On the structure of an evolutionary transition: dependence and cohesion in incipient endosymbioses

Gaurav S. Athreya[1,2], Peter Czuppon[3], Chaitanya S. Gokhale[2,4*]

[1]Indian Institute of Science Education and Research, Pune, India

[2]Research Group for Theoretical Models of Eco-evolutionary Dynamics, Department of Theoretical Biology, Max Planck Institute for Evolutionary Biology, August-Thienemann-Straße 2, 24306 Plön, Germany

[3]Institute for Evolution and Biodiversity, University of Münster, Hüfferstraße 1, 48149 Münster, Germany

[4]Center for Computational and Theoretical Biology, Julius-Maximilians University Würzburg, Klara-Oppenheimer-Weg 32, Campus Hubland Nord, 97074 Würzburg, Germany

* Corresponding author. Email: chaitanya.gokhale@uni-wuerzburg.de

Abstract

How did biological organisms become so complex? Endosymbiosis is the prototypical example of an egalitarian evolutionary transition and is central to the origins of many complex biological systems. Why do only some symbioses undergo evolutionary transitions, and how does the host-symbiont relationship change during this process? Here, we characterise endosymbiosis by two emergent collective-level properties: the relationship between the host and symbiont (mutual dependence) and reproduction (reproductive cohesion). Using methods from adaptive dynamics, we study the co-evolution of the traits underlying these properties. Our central result is a robust demonstration that even when investments in dependence and cohesion are uncorrelated, host-symbiont mutual dependence arises faster than reproductive cohesion. Further, we show that the collective's formation and shared fate, coupled with different generation times of the host and symbiont, leads to an emergent asymmetry in how much they invest in the collective. Lastly, we account for biological realism that previous models have ignored and show that this can preclude a successful evolutionary transition towards stable endosymbiosis. Together, this work uncovers a fundamental property of endosymbioses, highlighting the immense effects of simple ecological factors and providing a clear way forward for theoretical and empirical investigations.

Keywords: major transitions — symbiosis — mitochondria — microbiome — population dynamics — adaptive dynamics — invasion fitness — Red King effect — obligacy — adhesion

1 Introduction

Endosymbiosis is a phenomenon of central importance in evolutionary biology, leading to the origin of eukaryotes and several astonishing long-term associations between different species. It is the prototypical example of an egalitarian – i.e. different species coming together – transition in individuality: initially autonomous and unrelated entities – the host and symbiont – come together to give rise to a more complex, integrated entity (Szathmáry and Smith, 1995; Queller, 2000; Szathmáry, 2015; West et al., 2015). It has been argued that endosymbiosis, and the central role it plays in eukaryogenesis, is fundamentally responsible for overcoming bioenergetic constraints that impeded the success of prokaryotes (Lane, 2015).

Endosymbionts are present in most forms of life, even in unicellular prokaryotes (Wujek, 1979; Corsaro et al., 1999). Mitochondria and plastids have been famously shown to be endosymbionts arising from an ancient union between Archaeabacteria and an alphaproteobacterium (Sagan, 1967; Koonin and Yutin, 2014; Martijn et al., 2018; Fan et al., 2020). Many insects, such as the sap-sucking aphids, have been shown to be co-diversifying with their *Buchnera*, *Wigglesworthia*, and *Wolbachia* endosymbionts for millions of years (Shigenobu et al., 2000; Zientz et al., 2004; Wu et al., 2004; Hansen and Moran, 2011). There are many more examples - methanogenic endosymbionts in anaerobic ciliates (Embley and Finlay, 1994), nitrogen-fixing endosymbionts in the diatom *Rhopalodia* (Prechtl et al., 2004), consortia of chemosynthetic bacteria in gutless tubeworms (Woyke et al., 2006), cyanobacterial endosymbionts in sponges (Thacker, 2005). New, interesting kinds of endosymbiotic associations are being discovered even now, such as denitrifying endosymbionts in anaerobic ciliates (Graf et al., 2021). Despite this widespread presence, endosymbiosis and egalitarian evolutionary transitions have been given relatively lesser attention than their fraternal counterparts (where lower-level individuals are of the same species), although this is slowly changing (Law and Dieckmann, 1998; Estrela et al., 2016; Zachar and Szathmáry, 2017; Zachar et al., 2018; Nguyen and Baalen, 2020; Patel and West, 2022).

In this work, we are interested in the relationship between the host and symbiont and how it evolves over the course of an evolutionary transition. Some symbioses have undergone an evolutionary transition, whereas others have not. For example, proto-mitochondria and their ancestral hosts underwent a transition to form the modern eukaryotic cell (Sagan, 1967; Maynard Smith and Szathmáry, 1995). *Buchnera* endosymbionts are obligately dependent on their aphid hosts, vertically transmitted, and have small genomes (Bennett and Moran, 2015). On the other hand, the tubeworm *Riftia* is obligately dependent on its symbiont *Endoriftia*, but the latter has a free-living stage, and is transmitted horizontally (Nussbaumer et al., 2006). The bobtail squid *E. scolopes* and its bioluminescent *Vibrio* symbiont are both facultative, and the symbiont is acquired horizontally every generation (McFall-Ngai, 2014). What controls this difference in outcomes, and how is it impacted by the life-history traits of the host and symbiont? How do the properties of the collective co-evolve, and is there a difference between host and symbiont evolutionary trajectories? Such questions began to be addressed theoretically in the work of Law and Dieckmann (1998), followed by Nguyen and Baalen (2020). We construct similar, but more general models to answer the above questions.

We must first precisely define endosymbiosis. We define a host-symbiont collective to be obligate endosymbiosis if it exhibits three properties: (i) intracellular union; (ii) at least one

of the host or symbiont is obligately dependent on this interaction; and (iii) the collective can reproduce as a unit. Following Keeling and McCutcheon (2017), we use "symbiosis" to mean any sustained organismal interaction on the pathogenic-beneficial continuum. Synchronised collective reproduction endows the collective with a life cycle, previously proposed as the defining characteristic of an entity that can undergo an evolutionary transition (van Gestel and Tarnita, 2017). For example, this definition excludes a gut microbiome as there is no intracellular union or synchronised reproduction.

To study the conditions for the evolution of endosymbiosis, we use methods from adaptive dynamics (Metz et al., 1992; Dieckmann and Law, 1996; Geritz et al., 1998). The population dynamics and evolution is modelled using a separation of ecological and evolutionary timescales, which allows the application of "evolutionary invasion analysis" (Otto and Day, 2007).

2 An eco-evolutionary model

We consider three types of individuals - independent hosts (H), independent symbionts (S), and host-symbiont collectives (C). The main process of interest is the evolution of the growth rates of these types - we characterise obligate endosymbiosis by a positive growth rate for the collective and a zero growth rate for the independent types.

Inspired by the review of Estrela et al. (2016), we characterise a symbiosis by two emergent, collective-level properties: the degree of host-symbiont mutual dependence and the degree of their reproductive cohesion (see Figure 1). The evolution of symbiosis can thus be visualised as taking place in the plane of these two quantities. We say that an evolutionary transition in individuality occurs when there is full mutual dependence and full reproductive cohesion - the constituent individuals cannot live or reproduce without the other. The objective of this work is to understand evolutionary trajectories in this dependence-cohesion plane.

We model the evolutionary dynamics along these two phenomenological trait axes each for the host and symbiont (see Figure 1). First, mutual dependence: let f_H, f_S, f_C be the population growth rates of the host, symbiont, and collective, respectively. We introduce Ω_H, Ω_S, referred to hereafter as the "obligacy" of the host and symbiont, respectively. These are dimensionless numbers in $[0, 1]$ and denote the degree of dependence of the host and symbiont on the formation of the collectives and the benefits they gain. These traits formalise a tradeoff between individual and collective reproduction. Since an organism has only a finite amount of resources, any investment in collective reproduction comes at a cost to its reproduction (and vice versa). Further, to make explicit the co-localised nature of an endosymbiotic interaction, we assume that the benefits of endosymbiosis are only present when the organisms are part of the collective, implying that the growth rates of the host in isolation do not depend on the symbiont's investment. The mathematical translation of these statements is in terms of the partial derivatives of f_i along the obligacies and is stated precisely in Appendix A.1. Note that Ω_H and Ω_S can in general be different since the two participants may have asymmetric dependencies on each other.

