a question then arises, how the number of species involved in symbiosis is determined evolutionarily, and how it is affected by the potential risk of failure in assembling symbiotic partners.

Here, we theoretically explored the evolutionary dynamics describing the development of division of labour in mutualistic symbiosis and revealed when organisms evolve to specialize in and rely exclusively on symbiotic interaction. Since the evolution of

division of labour has been studied in the context of the evolution of multicellularity and sociality [3,17–20], the cost of failing to find partners has not received much attention (but see [3]), which will be shown in this paper to play a pivotal role in environmentally acquired symbiosis. Implication to the evolution of generalists versus specialists with dispersal in heterogeneous environment will be discussed later.

## Model

To examine the evolution of division of labour, we consider a mutualistic system where symbiotic species produce and share several essential metabolites. We assume that the amount of investment in producing the metabolites are the evolutionary traits of each species. We regard that division of labour is achieved in the symbiotic system if some of the metabolites essential to all the symbiotic species are produced only by a limited set of them. By contrast, if each symbiotic species produces all of its essential metabolites by itself, we assume that there is no division of labour with regards to the production of the metabolites. In the following, we investigate the coevolutionary dynamics for the metabolite production among the symbiotic species. Although we focus on metabolite sharing for simplicity, similar arguments should apply for other kinds of division of labour in mutualism by generalizing “production of metabolite” to “performance of tasks”. For example, for plant–pollinator mutualism, where mobility of pollinators is exchanged with energy rewards by plants [6], our model would give insights into the coevolution of nectar production by plants and diet switching to nectar by insects.

### Metabolite production

We first consider a general scenario of mutualistic symbiosis where  $n$  species interchange  $m$  kinds of metabolites (figure 1 is the schematic diagram of our model). We assume that each species has the potential ability to produce all metabolites by allocating their available resources to the production of these metabolites. Let  $f(\xi)$  be the production of a metabolite

when an individual allocates proportion  $\xi$  of its available resources for producing the metabolite. Thus, letting  $x_{ij}$  ( $0 \leq x_{ij} \leq 1$ ) be the amount of investment of species  $i$  in producing metabolite  $j$  ( $i = 1, 2, \dots, n; j = 1, 2, \dots, m$ ), an individual of the species produces  $f(x_{ij})$  units of metabolite  $j$ . The amount of investment  $x_{ij}$  is the traits to be evolved, and the total amount is assumed to be subject to a constraint in each species:  $x_{i1} + x_{i2} + \dots + x_{im} = 1$  for each  $i$ .

We assume that the production function  $f(\xi)$  is an increasing function ( $df/d\xi \geq 0$ ), i.e. more metabolites are produced if the species invests more to it, and that it vanishes when there is no investment to it ( $f(0) = 0$ ). We also assume that the production function is a convex (accelerating) function, i.e.  $d^2f/d\xi^2 > 0$ , meaning that more metabolites are produced per investment if the species specializes more in that metabolite (increased efficiency by specialization). If there is no convexity in  $f$ , then there is no merit to specialize either, and division of labour should never evolve, which is not the situation we want to focus on. More specifically, in our numerical illustrations of the results, we put  $f(\xi) = \varepsilon\xi + \xi^2$  as one of the simplest and generic forms of convex production function with  $f(0) = 0$ , where parameter  $\varepsilon$  ( $> 0$ ) is a positive constant. With an increased  $\varepsilon$ , the nonlinear in  $f$  is blurred, and specialization has less impact on production than when  $\varepsilon$  is small.

### Colonization process

We consider a population consisting of microhabitats to which propagules of each species colonize, and assume that a vacant site is colonized independently by a propagule of species  $i$  ( $i = 1, 2, \dots, n$ ) with probability  $p_i$  ( $0 < p_i \leq 1$ ). A set of species that succeeded in colonizing a particular microhabitat are assumed to engage in a symbiotic system and share their products evenly. By contrast, when only a single species colonizes a vacant site, it lives alone without a partner species and uses only its own production. Thus, denoting by  $s$  a combination of species, probability  $P_n(s)$  that a symbiotic system consists of a species

combination  $s$  is given by  $\prod_{i \in s} p_i \prod_{j \notin s} (1 - p_j)$ . For example, when the number of species is three ( $n = 3$ ), symbiosis with species 1 and 2 but not with species 3 ( $s = \{1, 2\}$ ) occurs with probability  $P_3(\{1, 2\}) = p_1 p_2 (1 - p_3)$ .

### Fitness function

We assume that the fitness of an individual of a species is multiplicative to the production of each essential metabolite, because the lack of any of such metabolites should severely deteriorate the fitness. We also assume that the produced metabolites are shared evenly among all the symbiotic partners. Thus, if an individual of species  $i$  invests its resource in the production of metabolites  $1, 2, \dots, m$  as  $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{im})$ , respectively, its fitness when engaged with a set  $s$  of species in a symbiotic system is

$$w(\mathbf{x}_i, \mathbf{x}_{-i}; s) = \prod_{j=1}^m \frac{\sum_{k \in s} f(x_{kj})}{|s|}, \quad (1)$$

where the argument  $\mathbf{x}_{-i}$  indicates a vector of trait values of all species except species  $i$  and  $|s|$  is the number of species constituting symbiotic system  $s$ , that is, the number of elements of  $s$ . For example, if there are two species and two essential metabolites, an individual's fitness of species 1 is

$$w(\mathbf{x}_1, \mathbf{x}_2; \{1\}) = f(x_{11})f(x_{12}), \quad (2a)$$

when it lives alone ( $s = \{1\}$  and  $|s| = 1$ ), and

$$w(\mathbf{x}_1, \mathbf{x}_2; \{1, 2\}) = \left( \frac{f(x_{11}) + f(x_{21})}{2} \right) \left( \frac{f(x_{12}) + f(x_{22})}{2} \right), \quad (2b)$$

when it is engaged with species 2 ( $s = \{1, 2\}$  and  $|s| = 2$ ), where  $\mathbf{x}_1 = (x_{11}, x_{12})$  and  $\mathbf{x}_2 = (x_{21}, x_{22})$  are traits of species 1 and 2, respectively, and  $\mathbf{x}_{-1} = \mathbf{x}_2$ .

From the above assumptions, the expected fitness of an individual holding a trait  $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{im})$  of species  $i$  is

$$W_i(\mathbf{x}_i, \mathbf{x}_{-i}) = \sum_{s \in S_i} P_n(s) w(\mathbf{x}_i, \mathbf{x}_{-i}; s), \quad (3)$$

where  $S_i$  is a set of all symbiotic systems including species  $i$ , for example, in the case where the number of species is three ( $n = 3$ ),  $S_1 = \{\{1\}, \{1, 2\}, \{1, 3\}, \{1, 2, 3\}\}$ .