Second, reproductive cohesion: let a be the rate of association of the host and symbiont to form the collective, and d that of dissociation of collectives to form independent host and

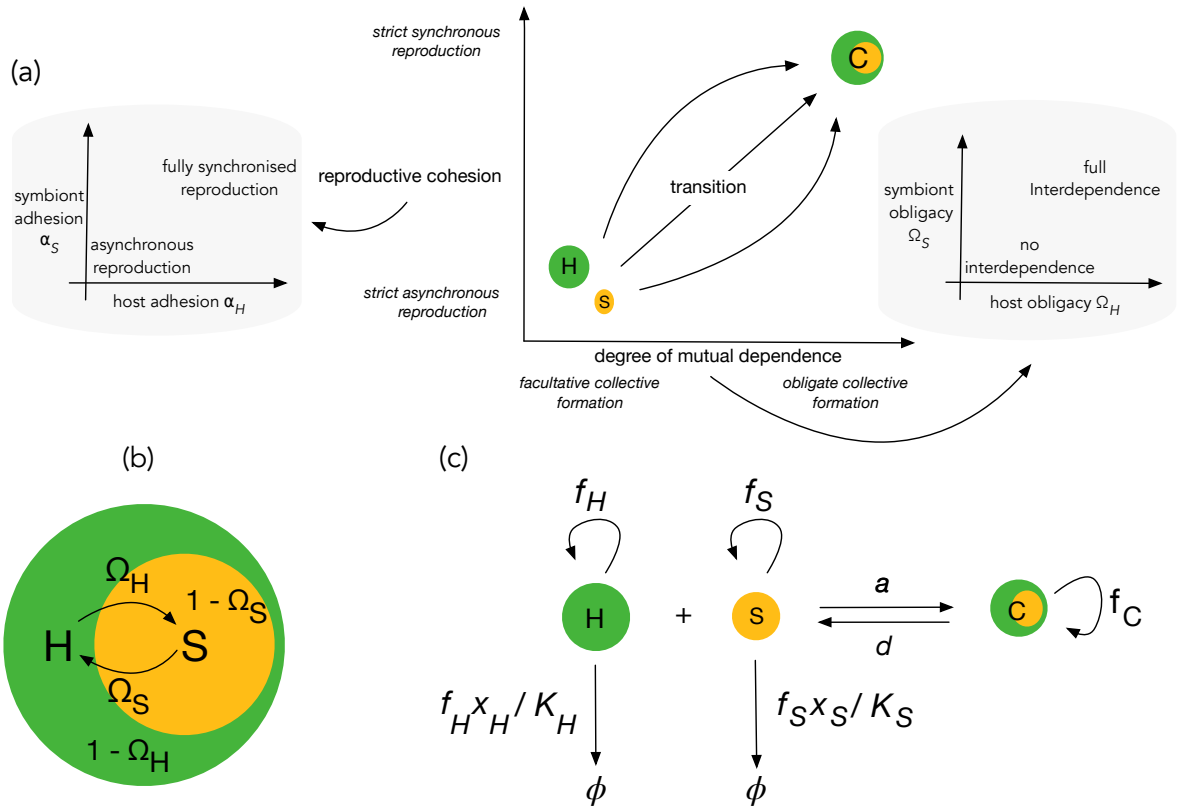

Figure 1: **(a)** We conceptualise the trajectory of an evolutionary transition in the plane of two quantities - the reproductive cohesion of the lower-level individuals, and the degree of mutual dependence between them. The adhesions α_H and α_S of the host and symbiont respectively control their investment in reproductive cohesion i.e., the degree of synchronised reproduction. The obligacy Ω_i describes the dependence of type i on forming a collective. In this study, we study the co-evolutionary trajectory of the adhesions with the obligacy, and also between the host and symbiont. **(b)** On a microscopic, physiological scale, it is useful to picture that there is – exclusively when they are both part of the collective – resource exchange between the host and symbiont. In this setting, the traits Ω_i control how much resource sharing the host/symbiont individuals are prone to. The trait α_i is more phenomenological, but is most concretely connected to the ratio of synchronised to asynchronised reproduction of species i. **(c)** The flows of the population dynamical model. The host and symbiont have a logistic growth rate, corresponding, for example, to an intrinsic birth rate f_i, and a density-dependent death rate $f_i x_i / C_i$. The collective has a density-independent (i.e. exponential) growth rate f_C. Further, the host and symbiont associate with each other at rate a, and dissociate with each other at rate d.

symbiont. We introduce α_H, α_S, henceforth referred to as the "adhesion" of the host and symbiont respectively. These traits are also dimensionless in $[0, 1]$, and a higher adhesion denotes a higher association rate, higher propensity of synchronised birth, and conversely a lower propensity of asynchronised birth. These traits hence induce a tradeoff between processes favouring the formation of the collective versus those favouring the individuals (see Appendix A.1 for mathematical details).

To differentiate between the Ω and α trait pairs, it is useful to focus on an example such as a fig-wasp mutualism (Herre et al., 2008). This association is an enmeshing of two life cycles – the fig tree depends on the wasp, the pollinator, and the wasp is dependent on the fig tree (the fruit, to be specific) to complete a part of its development. The two species here are therefore high in their dependence on each other, but they do not physically reproduce as a unit. Hence in this case the obligacies Ω_i are high but the adhesions α_i are not.

2.1 Ecological and evolutionary dynamics

We formalise the population dynamics taking place over short, mutation-free timescales as follows. Each type gives rise to offspring with an intrinsic growth rate $f_i, i \in \{H, S, C\}$, and the host and symbiont associate and dissociate from each other at rates a and d respectively, flowing into and out of the collective population (Figure 1(c)). The dynamics are given by:

$$\dot{x}_H = f_H x_H \left(1 - \frac{x_H}{K_H}\right) - a x_H x_S + d x_C, \tag{1a}$$

$$\dot{x}_S = f_S x_S \left(1 - \frac{x_S}{K_S}\right) - a x_H x_S + d x_C, \tag{1b}$$

$$\dot{x}_C = f_C x_C + a x_H x_S - d x_C. \tag{1c}$$

Some notes are in order. First, we consider that the intrinsic growth of hosts (H) and symbionts (S) is logistic, whereas that of the collective (C) is exponential. This has a biological motivation: symbioses allow the exploration of novel niches (Kleiner et al., 2012; Moran et al., 2003), and further, we are interested in the invasion of host-symbiont collectives. So we assume that during this phase, host-symbiont collectives are so rare that they do not limit their own growth rate.

Second, the parameter d is not as biologically explicit as can be. Previous works (Law and Dieckmann, 1998; Nguyen and Baalen, 2020) have considered explicit rates at which hosts or symbionts can - when part of the collective - give birth to independent offspring, or die. Both these mechanisms decrease the number of collectives and give rise to a free-living individual. Since we are only interested in the tradeoff between synchronised and asynchronised reproduction, we adopt the simpler picture given by a and d. One can rewrite the models of Law and Dieckmann (1998) and Nguyen and Baalen (2020) in terms of ours under a specific choice of parameters, see Appendix (A.3).

Third, the formation of the collective in our model can be formally shown to be comparable to a mutualism between the host and symbiont, see Appendix A.2.

The parameters in Figure 1(c) are functions of the obligacy and adhesion, and we are interested in their long-term co-evolution. We use adaptive dynamics, where the main ob-

ject of interest is the invasion fitness of a mutant in an environment generated by a resident population. To this end, it is assumed that the ecological and evolutionary timescales can be separated i.e., the (ecological) realisation of a mutant's fate takes place much faster compared to the (evolutionary) timescale on which the next mutant arises (Metz et al., 1992) .

To compute the invasion fitness of an arbitrary mutant, suppose that there is a resident population at ecological equilibrium. The invasion fitness is the growth rate of a small number of mutants in a resident population that is at ecological equilibrium (Metz et al., 1992). The mutant can invade if the growth rate is positive; if it is negative, the mutant goes to extinction (Dieckmann and Law, 1996; Dercole and Rinaldi, 2008). This quantity is not amenable to mathematical analysis in our case, so we use the next-generation theorem, which gives, under some conditions, an alternate characterisation of the same criterion (van den Driessche and Watmough, 2002; Hurford et al., 2010). We then make use of the canonical equation (Dieckmann and Law, 1996) to study the macroscopic behaviour of long-term evolutionary trajectories.

3 Results and Discussion

3.1 An invasion criterion

The system of equations (1) has four fixed points for the population densities, three of which are trivial (Appendix A.4). The fourth fixed point is internal, and feasible i.e., all equilibrium densities are non-negative, precisely when

$$d > f_C \left(1 + a\sqrt{\frac{K_H K_S}{f_H f_S}}.\right) \tag{2}$$

We shall henceforth refer to this lower bound on d as the "feasibility bound". It can be shown via the Routh-Hurwitz criterion (Edelstein-Keshet, 2005) that the feasibility bound is necessary also for linear (local) stability of the fixed point. We determined computationally that this inequality guarantees linear stability for a large range of parameters (see Appendix A.4).