### Evolutionary dynamics

By employing the adaptive dynamics framework [21], we then trace the phenotypic evolution of symbiotic specialization, the extent to which is determined by proportions  $x_{ij}$ , of which species  $i$  allocates its resource to producing metabolite  $j$ . In this framework, mutations in either species occur only infrequently, and hence populations of all species are assumed to be monomorphic in most of the times. Under this situation, we can examine the fate of a mutant occurred in each species separately. Thus, the invasion fitness for a mutant of species  $i$  with a set of trait values  $\mathbf{x}'_i = (x'_{i1}, x'_{i2}, \dots, x'_{im})$  in the resident populations with the set of trait values  $\mathbf{x}_i$  (of species  $i$ ) and  $\mathbf{x}_{-i} = \{\mathbf{x}_1, \dots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \dots, \mathbf{x}_n\}$  (of the other species) is described by:

$$F_i(\mathbf{x}'_i | \mathbf{x}_i, \mathbf{x}_{-i}) = W_i(\mathbf{x}'_i, \mathbf{x}_{-i}) - W_i(\mathbf{x}_i, \mathbf{x}_{-i}). \quad (4)$$

Next, we assume that the phenotypic difference between mutants and residents is sufficiently small. A mutant in species  $i$  can invade the population when  $F_i(\mathbf{x}'_i | \mathbf{x}_i, \mathbf{x}_{-i})$  is positive. Therefore, if there were no constraint, the fitness gradient

$$\nabla_{\mathbf{x}'_i} F_i \Big|_{\mathbf{x}'_i = \mathbf{x}_i} = \left( \frac{\partial F_i}{\partial x'_{i1}}, \frac{\partial F_i}{\partial x'_{i2}}, \dots, \frac{\partial F_i}{\partial x'_{im}} \right)_{\mathbf{x}'_i = \mathbf{x}_i}^T \quad (5)$$

determines the direction towards which species  $i$ 's trait  $\mathbf{x}_i$  evolves, where symbol  $T$  indicates transposition. However, since possible mutants are restricted to a constraint surface  $h(\mathbf{x}'_i) = 0$ , where  $h(\mathbf{x}'_i) = \sum_{j=1}^m x'_{ij} - 1$ , we need the elements of  $\nabla_{\mathbf{x}'_i} F_i \Big|_{\mathbf{x}'_i = \mathbf{x}_i}$  along the constraint surface to consider the evolutionary dynamics. According to Ito and Sasaki [22], the corresponding dynamics restricted on the constraint surface is given by integrating the fitness

function  $F_i$  and the constraint function  $h$  into the Lagrange fitness function

$$F_{L_i}(\mathbf{x}'_i | \mathbf{x}_i, \mathbf{x}_{-i}, \lambda_{x_i}) = F_i(\mathbf{x}'_i | \mathbf{x}_i, \mathbf{x}_{-i}) - \lambda_{x_i} \{h(\mathbf{x}'_i) - h(\mathbf{x}_i)\} \quad (6)$$

with parameter  $\lambda_{x_i} = (\nabla_{\mathbf{x}'_i} h(\mathbf{x}'_i) \cdot \nabla_{\mathbf{x}'_i} F_i) / |\nabla_{\mathbf{x}'_i} h(\mathbf{x}'_i)|^2$ , where operator “ $\cdot$ ” indicates the inner product of the two vectors and all derivatives are evaluated at  $\mathbf{x}'_i = \mathbf{x}_i$ . Noting that  $\nabla_{\mathbf{x}} h(\mathbf{x}) = \nabla_{\mathbf{x}} (\sum_{j=1}^m x_j - 1) = (1, 1, \dots, 1)^T$  and  $|\nabla_{\mathbf{x}} h(\mathbf{x})| = \sqrt{m}$  with this constraint, the long-term coevolutionary dynamics for the trait  $j$  of species  $i$  is described by using the gradient of the Lagrange fitness function (6) as

$$\begin{aligned} \dot{x}_{ij} &= \sigma_i \left( \nabla_{\mathbf{x}'_i} F_{L_i} \Big|_{\mathbf{x}'_i = \mathbf{x}_i} \right)_j \\ &= \sigma_i \left\{ \frac{\partial F_i(\mathbf{x}'_i, \mathbf{x}_i | \mathbf{x}_{-i})}{\partial x'_{ij}} \Big|_{\mathbf{x}'_i = \mathbf{x}_i} - \frac{1}{m} \sum_{k=1}^m \frac{\partial F_i(\mathbf{x}'_i, \mathbf{x}_i | \mathbf{x}_{-i})}{\partial x'_{ik}} \Big|_{\mathbf{x}'_i = \mathbf{x}_i} \right\}, \end{aligned} \quad (7)$$

( $i = 1, \dots, n; j = 1, \dots, m$ ), where the dot on a variable indicates the time derivative in a long-term evolutionary time scale and  $\sigma_i (> 0)$  is a parameter of species  $i$  that determines the speed of their evolution, which consists of the rate of mutation and the variance of its phenotypic effect.

## Results

### Two-species two-metabolite model

In this section, we show the coevolutionary dynamics of the simplest case where two species (species 1 and 2) interchange two kinds of metabolites (metabolite 1 and 2). In the case, the expected fitness of each species (equation (3)) is given as follows:

$$W_1(\mathbf{x}_1, \mathbf{x}_2) = p_1 p_2 \left( \frac{f(x_{11}) + f(x_{21})}{2} \right) \left( \frac{f(x_{12}) + f(x_{22})}{2} \right) + p_1 (1 - p_2) f(x_{11}) f(x_{12}), \quad (8a)$$

$$W_2(\mathbf{x}_1, \mathbf{x}_2) = p_1 p_2 \left( \frac{f(x_{11}) + f(x_{21})}{2} \right) \left( \frac{f(x_{12}) + f(x_{22})}{2} \right) + (1 - p_1) p_2 f(x_{21}) f(x_{22}), \quad (8b)$$

where  $\mathbf{x}_i = (x_{i1}, x_{i2})^T$  is traits of species  $i$ . From equation (7), the coevolutionary dynamics

is:

$$\dot{x}_{ik} = \sigma_i \left( \frac{\partial W_i(\mathbf{x}'_i, \mathbf{x}_j)}{\partial x'_{ik}} \Big|_{x'_i=x_i} - \frac{1}{2} \sum_{l=1}^2 \frac{\partial W_i(\mathbf{x}'_i, \mathbf{x}_j)}{\partial x'_{il}} \Big|_{x'_i=x_i} \right). \quad (9)$$

( $i, j, k = 1, 2; i \neq j$ ). In the case of two species ( $n = 2$ ) and two metabolites ( $m = 2$ ), the coevolutionary dynamics (9) can be described as two-dimensional dynamics with the simple transformation of variables such that  $y_1 = x_{11}$  and  $y_2 = x_{21}$  (with  $x_{12} = 1 - y_1$ ,  $x_{22} = 1 - y_2$ ):

$$\begin{aligned} \dot{y}_1 = & \frac{\sigma_1}{2} p_1 p_2 \left\{ \left( \frac{f'(y_1)}{2} \right) \left( \frac{f(1-y_1) + f(1-y_2)}{2} \right) - \left( \frac{f(y_1) + f(y_2)}{2} \right) \left( \frac{f'(1-y_1)}{2} \right) \right\} \\ & + \frac{\sigma_1}{2} p_1 (1-p_2) \{f'(y_1)f(1-y_1) - f(y_1)f'(1-y_1)\}, \end{aligned} \quad (10a)$$

$$\begin{aligned} \dot{y}_2 = & \frac{\sigma_2}{2} p_1 p_2 \left\{ \left( \frac{f'(y_2)}{2} \right) \left( \frac{f(1-y_1) + f(1-y_2)}{2} \right) - \left( \frac{f(y_1) + f(y_2)}{2} \right) \left( \frac{f'(1-y_2)}{2} \right) \right\} \\ & + \frac{\sigma_2}{2} (1-p_1) p_2 \{f'(y_2)f(1-y_2) - f(y_2)f'(1-y_2)\}, \end{aligned} \quad (10b)$$

where  $f'(\xi) = df(\xi)/d\xi$ .

We first show the simplest results where the probabilities of encountering partners are the same for all species ( $p_1 = p_2 = p$ ) and the production function does not have any linear terms ( $f(\xi) = \xi^2$ ). We then touch on more general results later when these simplifying assumptions are relaxed. Figure 2 shows typical coevolutionary dynamics in the simplest case.