After sufficiently long time, a mutant (without loss of generality, a host mutant) arises having obligacy $\widetilde{\Omega}_H$ and adhesion $\widetilde{\alpha}_H$. We show (Appendices A.1, A.5) that the mutant can invade if and only if

$$\frac{\widetilde{f}_H}{\widetilde{a}}\left(1 - \frac{\widetilde{d}}{\widetilde{f}_C}\right) > \frac{f_H}{a}\left(1 - \frac{d}{f_C}\right) \tag{3}$$

Since the host and symbiont are so far identical in everything but the label we impose on them, an analogous criterion exists for the fate of a symbiont mutant. This criterion is "separable" into two quantities of the same functional form, each depending only on either the mutant or resident. This implies that over the course of successive mutations, this quantity is maximised.

First, it implies that evolutionary branching is impossible (see Appendix A.7). Second, this criterion clearly demonstrates the conflict between the levels of selection in this problem.

To illustrate this, let us restrict ourselves to the simple case where adhesion does not evolve, and we assume that the parameters a, d do not depend on Ω_H, Ω_S. The invasion criterion (3) reduces to

$$\widetilde{f}_H \left(1 - \frac{d}{\widetilde{f}_C}\right) > f_H \left(1 - \frac{d}{f_C}\right) \tag{4}$$

For a mutant with trait $\tilde{\Omega}_H$, it invades by either increasing f_H or f_C. The conflict becomes clear when one realises that due to the constraints we set up earlier that arise from the meaning of the traits Ω_i and α_i, f_H and f_C change at the cost of one another - they cannot both increase. Given that the new mutant can only have a higher or lower obligacy, does this criterion select for $\tilde{\Omega}_H > \Omega_H$ (increased independent host growth rate f_H) or $\tilde{\Omega}_H < \Omega_H$ (increased collective growth rate f_C)? We answer this question exactly by means of the canonical equation of adaptive dynamics (Dieckmann and Law, 1996; Champagnat et al., 2006). By the same constraints as above, and the fact that $d > f_C$, one can show that the obligacy Ω_i increases monotonically (Appendix A.6). In fact, one can show analytically that all four traits $\Omega_H, \Omega_S, \alpha_H, \alpha_S$ must increase over evolutionary time.

Stepping away from the mathematical machinery, there is a clear explanation - the collective has a uniform growth benefit because of its exponential growth, and so investment in its growth is not unexpected. The apparent cost to independent growth is a hoax – the total number of hosts is the sum of those independent and those in collective. Thus more collective growth implies more hosts. While this result is therefore not surprising in itself, the analytical setup is conducive for asking more nuanced questions.

3.2 Evolution towards full dependence is inhibited by a feasibility bound

The generic analysis of the invasion criterion (4) has so far yielded many insights, but it is useful to computationally study a concrete example. Again, we will restrict evolution to only the obligacies, of which there are two. Let us fix the mapping

$$f_H(\Omega_H, \Omega_S) = r_H(1 - \Omega_H)\,, \tag{5a}$$
$$f_S(\Omega_H, \Omega_S) = r_S(1 - \Omega_S)\,, \tag{5b}$$
$$f_C(\Omega_H, \Omega_S) = r_C \Omega_H \Omega_S\,, \tag{5c}$$

obeying the required constraints, with constants r_H, r_S, r_{HS} which set the growth rate scale. For this mapping, the invasion criterion (4) becomes trivial: a mutant with obligacy $\tilde{\Omega}_i$ invades in a population that currently has trait value Ω_i when $\tilde{\Omega}_i > \Omega_i$.

In the adaptive dynamics framework, the population dynamical equilibrium must be feasible and stable for us to speak of successive mutants. The conditions for these to hold depend on the f_i, which in turn depend on the Ω_i that changes over evolutionary timescales. Determining the existence (or not) of a resident population across the evolutionary trajectory is therefore nontrivial. We track the evolution of Ω_i in a particular system with realistic parameters, while numerically checking if feasibility and stability hold over the course of evolution. The results of this exercise are shown in Figure 2.

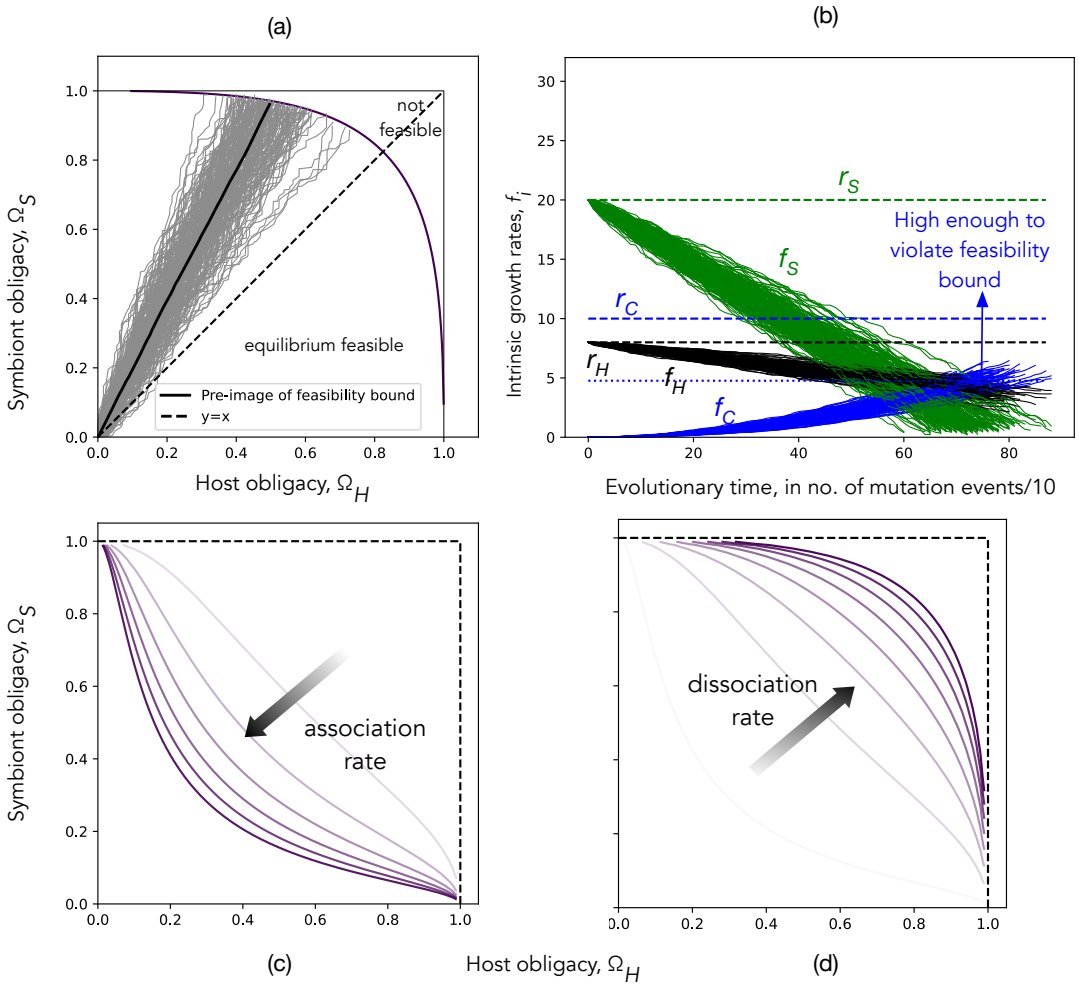

Figure 2: **Evolutionary trajectories are biased towards higher symbiont obligacy. (a)** For the mapping in (5), we simulated 200 independent evolutionary trajectories. The trajectories monotonically increase with Ω_S increasing faster, and then stop when the fixed point is no longer stable (it also becomes infeasible). In the unstable region, the collective population density increases to infinity and, due to dissociation, maintains a comparatively small population of independent hosts and symbionts. Stability is lost precisely at the feasibility bound (2), which is a lower bound on the here constant parameter d. See Appendix B for information on how the average trajectory is computed in this and following figures. **(b)** Using the mapping (5), we can also visualise the dynamics of the growth rates. We observe that the independent growth rates decrease, and the collective growth rate increases until it becomes high enough that the feasibility bound is violated. **(c,d)** The feasibility bound controls the maximal investment of the host and symbiont in collective reproduction before stability is lost. Here we visualise the shape of the boundary for different values of a (in panel **c**) and different values of d (in panel **d**). The bound moves closer to (0,0) for high a and low d. Since we expect high association rate and low dissociation rate in a highly integrated endosymbiosis, this predicts that the dependence cannot go above a very small number. This motivates the relaxation of the assumption of exponential growth of the collective. **Parameter values**. Common for all: $r_H = 8, r_S = 20, r_C = 10, K_H = 100, K_S = 200$; for (a) and (b): $a = 0.1, d = 50$; for (c): $d = 50$, $a \in [0.01, 5.0]$; for (d): $a = 2.1$, $d \in [25, 750]$.