*Perfect Division of Labour (PDL): the only coevolutionary outcome with a high chance of finding partners*

First, if the probability of encountering symbiotic partners  $p$  is higher than  $2/3$  (the right panel of figure 2), the coevolutionary dynamics only attain a state which we call *Perfect-division-of-labour* (PDL), where each species perfectly specializes in producing one of metabolite 1 and 2 so as to complement each other, corresponding to the equilibria  $(y_1, y_2) = (1, 0)$  and  $(0, 1)$  in the trait space (equilibria labelled ‘‘P’’ in figure 2). Under PDL, both

species rely exclusively on their symbiotic interaction and cannot live alone. As  $f'(0) = 0$ ,  $\dot{y}_i$  is always positive when  $y_i = 1$  and always negative when  $y_i = 0$ , indicating that these PDL equilibria are always stable however difficult it is obtaining partners (i.e. even when  $p$  is extremely low).

*Both Jack-of-All-Trade (JAT) and PDL maintained with intermediately high chance of finding partners*

Second, if the probability of encountering partners  $p$  becomes smaller than  $2/3$  but is still larger than  $4/11$  (the middle panel of figure 2), the coevolutionary dynamics becomes bi-stable between PDL and a state which we call *Jack-of-all-trades* (JAT), where both species do not specialize at all in their metabolite production, corresponding to the point  $(y_1, y_2) = (1/2, 1/2)$  in the trait space (equilibrium labelled “J” in figure 2). Under JAT, living with a partner does not differ at all from being free-living in terms of their fitness. The JAT equilibrium  $(y_1, y_2) = (1/2, 1/2)$  is stable when  $p < 2/3$ , where Jacobi matrix of equation (10) on JAT equilibrium has only eigenvalues with a negative real part.

*Unequal division of labour: Perfect in one species and partial in the other (HPDL) participates in coevolutionarily stable states with an even lower chance of finding partners*

Finally, if the probability of encountering partners  $p$  is smaller than  $4/11$  (the left panel of figure 2), a state which we call *Hemi-perfect-division-of-labour* (HPDL) is also stable in addition to PDL and JAT. In HPDL, one species completely specializes in the symbiotic interaction, but the other retains the ability to survive in a free-living state by partially investing the production of both metabolites. In this simplest case, it corresponds to the points  $(\hat{Y}_0, 0)$ ,  $(\hat{Y}_1, 1)$ ,  $(0, \hat{Y}_0)$ , and  $(1, \hat{Y}_1)$ , where  $\hat{Y}_0 = (3 - \sqrt{(4 - 11p)/(4 - 3p)})/4$  and  $\hat{Y}_1 = (1 + \sqrt{(4 - 11p)/(4 - 3p)})/4$ , (equilibria labelled “H” in figure 2). These HPDL equilibria are biologically feasible when  $p < 4/11$  and stable if they are feasible. The existence of a stable

HPDL equilibrium indicates that asymmetric dependency can evolve spontaneously even if there is no difference between species in the probability of encountering partners or in productivity of metabolites.

*Fluctuation in  $p$  can lead to irreversible transition towards perfect division of labour*

Since the coevolutionary dynamics show multistability, as shown above, it causes hysteresis along the fluctuation of the probability of encountering partners  $p$  (figure 2). For example, once the evolutionary transition from JAT to PDL is achieved along with the increase in probability  $p$ , the PDL state can be maintained evolutionarily even if probability  $p$  decreases again, rather than returning to the JAT state. In particular, since PDL is always maintained once it is achieved, any coevolutionary dynamics will eventually attain PDL through the fluctuation of probability  $p$  if it is sufficiently large on a long-timescale --- evolution towards further division of labour is a “contingent irreversibility” [23].

General convex production function  $f$  and nonidentical  $p$ 's

Even in the more general case where the probabilities of encountering partners  $p_1$  and  $p_2$  can differ between species or the production function  $f(\xi)$  may have a linear part, the dependency of the stability of each state on the probability of encountering partners, as well as the existence of multistability and hysteresis, remains qualitatively the same as the above results for the simplest case (electronic supplementary material, S1, where  $f(\xi) = \varepsilon\xi + \xi^2$  is assumed). However, the range of the stability of PDL becomes narrower by introducing the linear term in the production function  $f(\xi)$ . Indeed, the condition for the stability of PDL is that both  $p_1$  and  $p_2$  are greater than a threshold  $p_c$  (see electronic supplementary material, S2, for the derivation):

$$p_1, p_2 > p_c = \frac{4f'(0)}{f'(1) + 3f'(0)}. \quad (11)$$

That is, if there is a nonzero linear term ( $f'(0) > 0$ ), or if allocating an infinitesimally small amount of resource to a metabolite that has not been produced at all can improve its production, then PDL is no longer stable when finding symbiotic partners is difficult in one of the species (either  $p_1 < p_c$  or  $p_2 < p_c$ ). Thus, generally speaking, any coevolutionary trajectory can swing between PDL, JAT, and HPDL through the fluctuation of probability  $p$  on a long timescale (electronic supplementary material, S1). However, if the fluctuation of probability  $p$  does not reach significantly low values (strictly speaking,  $p$  does not reach as low as  $p_c = 4\varepsilon/(4\varepsilon + 2)$ ), which would be plausible in actual symbiotic interaction, then coevolutionary dynamics will eventually attain PDL through the fluctuation of  $p$  (electronic supplementary material, S1) as we argued in the simplest case (figure 2).

In contrast with PDL, the condition for the stability of JAT is not so different from the simplest case of  $\varepsilon = 0$  and  $p_1 = p_2$  (electronic supplementary material, S1). For the general case, JAT equilibrium  $(x_{11}, x_{21}) = (1/2, 1/2)$  is stable when (see electronic supplementary material, S2, for the derivation):

$$\frac{p_1 + p_2 - p_1 p_2}{(2 - p_1)(2 - p_2)} < 1 - \frac{f(\frac{1}{2})f''(\frac{1}{2})}{f'(\frac{1}{2})^2}. \quad (12)$$

Note that, as  $f(\xi)$  is a positive and convex function, the right side is always in the range of zero to one. The left side is also in the range of zero to one, and it increases as  $p_1$  or  $p_2$  increases. Therefore, JAT equilibrium becomes unstable when finding symbiotic partners is sufficiently easy (when either  $p_1$  or  $p_2$  is large so that the inequality (12) is violated).

The stability conditions of HPDL cannot be solved explicitly in the general case, however, numerical study shows that it tends to be stable for small  $p_i$ 's (electronic supplementary material, S1)

In any case, the coevolutionary dynamics can attain PDL, JAT, and HPDL and show multistability and hysteresis. Our results suggest that, even if obtaining partners is difficult,

obligate mutualism can be achieved and maintained easier than one may expect. Besides, whether symbiotic interaction evolves to be obligate and whether division of labour is achieved through the coevolution depend on the initial conditions of coevolutionary dynamics as well as the parameters that affect the efficiency of division of labour and difficulty in obtaining partners.

### More than two symbiotic species: $n$ -species $m$ -metabolite model

As the number of symbiotic species  $n$  and the number of metabolites  $m$  increase, the number of possible equilibria of the model rises rapidly (see Simulation results). Since it is difficult to examine all the equilibria of the dynamics analytically, we focus only on the stability of PDL and JAT equilibria. Figure 3a and 3b, respectively, show parameter regions where PDL and JAT are locally stable in dimension  $n = m = 2, 3, 4$ , and 5 (see electronic supplementary material, S2, for the analysis). Note that for comparing and illustrating high-dimensional cases, it is assumed in the figure that the number of species is the same as that of metabolites and that  $p_i$  and  $\sigma_i$  are the same for all the species. Interestingly, although we expected that assembling all the species into PDL would become difficult as the number of species increases, we found that the more species are involved, the more extended is the region where PDL is stable: PDL becomes easier to be maintained evolutionarily for large symbiotic systems with many species. In addition, JAT equilibrium tends to be unstable in the systems with many species and many metabolites.