Starting from the natural initial condition of $(\Omega_H, \Omega_S) = (0, 0)$ i.e., no dependence, we observe first that both obligacies increase, as expected. However, the trajectory does not reach $(\Omega_H, \Omega_S) = (1, 1)$. The obligacies increase upto a certain point, after which the ecological equilibrium becomes unstable (and also infeasible). The feasibility bound (2) is violated, leading to the collective growth rate becoming too high to sustain independent host/symbiont populations. The resident population ceases to exist. The collective's abundance then grows to infinity, also maintaining a relatively small population of independent hosts and symbionts via association and dissociation.

Second, the symbiont obligacy Ω_S is almost always larger than the host obligacy Ω_H. This stems from the parameter choices we have made: r_S is chosen to be larger than r_H to reflect the symbiont's smaller generation time, and the carrying capacity of the symbiont is chosen to be twice that of the host. These parameters induce a bias in investment since they determine the equilibrium population size of the host and symbiont, which determines how quickly mutants arise, assuming other factors like mutation rate are constant. This result has a clear explanation. It is always beneficial for both the host and symbiont to invest more in collective reproduction, and the symbiont invests more since it gets more opportunities (mutations) to do so.

3.3 Logistic growth of the collective induces new phenomena

In the most general model, the growth of the collective cannot be exponential as we previously assumed; it must be self-limiting. Here, we recognise this by introducing a logistic growth function in the equation describing the change in the collective's population density. The logistic model of population dynamics will refer to the following system of ordinary differential equations:

$$\dot{x}_H = f_H x_H \left(1 - \frac{x_H}{K_H}\right) - a x_H x_S + d x_C \,, \tag{6a}$$

$$\dot{x}_S = f_S x_S \left(1 - \frac{x_S}{K_S}\right) - a x_H x_S + d x_C \,, \tag{6b}$$

$$\dot{x}_C = f_C x_C \left(1 - \frac{x_C}{K_C}\right) + a x_H x_S - d x_C \,, \tag{6c}$$

with the additional parameter K_C, set in accordance with intuition to be larger than the host and symbiont carrying capacities. We keep all common parameters and initial conditions identical to the previous case to make comparisons precise. A full characterisation of the logistic model is not attempted; our goal is merely to make 'existence' statements regarding novel phenomena that become possible within the logistic model.

The first qualitatively novel phenomenon that arises is the evolution of full dependence. The evolutionary trajectory $(\Omega_H(t), \Omega_S(t))$ reaches the maximum value (1,1), while it did not do so in the exponential case due to the feasibility boundary. The inclusion of logistic growth thus destroys the feasibility boundary, since the resident equilibrium remains a demographic attractor in the absence of mutants until the evolution of full dependence.

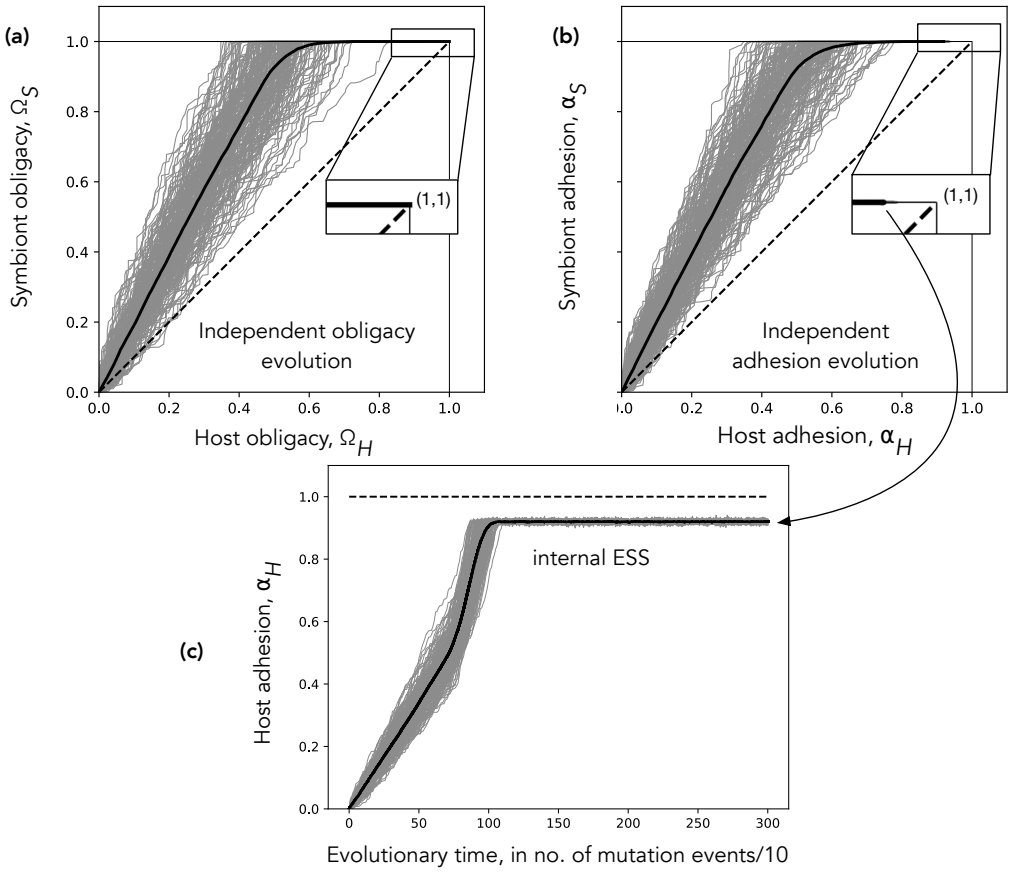

Figure 3: **Qualitatively new phenomena are induced by logistic growth of the collective.** The relaxation of the exponential growth assumption for the collective gives rise to the following phenomena. All panels contain results of 200 independent stochastic simulations. **(a)** The obligacies, which were inhibited by the feasibility boundary, now evolve to (1,1) - full mutual dependence between the host and symbiont. **(b)** Adhesion evolution was, in the exponential model, also inhibited by a boundary. Under the logistic model, we observe that the adhesions evolve to high values, but the traits (α_H, α_S) settle to a value different from (1,1). After sufficiently long, $\alpha_S = 1$ whereas $\alpha_H < 1$ i.e. the host invests lesser than the symbiont. **(c)** This panel shows that the long-run value of α_H is indeed an evolutionary stable strategy - one observes that the trajectories approach this equilibrium value (so it is convergence stable) and the trajectories are not repelled once they get there (it is uninvadable).
Parameter values. Common for all: $K_H = 100, K_S = 200, K_C = 500, a = 0.1$; for **(a)**: $r_H = 8, r_S = 20, r_C = 10, d = 50.0$; for **(b)** and **(c)**: $f_H = 8, f_S = 20, r_C = 10, d = 50$.

The second departure from the exponential model is exhibited by the independent evolution of the adhesions α_i, which also evolve to higher values than in the exponential case. Instead of stopping at a boundary, one observes that given sufficient time, the final state is an ESS where $\alpha_S = 1$, and α_H settles at some high value smaller than one. This result is provably impossible in the exponential model. Notice here that it is $\alpha_S = 1$ whereas $\alpha_H < 1$ i.e., the slower-reproducing species invests lesser in synchronised reproduction. This situation may be likened to the evolution of vertical transmission.

The proximate reason for why the boundary disappears in the logistic model is as follows. The crux of the matter is the intrinsic growth term of the collective's abundance - $f_C x_C$ in the exponential model, and $f_C x_C(1 - x_C/K_C)$ in the logistic model. When this term becomes "too high", the equilibrium disappears. The difference then is that $f_C x_C$ increases much faster than $f_C x_C(1 - x_C/K_C)$. Even at intermediate values of the Ω_i, the intrinsic growth of the collective can become too high. In the logistic model however, the growth of the collective is self-limiting and therefore increases slower. Further, the population cannot increase indefinitely because of the carrying capacity that is now present. Both of these factors, which are biologically more "correct" than the exponential model, induce these new, richer, more realistic phenomena.

3.4 Does mutual dependence evolve before or after reproductive cohesion?

Now we are equipped to finally understand the co-evolution of the traits (Ω_H, α_H) and (Ω_S, α_S). We have thus far studied the evolution along each of these axes separately, and in two related models - exponential and logistic. Studying these simplified cases enables us to build intuition and better interpret their co-evolution.

Representative results for the exponential model are shown in Figure (4) for a specific mapping of traits to parameters. This mapping is similar to Equations (5), and is given by

$$f_H(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_H(1 - \Omega_H) \tag{7a}$$
$$f_S(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_S(1 - \Omega_S) \tag{7b}$$
$$f_C(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_C \Omega_H \Omega_S \alpha_H \alpha_S \tag{7c}$$
$$d(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = d_0(1 - \alpha_H \alpha_S) \tag{7d}$$
$$a(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = a \in \mathbb{R}_{\geq 0} \tag{7e}$$

We represent the degree of dependence between the host and symbiont by the product $\Omega_H \Omega_S$ of their individual obligacies, and the degree of reproductive cohesion by $\alpha_H \alpha_S$. [1] These results show two important facts. First is that for both the host and symbiont, it is adaptive to evolve such that $\Omega_i > \alpha_i$ i.e. being more obligate than cohesive is a result, in this model, of natural selection. Second, if one studies the quantities $\Omega_H \Omega_S$ and $\alpha_H \alpha_S$, one concludes

[1] The multiplicative form here is connected to how they map to ecological parameters above - a multiplicative effect of the Ω_i on collective reproduction f_C motivates the interpretation of $\Omega_H \Omega_S$ as the degree of mutual dependence as opposed to other potential measures such as $\Omega_H + \Omega_S$.

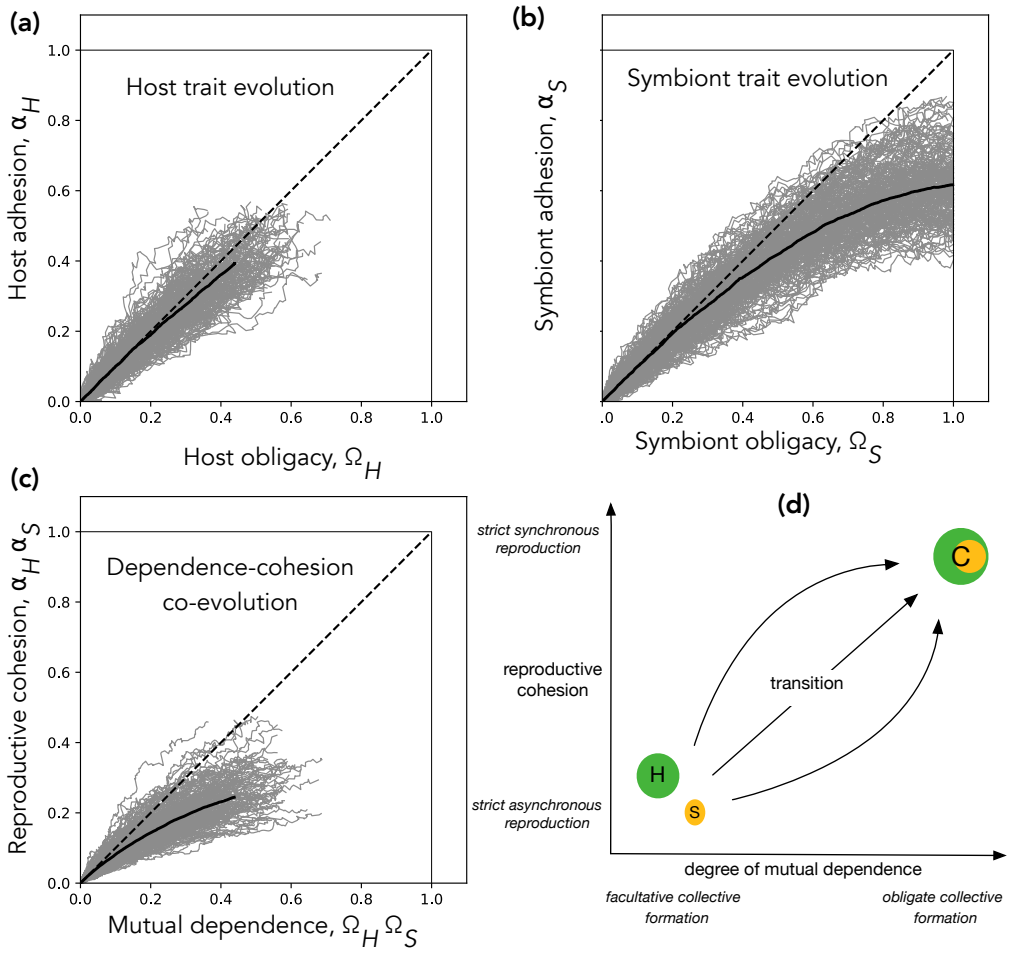

Figure 4: **Mutual dependence evolves before collective reproduction.** Here we consider the co-evolution of the 4 traits Ω_H, α_H and Ω_S, α_S under the exponential model. Recall that this implies the existence of the feasibility boundary that we saw in Equation (2), then in Figure 2. **(a, b)** Both host and symbiont traits evolve such that, on average, $\Omega_i > \alpha_i$. Notice that symbiont evolution (panel **(b)**) shows that symbiont traits evolve to much higher values than host traits. This is consistent with the Red King-type argument – both these traits align reproductive traits of the host and symbiont, and since the symbiont has a higher evolutionary rate, it invests more. **(c)** Given the individual evolutionary trajectories of the host and symbiont, one can collate information to obtain measures that describe the collective as a whole - the degree of mutual dependence is represented here by the product $\Omega_H \Omega_S$ and the degree of reproductive cohesion by $\alpha_H \alpha_S$. We observe that evolution of mutual dependence is faster than that of reproductive cohesion. The trajectories stop when they hit the feasibility bound; it is nontrivial to visualise its projection in this plane and is hence absent. **(d)** We reproduce this panel from Figure 1 to reiterate the objective of this work and illustrate the progress we have made in answering the question posed in the Introduction. **Parameter values**. $K_H = 100, K_S = 200, K_C = 500, a = 0.1, r_H = 8, r_S = 20, r_C = 10, d = 50.0$

that evolutionary trajectories are biased in the direction of more mutual dependence than reproductive cohesion. This is a central result. It shows that, over time, one expects that a host-symbiont collective evolves such that the partners are closer to full mutual dependence than they are to reproducing synchronously.

Fundamentally, this is because the functional effect of the pairs of traits are different. Recall that both trait pairs increase because of the exponential growth of the collective. However, we assume that association and dissociation are processes that depend on both the α_H and α_S whereas independent growth of type i ($\in \{H, S\}$) is a function only of Ω_i (see Equation (7)). Therefore, a different Ω_i has more functional effect (in changing the invasion fitness) than a different α_i since the latter's effect always depends on the other adhesion as well. See Appendix A.9 for a more comprehensive explanation.

This result is robust to many of the specifics - choice of parameters, different generation times, a multiplicative effect of the individual traits on the collective, etc. (Supplementary Figure 10). The bias persists, and mutual dependence is always faster to evolve than reproductive cohesion. This suggests that endosymbioses in nature are more likely to be more mutually dependent than cohesive.

We performed the same analysis also for the logistic model, and observed results that are consistent with the exponential model, but with additional complications (Supplementary Figure 11). Note that the exponential model is an approximation of the logistic model, and so it describes well the dynamics roughly before the feasibility boundary.

4 Conclusions and conjectures

Endosymbiosis and the advances in complexity it made possible are astonishing. An endosymbiotic association gave rise to eukaryotes, and many other fantastical associations between unrelated species. In this work we endeavour to give a precise view of endosymbiosis as an egalitarian evolutionary transition in individuality, and study the effect of some important ecological factors on its origins.

We study two major characteristics of an endosymbiotic collective undergoing an evolutionary transition - the reproductive cohesion of the host and symbiont ("adhesion", α_H and α_S), and the level of mutual dependence between them ("obligacy", Ω_H and Ω_S). How they co-evolve is not clear; for example, which is faster to evolve? This top-down approach – identifying emergent properties and then studying their evolutionary origins – is the natural first step, but one must more generally construct questions using a bottom-up approach instead (van Gestel and Tarnita, 2017). The top-down approach is nevertheless useful to ask pointed questions about observable quantities.