### *Simulation results*

Since the coevolutionary dynamics (equation 7) tend to show multistability as we have seen above, examining local stability of all the possible equilibria becomes intractable as the numbers of symbiotic species and essential metabolites increase in our model. We, therefore, examined by numerical simulations to assess the likelihood with which perfect division of labour (PDL) or similar highly specialized functional partitioning among symbiotic species is

attained.

To easily guarantee the positivity constraint in our numerical simulations, we replaced  $x_{ij}$  in equation (7) by  $z_{ij}^2$  and calculated not the coevolutionary dynamics of trait  $x_{ij}$  directly but that of trait  $z_{ij}$  alternatively. The reason for adopting this transformation is that there is no, in our knowledge, simple and rigorous way to restrict the coevolutionary trajectory not to cross the simplex boundaries,  $x_{ij} = 0$  or  $x_{ij} = 1$ , where selection pressure on trait  $x_{ij}$  sometimes works in such a way as to push it out from the feasible region.

Figure 4 shows the fraction, for randomly varied initial conditions, eventually fell into perfect division of labour (PDL:  $x_{ij} = 1$  if  $i = j$  and  $x_{ij} = 0$  if  $i \neq j$  for  $i, j = 1, 2, \dots, n$  shown in black) or no division of labour (JAT:  $x_{ij} = 1/n$  for  $i, j = 1, 2, \dots, n$  shown in white) in the coevolutionary dynamics equation (7) plotting against the number of symbiotic species and essential metabolites varied from  $n = m = 2$  to 5. The frequency by which a specific pattern of specialization is attained coevolutionarily for randomly varying initial patterns can be regarded as the relative size of its “basin of attraction”. The measure is called “basin stability” and can be interpreted not only as the likelihood with which an equilibrium is attained but also as *how the likelihood of return to the state after any random—possibly non-small—perturbation* [24].

As the number of species and metabolites increases, the frequency of the evolutionary trajectories that fall into PDL decreases. However, the basin stability of PDL stays high even in high-dimensional cases if the probability of encountering partners is not too low (figure 4a). With  $p = 0.7$  and  $n = 5$ , the probability that all five species constituting the symbiotic system will be found in a colonizing site is only 17%. Nonetheless, the five-species symbiotic system evolved in around three-quarters of cases to a PDL where each of five essential metabolites is produced only by one specialized symbiotic species. Here, the fraction evolving to PDL decreases from nearly 100% in  $n = 2$  to around 75% in  $n = 5$ . However, when we

look at the second most frequent equilibrium patterns of specialization other than PDL (which is the most frequent), we found that in almost all the other cases, the system evolved to a redundant, but nearly perfect division of labour (figure 5): for  $n = m = 5$ ,

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} & x_{15} \\ x_{21} & x_{22} & x_{23} & x_{24} & x_{25} \\ x_{31} & x_{32} & x_{33} & x_{34} & x_{35} \\ x_{41} & x_{42} & x_{43} & x_{44} & x_{45} \\ x_{51} & x_{52} & x_{53} & x_{54} & x_{55} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \end{pmatrix}, \quad (13a)$$

for  $n = m = 4$ ,

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} \\ x_{21} & x_{22} & x_{23} & x_{24} \\ x_{31} & x_{32} & x_{33} & x_{34} \\ x_{41} & x_{42} & x_{43} & x_{44} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix}, \quad (13b)$$

and for  $n = m = 3$ ,

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1/2 & 1/2 \end{pmatrix}. \quad (13c)$$

Such redundancy can mitigate the difficulty in assembling all the partners in the high-dimensional case and would balance its risk with the benefit of specialization.

If the probability of encountering symbiotic species is low, the frequency that PDL is attained through the coevolutionary dynamics becomes progressively lower as the number of species increases (figure 4b). However, the most frequent patterns of specialization attained in the coevolutionary dynamics are nearly perfect specialization, although there is more redundancy in metabolite production between species than in the cases with higher  $p$  (figure 5; the extents of specialization and redundancy of the attained states are summarized in electronic supplementary material, S3). For example, when  $n = 5$  and  $p = 0.3$ , the most frequent pattern of specialization attained by the coevolutionary dynamics is:

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} & x_{14} & x_{15} \\ x_{21} & x_{22} & x_{23} & x_{24} & x_{25} \\ x_{31} & x_{32} & x_{33} & x_{34} & x_{35} \\ x_{41} & x_{42} & x_{43} & x_{44} & x_{45} \\ x_{51} & x_{52} & x_{53} & x_{54} & x_{55} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \end{pmatrix}. \quad (14)$$

In this example, although all five species rely exclusively on symbiosis, a suitable choice of three species from each functionally distinct group (one from species 1 or 2, one from species 3 or 4, and species 5) is sufficient for each of three species to live. In this example, since species 5's risk of failing to gather partners is the lowest, it enjoys the highest fitness among these species. It means that although all species share symbiotic benefits evenly in each interaction, the disparity arises in their eagerness to encounter partners. In any way, although PDL tends not to be established with a large number of interacting species, almost all species consisting of the predominantly attained states still specialize completely in and rely exclusively on symbiotic interactions even if obtaining partners is not easy. As a result, redundant interactions evolve frequently, and it can mitigate the difficulty in assembling all the partners in symbiotic systems consisting many species.

The horizontal axis in figure 5 represents the Euclidian distance between the equilibrium  $(x_{11}, \dots, x_{1n}, x_{21}, \dots, x_{2n}, \dots, x_{n1}, \dots, x_{nn})$  attained in the coevolutionary dynamics and JAT equilibrium  $(1/n, \dots, 1/n)$ . Since the attained equilibria represent the coevolutionarily attained pattern of specialization, the distances from JAT equilibrium, therefore, quantify the degree of specialization in essential metabolite production in symbiotic species. When we randomly assign the initial specialization pattern in simplex trait space in our simulation, this distance measure is not uniformly distributed initially (see grey histogram in figure 5). This is due to geometry of high-dimensional space: in high-dimensional space, the volume of its peripheral region is much larger than that of its central region[25]. We excluded this bias from initial distance distribution in sampling from our high-dimensional simplex trait space (electronic supplementary material, S4), but fail to show any qualitative difference from our results shown in figure 4 or 5. In addition, we confirmed that making the merit of specialization less remarkable, or increasing parameter  $\varepsilon$  of production function  $f(\xi) = \varepsilon\xi + \xi^2$ , affects the results only slightly (electronic supplementary material, S5).

## Discussion

Mutualistic symbiosis can be regarded as interspecific division of labour. In this study, we theoretically explored the coevolutionary dynamics of division of labour in mutualistic symbiosis and revealed when organisms evolve to specialize in the symbiotic interaction. During the analysis, we focused on two states: *perfect-division-of-labour* (PDL) where each species perfectly specializes in producing one of the metabolites and relies exclusively on symbiotic interactions; and *Jack-of-all-trades* (JAT) where each species does not specialize at all. Our study has shown that, under the large advantage of specialization, PDL can be stable even if obtaining partners is difficult, and that it often shows multistability and hysteresis with other states including JAT. Moreover, although assembling all species becomes difficult when the number of species increases, the more species are involved, the easier PDL state can be maintained evolutionarily. In such cases, the coevolutionary dynamics frequently attains to redundant states with nearly perfect division of labour. In this way, our study can explain why several species rely exclusively on symbiotic interactions despite the lethal risk of failing to encounter partners. Our study also suggests that whether symbiotic interaction evolves to be obligate, that is, whether division of labour is achieved through coevolution, depends on initial conditions of the coevolution as well as on the cost and benefit of specialization. In other words, even if obtaining partners is difficult, obligate mutualism can be achieved and maintained easier than one may expect.