We derive explicit invasion criteria which show that Ω_i and α_i are expected to increase with time, with that of the symbiont increasing faster. The traits, however, increase only upto a finite quantity less than the maximum due to the existence of a feasibility boundary. Relaxing the assumption of exponential growth of the collective shows that this boundary is an artefact; we observe evolution of full dependence and the existence of internal (< 1) evolutionarily stable strategies (ESSs) for the dynamics of adhesion. This additional biological realism produces the possibility that an evolutionary transition does not take place even if the collective has a

clear growth benefit, as evidenced by the internal ESS in Figure (3). We then show that the density of permitted evolutionary trajectories in the dependence-cohesion plane is not uniform: when Ω_i and α_i co-evolve, both species irrespective of their evolutionary rate evolve such that they are in general more obligate than adhesive.

This central result can be confronted with biological examples. We compiled a short, non-exhaustive list of well-studied endosymbioses (Table 1) where there is information on the level of vertical transmission (a proxy for reproductive cohesion) and the degree of mutual dependence. However, this data is crude, and due to its qualitative nature, inconclusive. Some cases are clear-cut: in the *Riftia-Endoriftia* endosymbiosis, there is high dependence but with horizontal transmission – supporting our results. However, the *Dictyostelium-Burkholderia* farming symbiosis is facultative from both sides, and has a mixed mode of transmission (DiSalvo et al., 2015). It is impossible to compare the degree of dependence and cohesion here due to the non-quantitative nature of this data. To our knowledge, very few studies have precisely estimated quantities related to the level of vertical and horizontal transmission (Darby and Douglas, 2003; Hosokawa et al., 2007). This survey and its limitations highlight the need to conduct an experiment where the evolutionary transition to stable endosymbiosis can be observed in real time.

More broadly, one hopes to understand why most symbioses do not undergo evolutionary transitions. A precise question in this context is to explain the existence of internal ESSs – putative end-points of evolution – where the participants of the symbiosis are not completely obligate or not completely adhesive. Our analysis identifies three distinct sources of an internal ESS. First, considering the density-dependent nature of the collective populations' growth can lead to internal ESSs. Second, Nguyen and Baalen (2020) discovered that there exist internal ESSs where the symbiont is not fully obligately dependent on the host. The only respect in which our model (1) differs from theirs is in the modelling of dissociation - they explicitly model the asynchronous growth processes described in Appendix A.3. Our exponential model (provably) cannot have such ESSs. Thus considering more biological realism produces internal ESSs. Third, The monotonic increase of the Ω_i and α_i is predicated on perfect tradeoffs – increased collective growth is always negatively correlated with independent growth, etc. Strong enough imperfections in these tradeoffs – perhaps independent and collective growth are sometimes correlated – can give rise to internal ESSs. This is not unreasonable given the diversity of life; perfect tradeoffs are the most interesting case since it is hardest here to undergo an evolutionary transition. Overall, more biological considerations make it increasingly unlikely that a symbiosis undergoes a major transition. It is a necessary and important extension of this work to understand, in a general sense, the causes of internal ESSs in evolutionary models of symbioses.

There are more caveats that come with the ecological considerations of this model. Perhaps most importantly, we consider that the collective must always consist of a fixed density of symbionts that are acquired all together and only once. This does not allow for successive symbiont acquisition events – in reality, each host has a dynamically changing symbiont population. Additionally, we restrict ourselves to the case where interaction becomes beneficial only in close proximity, but in doing so we do not consider benefits of interaction outside of the collective. We also do not consider that there might be competition between the independent types and the collective due to niche overlap, or that there might be competition for

resources between the mutant and resident collective. Relaxing each of these assumptions are worthy directions for future work.

In the evolutionary trajectories of obligacy and adhesion, the symbiont trait increased faster than the host's. This is an instantiation of the Red King effect: when host and symbiont interests are aligned and defection leads to lower payoff, such a bias must be expected (Bergstrom and Lachmann, 2003; Frean and Abraham, 2004). In the original Red King effect, the interest-aligning mechanism is a mutualism, an ecological interaction. Our analysis shows that the formation (and shared fate) of the collective is itself another, more basic, interest-aligning mechanism in evolutionary transitions. When coupled with different generation times for the host and symbiont, it leads to an emergent host-symbiont asymmetry. While the result – symbiont investment higher than host's – may be a posteriori compatible with exploitation of the symbiont by the host, such constructive models show that the symbiont invests more merely because, over evolutionary time, more generous symbiont mutants arise than host mutants - and they invade because more investment is beneficial. Further, emergent asymmetry between the host and symbiont is a recurring theme of symbioses. This asymmetry is usually in one direction – the symbiont seems to be prone to giving more than it receives (Foster et al., 2017; Bennett and Moran, 2015; Moran and Bennett, 2014; McCutcheon and Moran, 2011). In addition to Red King-type effects, gene loss due to Muller's ratchet (Moran, 1996) and environment-induced redundancy (McCutcheon and Moran, 2010; Husnik et al., 2013), and gene transfer from symbiont to host lead to reductive genome evolution of the symbiont. The asymmetry here is because symbionts occur inside their hosts and not the other way around. Simple factors such as generation times and the nested structure of endosymbioses are therefore clearly important, and further work is required to delineate their consequences from that of complicated strategies such as partner choice/sanction, or zero-determinant strategies (Press and Dyson, 2012). In simpler words, what are the natural evolutionary consequences of living inside a cell of another organism?

5 Acknowledgements

GSA thanks Dr. Sutirth Dey and Dr. Phuong Linh Nguyen for helpful discussions. GSA thanks the SHE scholarship of the DST-INSPIRE programme for financial support, and the Max Planck Society for supporting the research visit. CSG and GSA thank the Max Planck Society for generous funding.

6 Data and Code Availability

This study has no empirical data associated to it. All Mathematica notebooks, scripts, and figure generation pipelines can be found in the associated Github repository: https://github.com/tecoevo/incipient_endosymbiosis.

References

Bennett, G. M. and Moran, N. A. Heritable symbiosis: The advantages and perils of an evolutionary rabbit hole. *Proceedings of the National Academy of Sciences*, 112(33):10169–10176, 2015.

Bergstrom, C. T. and Lachmann, M. The red king effect: When the slowest runner wins the coevolutionary race. *Proceedings of the National Academy of Sciences*, 100(2):593–598, 2003.

Buchner, P. *Endosymbiosis of Animals with Plant Microorganisms*. Wiley, 1965. ISBN 9780470115176.

Champagnat, N., Ferrière, R., and Méléard, S. Unifying evolutionary dynamics: From individual stochastic processes to macroscopic models. *Theoretical Population Biology*, 69(3):297–321, 2006.

Corsaro, D., Venditti, D., Padula, M., and Valassina, M. Intracellular life. *Critical Reviews in Microbiology*, 25 (1):39–79, 1999.

Darby, A. C. and Douglas, A. E. Elucidation of the transmission patterns of an insect-borne bacterium. *Applied and Environmental Microbiology*, 69(8):4403–4407, 2003.

Denison, R. F. and Kiers, E. T. Life histories of symbiotic rhizobia and mycorrhizal fungi. *Current Biology*, 21 (18):R775–R785, 2011.

Dercole, F. and Rinaldi, S. *Analysis of Evolutionary Processes: The Adaptive Dynamics Approach and Its Applications*. Princeton University Press, 2008.

Dieckmann, U. and Law, R. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *Journal of Mathematical Biology*, 34:579–612, 1996.

DiSalvo, S., Haselkorn, T. S., Bashir, U., Jimenez, D., Brock, D. A., Queller, D. C., and Strassmann, J. E. Burkholderia bacteria infectiously induce the proto-farming symbiosis of dictyostelium amoebae and food bacteria. *Proceedings of the National Academy of Sciences*, 112(36), 2015.

Edelstein-Keshet, L. *Mathematical Models in Biology*. Society for Industrial and Applied Mathematics, 2005.

Embley, T. M. and Finlay, B. J. Systematic and morphological diversity of endosymbiotic methanogens in anaerobic ciliates. *Antonie van Leeuwenhoek*, 64(3-4):261–271, 1994.

Estrela, S., Kerr, B., and Morris, J. J. Transitions in individuality through symbiosis. *Current opinion in microbiology*, 31:191–198, 2016.

Fan, L., Wu, D., Goremykin, V., Xiao, J., Xu, Y., Garg, S., Zhang, C., Martin, W. F., and Zhu, R. Phylogenetic analyses with systematic taxon sampling show that mitochondria branch within alphaproteobacteria. *Nature Ecology & Evolution*, 4(9):1213–1219, 2020.

Felbeck, H. and Jarchow, J. Carbon release from purified chemoautotrophic bacterial symbionts of the hydrothermal vent tubeworm riftia pachyptila. *Physiological Zoology*, 71(3):294–302, 1998.