Our study suggests that in mutualistic symbiosis, the evolution of specialization by division of labour depends not only on the ease of acquiring symbiotic partners but also on the evolutionary history, or the initial conditions. Intuitively, one may expect that a symbiotic species specializes more in the interaction if it is easier to encounter partners. However, our model revealed that the coevolution can lead to a PDL in a large range of  $p$ , the probability of encountering a symbiotic partner. Even when the chance of meeting symbiotic partners is

remote, PDL can be jointly attracting with the JAT state—to which state the coevolution leads depends on the initial traits of each species. It would explain why obligate symbiosis is stably maintained even in environmentally acquired mutualism, for example, coral and symbiotic algae [8], lichen [7], and tubeworm and symbiotic bacteria [26]. Our model also shows that it is difficult to achieve moderate specialization, although perfect specialization must pay an expensive toll due to the lack of free-living ability. This might correspond to the fact that mutualism in which essential nutrients are exchanged tends to be obligate [27].

Moreover, the robustness of bi-stability in our model might explain that mutualism tends to be an evolutionary dead-end [28,29] or contingent irreversible [23]. In our model, once the evolutionary transition from JAT to PDL is achieved by the increased chance of encountering partners, the PDL state can be maintained evolutionarily without returning to the JAT state even if encountering partners becomes difficult again. It suggests that once obligate mutualism is established, its breakdown tends to be difficult. In previous studies, the robustness of mutualism is often explained by the difficulty of genetically restoring free-living ability [28]. However, our model suggests that even if such recovery is easy, the selective force can trap symbiotic interactions in the evolutionary dead-end, obligate mutualism.

Our study also provides insight into how the number of species involved in an obligate symbiosis is determined. In our model, as the number of involved species increases, possible coevolutionary outcomes are diversified. As a result, the more species are involved, the easier PDL becomes to be maintained evolutionarily once established, although the more difficult it becomes to be achieved. In this way, what state is achieved will depend strongly on the initial conditions of the coevolution. Interestingly, even in the coevolutionarily attained states other than PDL, all species also tended to rely exclusively on symbiotic interaction. Indeed, they spontaneously formed several functional groups like “guilds”, where species in a group hold the same pattern of metabolite production. It reflects the fact that, with the difficulty in

assembling all species, such redundancy helps improve their chance of survival even if not all species are assembled. Moreover, the redundancy can help evolutionary replacements of symbiotic partners within a group of species with the same functional traits. That the redundant patterns arise very frequently in our model might provide an insight into the endosymbiont replacements frequently observed, for example, in the symbiosis between several clades of hemipteran insects and endosymbiotic bacteria [30], weevils and endosymbiotic bacteria [31], legumes and root-nodule bacteria [32], and dinoflagellates and plastids [33].

The robustness of multistability found in our model for the coevolution of division of labour in mutualism is due to the improved efficiency in producing metabolites with more specialization in investment towards them. Such property of production in general can be described by an accelerating function of invested resources [3], as we assumed as  $d^2 f(\xi)/d\xi^2 > 0$ , and then benefit for each species engaged in a symbiotic system tends to be maximized when specialization is perfect. Comparing this with the fact that the benefit should be maximized under no specialization when each species lives alone, the combined benefit with a given chance of finding symbiotic partners should be locally maximized either at the perfect specialization or at non-specialization, and bottomed out when specialization is intermediate. Thus, our qualitative results by no means rely on a specific form of productivity–investment relationship or symmetry of metabolites and species we assumed. Moreover, our results are not limited to division of labour in metabolite exchange but can be generalized to other kinds of division of labour, for example, the exchange of mobility for nutritional compensation in plant–pollinator mutualism [6], as we touched in Model.

Our study can also give insights into the classical problem of the evolution of specialist versus generalist in organisms that disperse to heterogeneous habitats in their life stage (e.g.[34]). Heterogeneous habitats an organism faces in a dispersed patch, e.g. hosts with variable resistance for an infecting parasite [35], would correspond to a random combination

of a symbiotic party in our model. Local adaptation to a specific patch would then evolve if it acceleratingly increases the productivity from the patch, and if the probability of dispersing to the right patches is sufficiently large, enough to compensate for low fitness when dispersed to mismatched patches. Increasing return to specific investment would also promote multistability in the evolution of specialization as in our model.

Although not explicitly analysed in the present paper, all such factors including as unfair redistribution of metabolite between symbiotic species, the ability of co-colonization of symbionts (vertical transmission), and phenotypic plasticity for specialized/non-specialized production affect the evolution of division of labour. With our assumption of equal redistribution of products among symbiotic species, free riding in mutualistic interactions never occurs. The pattern of redistribution itself would evolve as a direct reciprocity in response to a partner's investments. If the tendency to co-colonization evolves, perfect division of labour would evolve easier with a higher chance of assembling symbiotic partners. The division of labour would evolve easier in vertically acquired mutualism if a phenotypically plastic response is allowed to switch between specialized and non-specialized metabolite productions according to the availability of symbiotic partners.

Mutualistic symbiosis can be regarded as interspecific division of labour, and specialization in the interaction will improve the productivity of metabolites at the expense of survivorship when finding partners fails. Our study suggests that for the evolution of division of labour in environmentally acquired mutualism, its benefit of increased productivity must be weighed more than the risk of failure in finding partners. There, however, is inherent multistability in coevolutionary dynamics that leads to evolutionary hysteresis and contingent irreversibility towards division of labour: high specialization in mutualism (obligate mutualism) can then be maintained evolutionarily even if meeting partners is difficult.

## Acknowledgements

We are grateful to Drs Ryosuke Iritani, Hiroshi Ito, Hisashi Ohtsuki, and Yohey Terai for discussions and helpful comments. This study was financially supported by the Graduate University for Advanced Studies, SOKENDAI.

## References

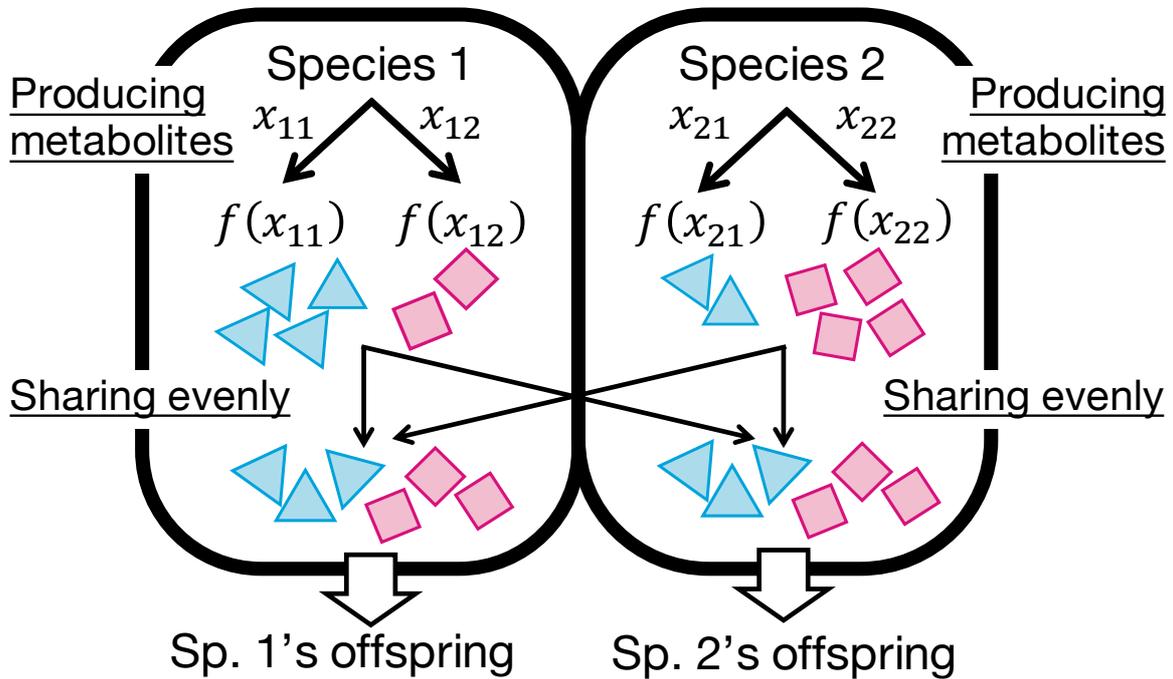
1. Smith A. 1776 *An inquiry into the nature and causes of the wealth of nations*. London: W. Strahan and T. Cadell.
2. Michod RE. 2007 Evolution of individuality during the transition from unicellular to multicellular life. *Proc. Natl. Acad. Sci.* **104**, 8613–8618. (doi:10.1073/pnas.0701489104)
3. Rueffler C, Hermisson J, Wagner GP. 2012 Evolution of functional specialization and division of labor. *Proc. Natl. Acad. Sci.* **109**, E326–E335. (doi:10.1073/pnas.1110521109)
4. Beshers SN, Fewell JH. 2001 Models of division of labor in social insects. *Annu. Rev. Entomol.* **46**, 413–440. (doi:10.1146/annurev.ento.46.1.413)
5. Clutton-Brock TH. 2006 Cooperative breeding in mammals. In *Cooperation in Primates and Humans: Mechanisms and Evolution* (eds PM Kappeler, CP van Schaik), pp. 173–190. Berlin, Heidelberg: Springer. (doi:10.1007/3-540-28277-7\_10)
6. Leigh Jr. EG. 2010 The evolution of mutualism. *J. Evol. Biol.* **23**, 2507–2528. (doi:10.1111/j.1420-9101.2010.02114.x)
7. Nash TH. 2008 *Lichen biology*. Second Edition. New York: Cambridge University Press.
8. Hartmann AC, Baird AH, Knowlton N, Huang D. 2017 The paradox of environmental symbiont acquisition in obligate mutualisms. *Curr. Biol.* **27**, 3711–3716.e3. (doi:10.1016/j.cub.2017.10.036)
9. Paracer S, Ahmadjian V. 2000 *Symbiosis: an Introduction to biological associations*. Second Edition. Oxford, New York: Oxford University Press.
10. Fadlallah YH. 1983 Sexual reproduction, development and larval biology in scleractinian corals. *Coral Reefs* **2**, 129–150. (doi:10.1007/BF00336720)
11. Yamashita H, Suzuki G, Kai S, Hayashibara T, Koike K. 2014 Establishment of coral–algal symbiosis requires attraction and selection. *PLOS ONE* **9**, e97003. (doi:10.1371/journal.pone.0097003)
12. Hollingsworth LL, Kinzie RA, Lewis TD, Krupp DA, Leong J-AC. 2005 Phototaxis of motile zooxanthellae to green light may facilitate symbiont capture by coral larvae. *Coral Reefs* **24**, 523–523. (doi:10.1007/s00338-005-0063-8)
13. Bright M, Bulgheresi S. 2010 A complex journey: transmission of microbial symbionts. *Nat. Rev. Microbiol.* **8**, 218–230. (doi:10.1038/nrmicro2262)
14. Spribille T *et al.* 2016 Basidiomycete yeasts in the cortex of ascomycete macrolichens. *Science* **353**, 488–492. (doi:10.1126/science.aaf8287)

15. Tuovinen V, Ekman S, Thor G, Vanderpool D, Spribille T, Johannesson H. 2019 Two basidiomycete fungi in the cortex of wolf lichens. *Curr. Biol.* (doi:10.1016/j.cub.2018.12.022)
16. Kwong WK, del Campo J, Mathur V, Vermeij MJA, Keeling PJ. 2019 A widespread coral-infecting apicomplexan with chlorophyll biosynthesis genes. *Nature* **568**, 103–107. (doi:10.1038/s41586-019-1072-z)
17. Duarte A, Pen I, Keller L, Weissing FJ. 2012 Evolution of self-organized division of labor in a response threshold model. *Behav. Ecol. Sociobiol.* **66**, 947–957. (doi:10.1007/s00265-012-1343-2)
18. Iaroslav I, Martin A, Michael D. 2012 Division of labour and the evolution of multicellularity. *Proc. R. Soc. B Biol. Sci.* **279**, 1768–1776. (doi:10.1098/rspb.2011.1999)
19. Vásárhelyi Z, Meszéna G, Scheuring I. 2015 Evolution of heritable behavioural differences in a model of social division of labour. *PeerJ* **3**, e977. (doi:10.7717/peerj.977)
20. Mullon C, Lehmann L. 2019 An evolutionary quantitative genetics model for phenotypic (co)variances under limited dispersal, with an application to socially synergistic traits. *Evolution* **73**, 1695–1728. (doi:10.1111/evo.13803)
21. Geritz SAH, Kisdi É, Meszéna G, Metz JAJ. 1998 Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol.* **12**, 35–57. (doi:10.1023/A:1006554906681)
22. Ito H, Sasaki A. 2016 Evolutionary branching under multi-dimensional evolutionary constraints. *J. Theor. Biol.* **407**, 409–428. (doi:10.1016/j.jtbi.2016.07.011)
23. Maynard Smith J, Szathmáry E. 1995 *The major transitions in evolution*. Oxford, New York: Oxford University Press.
24. Menck PJ, Heitzig J, Marwan N, Kurths J. 2013 How basin stability complements the linear-stability paradigm. *Nat. Phys.* **9**, 89–92. (doi:10.1038/nphys2516)
25. Bishop CM. 2006 *Pattern recognition and machine learning*. New York: Springer-Verlag.
26. Nussbaumer AD, Fisher CR, Bright M. 2006 Horizontal endosymbiont transmission in hydrothermal vent tubeworms. *Nature* **441**, 345–348. (doi:10.1038/nature04793)
27. Fisher RM, Henry LM, Cornwallis CK, Kiers ET, West SA. 2017 The evolution of host-symbiont dependence. *Nat. Commun.* **8**, 15973. (doi:10.1038/ncomms15973)
28. Moran NA, Wernegreen JJ. 2000 Lifestyle evolution in symbiotic bacteria: insights from genomics. *Trends Ecol. Evol.* **15**, 321–326. (doi:10.1016/S0169-5347(00)01902-9)
29. Sachs JL, Simms EL. 2006 Pathways to mutualism breakdown. *Trends Ecol. Evol.* **21**, 585–592. (doi:10.1016/j.tree.2006.06.018)

30. Sudakaran S, Kost C, Kaltenpoth M. 2017 Symbiont acquisition and replacement as a source of ecological innovation. *Trends Microbiol.* **25**, 375–390. (doi:10.1016/j.tim.2017.02.014)
31. Lefèvre C, Charles H, Vallier A, Delobel B, Farrell B, Heddi A. 2004 Endosymbiont Phylogenesis in the Dryophthoridae Weevils: Evidence for Bacterial Replacement. *Mol. Biol. Evol.* **21**, 965–973. (doi:10.1093/molbev/msh063)
32. Martínez-Romero E. 2009 Coevolution in Rhizobium-Legume Symbiosis? *DNA Cell Biol.* **28**, 361–370. (doi:10.1089/dna.2009.0863)
33. Saldarriaga JF, Taylor FJR, Keeling PJ, Cavalier-Smith T. 2001 Dinoflagellate Nuclear SSU rRNA Phylogeny Suggests Multiple Plastid Losses and Replacements. *J. Mol. Evol.* **53**, 204–213. (doi:10.1007/s002390010210)
34. Levins R. 1968 *Evolution in Changing Environments*. Princeton University Press.
35. Antonovics J, Boots M, Ebert D, Koskella B, Poss M, Sadd BM. 2013 The Origin of Specificity by Means of Natural Selection: Evolved and Nonhost Resistance in Host–Pathogen Interactions. *Evolution* **67**, 1–9. (doi:10.1111/j.1558-5646.2012.01793.x)