Foster, K. R., Schluter, J., Coyte, K. Z., and Rakoff-Nahoum, S. The evolution of the host microbiome as an ecosystem on a leash. *Nature*, 548(7665):43–51, 2017.

Frean, M. R. and Abraham, E. R. Adaptation and enslavement in endosymbiont-host associations. *Phys. Rev. E*, 69:051913, 2004.

Gehrer, L. and Vorburger, C. Parasitoids as vectors of facultative bacterial endosymbionts in aphids. *Biology Letters*, 8(4):613–615, 2012.

Geritz, S. A. H., Kisdi, E., Meszéna, G., and Metz, J. A. J. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology Research*, 12:35–57, 1998.

Graf, J. S., Schorn, S., Kitzinger, K., Ahmerkamp, S., Woehle, C., Huettel, B., Schubert, C. J., Kuypers, M. M. M., and Milucka, J. Anaerobic endosymbiont generates energy for ciliate host by denitrification. *Nature*, 591(7850):445–450, 2021.

Gruber-Vodicka, H. R., Dirks, U., Leisch, N., Baranyi, C., Stoecker, K., Bulgheresi, S., Heindl, N. R., Horn, M., Lott, C., Loy, A., Wagner, M., and Ott, J. Paracatenula, an ancient symbiosis between thiotrophic alphaproteobacteria and catenulid flatworms. *Proceedings of the National Academy of Sciences*, 108(29): 12078–12083, 2011.

Gündüz, E. A. and Douglas, A. Symbiotic bacteria enable insect to use a nutritionally inadequate diet. *Proceedings of the Royal Society B: Biological Sciences*, 276(1658):987–991, 2008.

Hansen, A. K. and Moran, N. A. Aphid genome expression reveals host-symbiont cooperation in the production of amino acids. *Proceedings of the National Academy of Sciences*, 108(7):2849–2854, 2011.

Herre, E. A., Jandér, K. C., and Machado, C. A. Evolutionary ecology of figs and their associates: Recent progress and outstanding puzzles. *Annual Review of Ecology, Evolution, and Systematics*, 39(1):439–458, 2008.

Hosokawa, T., Kikuchi, Y., and Fukatsu, T. How many symbionts are provided by mothers, acquired by offspring, and needed for successful vertical transmission in an obligate insect-bacterium mutualism? *Molecular Ecology*, 16(24):5316–5325, 2007.

Hurford, A., Cownden, D., and Day, T. Next-generation tools for evolutionary invasion analyses. *Journal of The Royal Society Interface*, 7(45):561–571, 2010.

Husnik, F., Nikoh, N., Koga, R., Ross, L., Duncan, R. P., Fujie, M., Tanaka, M., Satoh, N., Bachtrog, D., Wilson, A. C., von Dohlen, C. D., Fukatsu, T., and McCutcheon, J. P. Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell*, 153(7):1567–1578, 2013.

Keeling, P. J. and McCutcheon, J. P. Endosymbiosis: The feeling is not mutual. *Journal of Theoretical Biology*, 434:75–79, 2017.

Kiers, E. T., Rousseau, R. A., West, S. A., and Denison, R. F. Host sanctions and the legume–rhizobium mutualism. *Nature*, 425(6953):78–81, 2003.

Kleiner, M., Wentrup, C., Lott, C., Teeling, H., Wetzel, S., Young, J., Chang, Y.-J., Shah, M., VerBerkmoes, N. C., Zarzycki, J., Fuchs, G., Markert, S., Hempel, K., Voigt, B., Becher, D., Liebeke, M., Lalk, M., Albrecht, D., Hecker, M., Schweder, T., and Dubilier, N. Metaproteomics of a gutless marine worm and its symbiotic microbial community reveal unusual pathways for carbon and energy use. *Proceedings of the National Academy of Sciences*, 109(19), 2012.

Koonin, E. V. and Yutin, N. The dispersed archaeal eukaryome and the complex archaeal ancestor of eukaryotes. *Cold Spring Harbor Perspectives in Biology*, 6(4):a016188–a016188, 2014.

Lane, N. *The Vital Question: Why is Life the Way it Is?* Profile Books, 2015. ISBN 9781781250365.

Law, R. and Dieckmann, U. Symbiosis through exploitation and the merger of lineages in evolution. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1402):1245–1253, 1998.

Losey, G., Grutter, A., Rosenquist, G., Mahon, J., and Zamzow, J. Cleaning symbiosis: a review. In *Behaviour and Conservation of Littoral Fishes*. ISPA, 1999.

Martijn, J., Vosseberg, J., Guy, L., Offre, P., and Ettema, T. J. G. Deep mitochondrial origin outside the sampled alphaproteobacteria. *Nature*, 557(7703):101–105, 2018.

Maynard Smith, J. and Szathmáry, E. *The major transitions in evolution*. W. H. Freeman, Oxford, 1995.

McCutcheon, J. P. and Moran, N. A. Functional convergence in reduced genomes of bacterial symbionts spanning 200 my of evolution. *Genome Biology and Evolution*, 2:708–718, 2010.

McCutcheon, J. P. and Moran, N. A. Extreme genome reduction in symbiotic bacteria. *Nature Reviews Microbiology*, 10(1):13–26, 2011.

McCutcheon, J. P., McDonald, B. R., and Moran, N. A. Convergent evolution of metabolic roles in bacterial co-symbionts of insects. *Proceedings of the National Academy of Sciences*, 106(36):15394–15399, 2009.

McFall-Ngai, M. J. The importance of microbes in animal development: Lessons from the squid-vibrio symbiosis. *Annual Review of Microbiology*, 68(1):177–194, 2014.

Metz, J., Nisbet, R., and Geritz, S. How should we define 'fitness' for general ecological scenarios? *Trends in Ecology & Evolution*, 7(6):198–202, 1992.

Moran, N. A. Accelerated evolution and muller's rachet in endosymbiotic bacteria. *Proceedings of the National Academy of Sciences*, 93(7):2873–2878, 1996.

Moran, N. A. and Bennett, G. M. The tiniest tiny genomes. *Annual Review of Microbiology*, 68(1):195–215, 2014.

Moran, N. A., Plague, G. R., Sandström, J. P., and Wilcox, J. L. A genomic perspective on nutrient provisioning by bacterial symbionts of insects. *Proceedings of the National Academy of Sciences*, 100(suppl_2):14543–14548, 2003.

Nguyen, P. L. and Baalen, M. v. On the difficult evolutionary transition from the free-living lifestyle to obligate symbiosis. *PLOS ONE*, 15(7):e0235811, 2020.

Nussbaumer, A. D., Fisher, C. R., and Bright, M. Horizontal endosymbiont transmission in hydrothermal vent tubeworms. *Nature*, 441(7091):345–348, 2006.

Otto, S. P. and Day, T. *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution*. Princeton University Press, Princeton, New Jersey, 2007.

Patel, M. and West, S. Microbial warfare and the evolution of symbiosis. *Biology Letters*, 18(12), 2022.

Prechtl, J., Kneip, C., Lockhart, P., Wenderoth, K., and Maier, U.-G. Intracellular spheroid bodies of rhopalodia gibba have nitrogen-fixing apparatus of cyanobacterial origin. *Molecular Biology and Evolution*, 21(8):1477–1481, 2004.

Press, W. H. and Dyson, F. J. Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences*, 109(26):10409–10413, 2012.

Queller, D. C. Relatedness and the fraternal major transitions. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 355(1403):1647–1655, 2000.

Sagan, L. On the origin of mitosing cells. *Journal of Theoretical Biology*, 14(3):225–IN6, 1967.

Scarborough, C. L., Ferrari, J., and Godfray, H. C. J. Aphid protected from pathogen by endosymbiont. *Science*, 310(5755):1781–1781, 2005.

Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y., and Ishikawa, H. Genome sequence of the endocellular bacterial symbiont of aphids buchnera sp. APS. *Nature*, 407(6800):81–86, 2000.

Szathmáry, E. Toward major evolutionary transitions theory 2.0. *Proceedings of the National Academy of Sciences USA*, 12(33):10104–10111, 2015.

Szathmáry, E. and Smith, J. M. The major evolutionary transitions. *Nature*, 374(6519):227–232, 1995.

Thacker, R. W. Impacts of shading on sponge-cyanobacteria symbioses: A comparison between host-specific and generalist associations. *Integrative and Comparative Biology*, 45(2):369–376, 2005.

van den Driessche, P. and Watmough, J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180(1-2):29–48, 2002.

van Gestel, J. and Tarnita, C. E. On the origin of biological construction, with a focus on multicellularity. *Proceedings of the National Academy of Sciences USA*, 114(42):11018–11026, 2017.