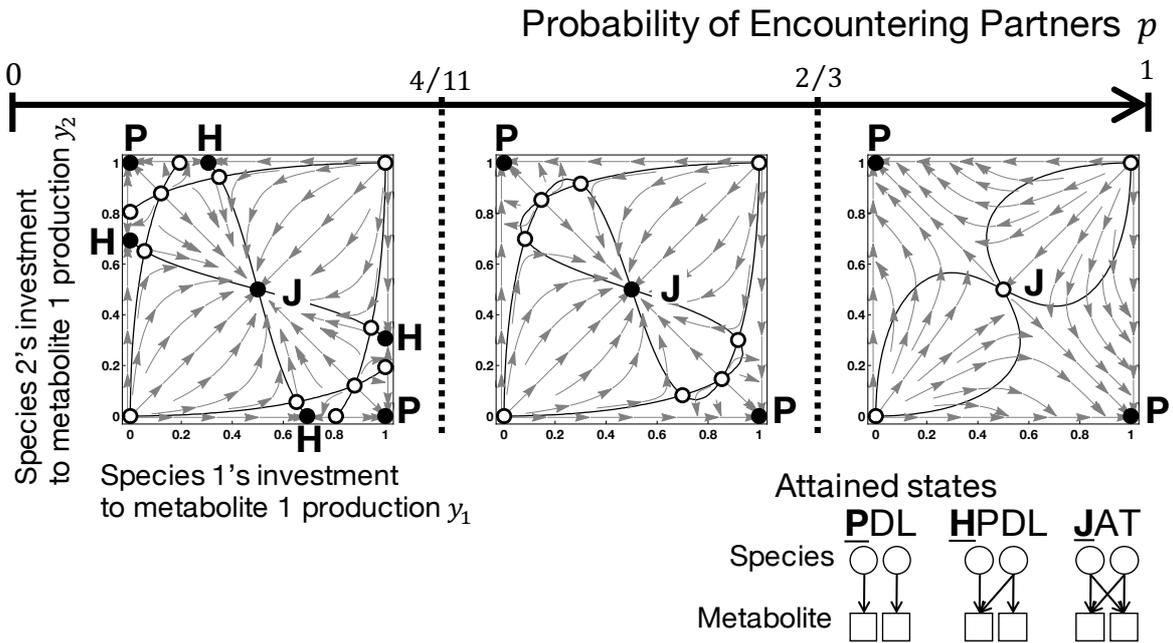
## Figures

Figure 1



**Figure 1.** The schematic diagram of our model when numbers of species and metabolite are two. Individuals of both species are engaged in symbiosis with probability  $p_1 p_2$ . Species 1 invests the amount  $x_{11}$  to producing metabolite 1 (blue triangles) and  $x_{12}$  to producing metabolite 2 (red squares), while the species 2 invests the amount  $x_{21}$  to producing metabolite 1 and  $x_{22}$  to producing metabolite 2. The total amount of investment is assumed to be subject to a constraint:  $x_{11} + x_{12} = 1$  and  $x_{21} + x_{22} = 1$ . With the amount  $\xi$  of investment for producing a metabolite, its production is given by a production function  $f(\xi)$ . The produced metabolites are shared evenly between species 1 and 2. The fitness of an individual of each species is assumed to be given by multiplying the total amount of essential metabolites produced in the symbiotic system:  $\{(f(x_{11}) + f(x_{21}))/2\}\{(f(x_{12}) + f(x_{22}))/2\}$ .

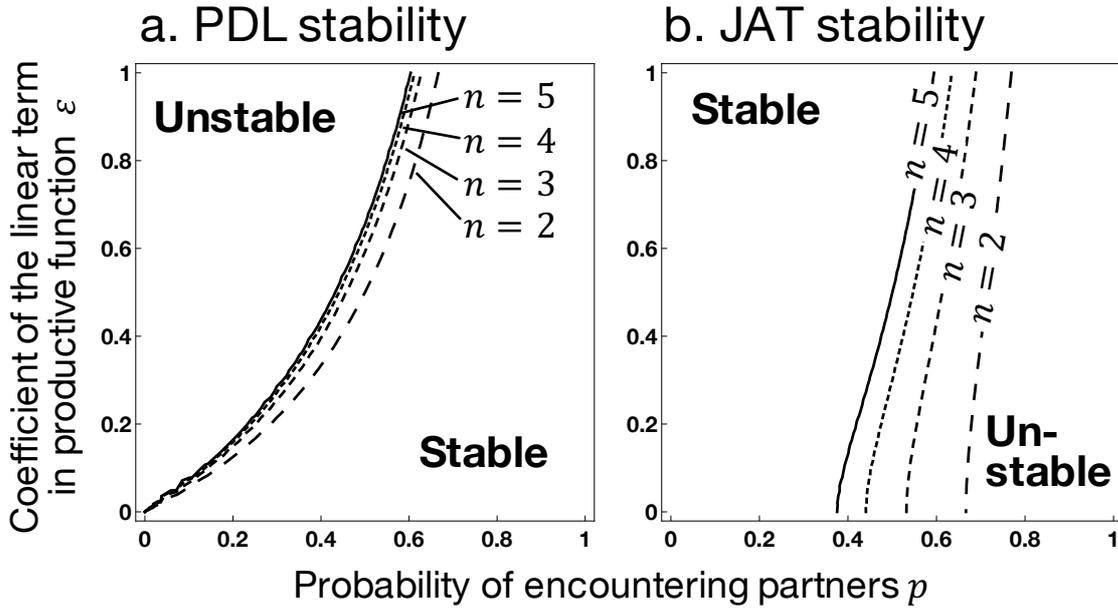
Figure 2



**Figure 2.** Typical phase portraits of the coevolutionary dynamics of two-species two-metabolites model. (Left panel) When the chance of obtaining symbiotic partners is remote ( $p$  being less than four eleventh;  $p = p_1 = p_2 = 0.35$  in the panel), the coevolutionary dynamics shows multistability of *perfect-division-of-labour* (PDL; labelled “P”), *hemi-perfect-division-of-labour* (HPDL; labelled “H”), and *Jack-of-all-trades* (JAT; labelled “J”) states. (Middle panel) When the chance of obtaining partners is intermediate ( $4/11 < p < 2/3$ ;  $p = 0.4$  in the panel), HPDL becomes unstable, leaving only bi-stability between PDL and JAT. (Right panel) When the chance of obtaining partners is higher than two thirds ( $p = 0.8$  in the panel), JAT becomes unstable, leaving only PDL state as the equilibrium of coevolutionary dynamics. Closed and open circle respectively represents stable and unstable equilibrium, and solid curves are null isoclines of coevolutionary dynamics. The grey lines with arrows indicate the trajectories of coevolutionary dynamics. The other parameters:  $\sigma_i = 1$  in all panels. The production function is quadratic  $f(\xi) = \xi^2$ . Schematic diagrams of the bottom of figure 2 represent the pattern of specialization of PDL, HPDL, and JAT states. Circles and squares represent species and metabolites, respectively. An arrow indicates which species produce

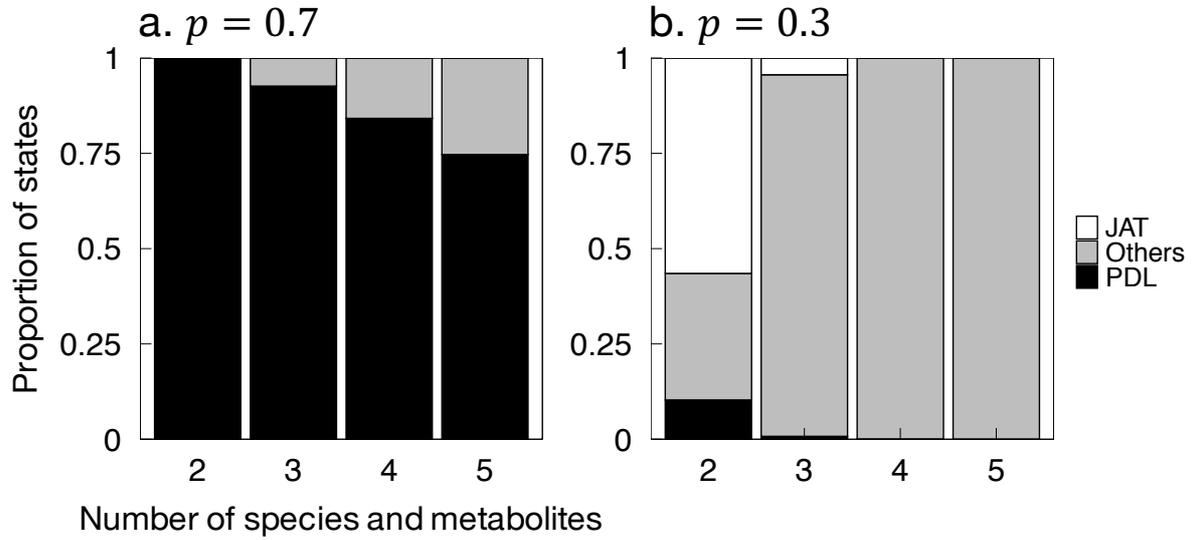
which metabolites (strictly speaking, an arrow is drawn if the corresponding investment  $x_{ij}$  from species  $i$  to metabolite  $j$  is positive at the equilibrium).

Figure 3



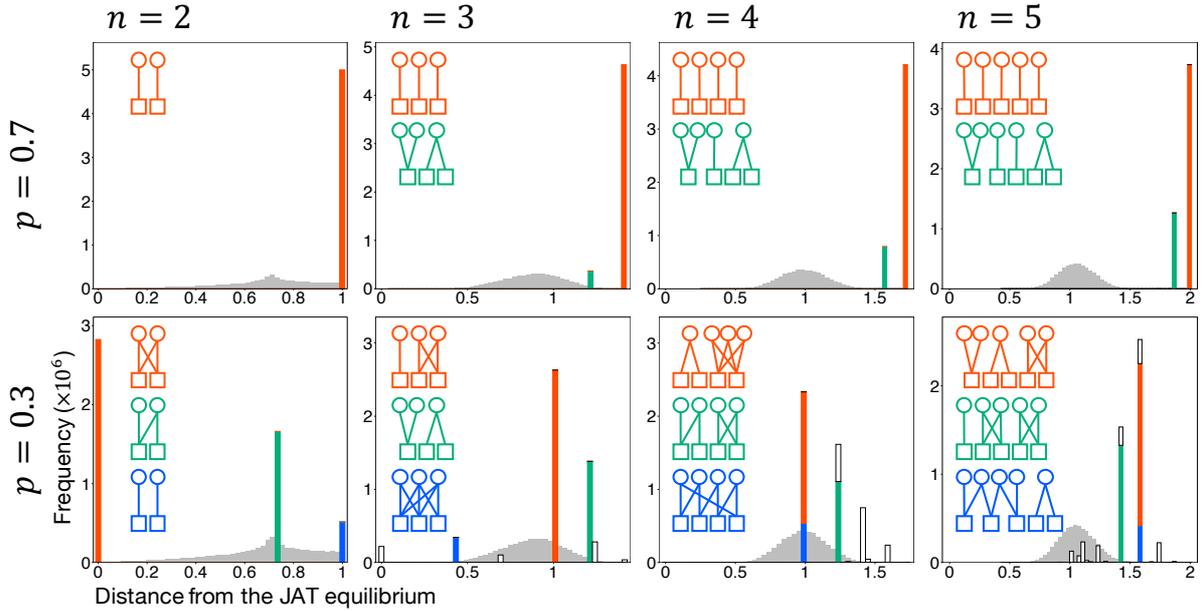
**Figure 3.** The boundaries for local stability of *perfect-division-of-labour* (PDL; panel a) and *Jack-of-all-trades* (JAT; panel b) in  $n$ -species  $n$ -metabolites model ( $n = 2,3,4,5$ ). (Panel a) The regions for the stability of *perfect-division-of-labour* (PDL) are plotted against the probability of encountering partners  $p$  (horizontal axis) and the coefficient  $\varepsilon$  for the linear component in production function  $f$  (vertical axis). A dashed line represents the boundary of the regions and becomes finer along  $n = 2,3,4$ . The solid line is the boundary with  $n = 5$ . (Panel b) The same as a) but showing the boundaries for the local stability of *Jack-of-all-trades* (JAT). Note that  $p_1 = p_2 = p$  and  $\sigma_1 = \sigma_2$  are assumed. The production function is  $f(\xi) = \varepsilon\xi + \xi^2$ .

Figure 4



**Figure 4.** Fraction, for randomly varied initial conditions, eventually fell into *perfect-division-of-labour* (PDL; black), *Jack-of-all-trades* (JAT; white), and the others (grey) in the coevolutionary dynamics with  $n$ -species and  $n$ -metabolites ( $n = 2,3,4,5$ ). (Panel a) when obtaining symbiotic partners is easy ( $p = 0.7$ ), the frequency that PDL is attained through the coevolutionary process stays high even when the numbers of species and essential metabolites are increased up to 5. (Panel b) When obtaining symbiotic partners is difficult ( $p = 0.3$ ), the fraction of PDL becomes progressively lower as the number of species increases. Note that  $p_1 = p_2 = p$  and  $\sigma_1 = \sigma_2$  are assumed. The production function is  $f(\xi) = \xi^2$ .

Figure 5



**Figure 5.** Distributions of Euclidian distances between an attained equilibrium and JAT equilibrium in  $n$ -species  $n$ -metabolites model ( $n = 2,3,4,5$ ). The distances quantify the degree of specialization in essential metabolite production in symbiotic species (see section Simulation results in the main text). Upper and lower panels show the cases when obtaining symbiotic partners is easy ( $p = 0.3$ ) and difficult ( $p = 0.7$ ), respectively. In each panel, the histograms indicate the distribution of Euclidean distance from JAT equilibrium to initial points which are sampled randomly (grey) and attained equilibria (red, green, blue, and white) of the coevolutionary dynamics. The colours of attained states correspond to their abundance: red indicates the most frequent state, green does the second most frequent, blue does the third most frequent, and white does all the other states. Small schematic diagrams in the inset of panels show the pattern of specialization of the equilibria (the same as in figure 2). The colours of the diagrams correspond to that of histograms. The number of sampled initial points are  $5 \times 10^6$ , the production function is  $f(\xi) = \xi^2$  (that is,  $\varepsilon = 0$  in  $f(\xi) = \varepsilon\xi + \xi^2$ ), any probabilities of encountering partners and any rates of evolution are the same as the others (i.e.  $p_i = p$  and  $\sigma_i = \sigma$  for all  $i$ ).