Vogel, K. J. and Moran, N. A. Sources of variation in dietary requirements in an obligate nutritional symbiosis. *Proceedings of the Royal Society B: Biological Sciences*, 278(1702):115–121, 2010.

West, S. A., Fisher, R. M., Gardner, A., and Kiers, E. T. Major evolutionary transitions in individuality. *Proceedings of the National Academy of Sciences*, 112(33):10112–10119, 2015.

Woyke, T., Teeling, H., Ivanova, N. N., Huntemann, M., Richter, M., Gloeckner, F. O., Boffelli, D., Anderson, I. J., Barry, K. W., Shapiro, H. J., Szeto, E., Kyrpides, N. C., Mussmann, M., Amann, R., Bergin, C., Ruehland, C., Rubin, E. M., and Dubilier, N. Symbiosis insights through metagenomic analysis of a microbial consortium. *Nature*, 443(7114):950–955, 2006.

Wu, M., Sun, L. V., Vamathevan, J., Riegler, M., Deboy, R., Brownlie, J. C., McGraw, E. A., Martin, W., Esser, C., Ahmadinejad, N., Wiegand, C., Madupu, R., Beanan, M. J., Brinkac, L. M., Daugherty, S. C., Durkin, A. S., Kolonay, J. F., Nelson, W. C., Mohamoud, Y., Lee, P., Berry, K., Young, M. B., Utterback, T., Weidman, J., Nierman, W. C., Paulsen, I. T., Nelson, K. E., Tettelin, H., O'Neill, S. L., and Eisen, J. A. Phylogenomics of the reproductive parasite wolbachia pipientis wMel: A streamlined genome overrun by mobile genetic elements. *PLoS Biology*, 2(3):e69, 2004.

Wujek, D. E. Intracellular bacteria in the blue-green alga pleurocapsa minor. *Transactions of the American Microscopical Society*, 98(1):143, 1979.

Zachar, I. and Szathmáry, E. Breath-giving cooperation: critical review of origin of mitochondria hypotheses. *Biology Direct*, 12(1), 2017.

Zachar, I., Szilágyi, A., Számadó, S., and Szathmáry, E. Farming the mitochondrial ancestor as a model of endosymbiotic establishment by natural selection. *Proceedings of the National Academy of Sciences*, 115(7), 2018.

Zientz, E., Dandekar, T., and Gross, R. Metabolic interdependence of obligate intracellular bacteria and their insect hosts. *Microbiology and Molecular Biology Reviews*, 68(4):745–770, 2004.

A Appendix

A.1 Mathematical details of the model

First, we make explicit the interpretations of the traits Ω_i and α_i. Recall that the obligacies Ω_i formalise a tradeoff between individual and collective reproduction. Further, the obligacy of an independent host does not affect the growth rate of the independent symbiont; the two traits only interact through their impact on the growth rate of the collective. Hence, if the growth rates of the three types are f_H, f_S, f_C, then the corresponding constraints are

$$\frac{\partial f_i}{\partial \Omega_i} \leq 0\,; \qquad \frac{\partial f_C}{\partial \Omega_i} \geq 0\,; \qquad \frac{\partial f_i}{\partial \Omega_j} = 0\,, \qquad i \neq j \in \{H, S\}\,. \tag{8}$$

Secondly, the adhesions α_i formalise a tradeoff between synchronised and asynchronised reproduction. Therefore a higher adhesion increases the rate of association, increases the rate of collective reproduction, and decreases the rate of dissociation. If a is the rate of association and d is the rate of dissociation, we impose the constraints

$$\frac{\partial f_C}{\partial \alpha_i} \geq 0\,; \qquad \frac{\partial a}{\partial \alpha_i} \geq 0\,; \qquad \frac{\partial d}{\partial \alpha_i} \leq 0\,, \qquad i \in \{H, S\}\,. \tag{9}$$

The population dynamical equations are given in the main text, but we reproduce the exponential model here for convenience:

$$\dot{x}_H = f_H x_H \left(1 - \frac{x_H}{K_H}\right) - a x_H x_S + d x_C, \tag{10a}$$

$$\dot{x}_S = f_S x_S \left(1 - \frac{x_S}{K_S}\right) - a x_H x_S + d x_C, \tag{10b}$$

$$\dot{x}_C = f_C x_C + a x_H x_S - d x_C. \tag{10c}$$

Next, to compute the invasion fitness of an arbitrary mutant, suppose that there is a resident population with host and symbiont traits (Ω_H, α_H) and (Ω_S, α_S), respectively. After a sufficiently long time, a mutant arises when the resident population is at equilibrium. Due to the separation between ecological and evolutionary timescales, there can only be one mutant at a time, without loss of generality say a host mutant. This mutant has trait value $(\tilde{\Omega}_H, \tilde{\alpha}_H)$ different from the resident, giving rise to different values of the parameters $\tilde{f}_H, \tilde{f}_C, \tilde{a}, \tilde{d}$ for the mutant host. The invasion condition is formalised as follows: we augment the model (1) with additional equations tracking each type that arises due to the introduction of a mutant. Here, this always amounts to two more – the mutant host y_H itself and the mutant host-resident

symbiont collective y_C. The augmented model takes the form

$$\dot{x}_H = f_H x_H \left(1 - \frac{x_H + y_H}{K_H}\right) - a x_H x_S + d x_C \,, \tag{11a}$$

$$\dot{x}_S = f_S x_S \left(1 - \frac{x_S}{K_S}\right) - a(x_H + y_H) x_S + d(x_C + y_C) \,, \tag{11b}$$

$$\dot{x}_C = f_C x_C + a x_H x_S - d x_C \,, \tag{11c}$$

$$\dot{y}_H = \tilde{f}_H x_H \left(1 - \frac{x_H + y_H}{K_H}\right) - \tilde{a} y_H x_S + \tilde{d} y_C \,, \tag{11d}$$

$$\dot{y}_C = \tilde{f}_C y_C + \tilde{a} y_H x_S - \tilde{d} y_C \,. \tag{11e}$$

The invasion fitness is interpreted to be the growth rate of a rare mutant in a resident population that is at stable equilibrium (Metz et al., 1992). Suppose the resident equilibrium is (x_H^*, x_S^*, x_C^*). Mathematically, the invasion fitness is the dominant eigenvalue of the "sub-Jacobian" corresponding to the mutant equations, evaluated at (x_H^*, x_S^*, x_C^*):

$$\begin{bmatrix} \tilde{f}_H x_H^* (1 - \frac{x_H^* + 2y_H}{K_H}) - \tilde{a} x_S^* & \tilde{d} \\ \tilde{a} x_S^* & \tilde{f}_{HS} - \tilde{d} \end{bmatrix} .$$

If it is positive, the equilibrium is destabilised in the presence of the mutant and the mutant can invade; if it is negative, the mutant goes to extinction (Dieckmann and Law, 1996; Dercole and Rinaldi, 2008). This quantity is unwieldy and not amenable to mathematical analysis in our case, so we turn to other methods to quantify (in)stability of the resident population in response to mutants. In particular, we use the next-generation theorem, which gives, under some conditions, an alternate characterisation of the standard stability condition of all eigenvalues having negative real parts (van den Driessche and Watmough, 2002; Hurford et al., 2010).

A.2 Correspondence to a symmetric mutualism model

It is useful to establish the ecological context of our model and understand exactly how the presence of a collective affects the independent types. One can do this by comparing an extreme version (colloquially, "take a limit") of our model to other well-understood models in the following way. In all of our analyses, we are only interested in long-term symbioses like the nutrient-exchange symbioses in anaerobic ciliates (Graf et al., 2021), and not symbioses where the interaction is short-lived, like the parasite-cleaning mutualisms (Losey et al., 1999). If the interaction between the participants is ephemeral, the collective does not have as much of an independent existence. In the limiting case of instantaneous interaction, we can set $\dot{x}_C = 0$. This has the effect of reducing the system of equations (1) to

$$\dot{x}_H = f_H x_H (1 - \frac{x_H}{K_H}) + \left(\frac{ad}{d - f_C} - a\right) x_H x_S \,, \tag{12a}$$

$$\dot{x}_S = f_S x_S (1 - \frac{x_S}{K_S}) \; + \left(\frac{ad}{d - f_C} - a\right) x_H x_S \,. \tag{12b}$$

It is easily read off from this set of equations that there is now a constant k preceding the cross-term $x_H x_S$ in both equations, with

$$k = a\left(\frac{d}{d - f_C} - 1\right).$$

Since $f_C > 0$ by assumption, $d > d - f_C$ and therefore $k > 0$ as long as $d > f_C$. This last condition implies that the interaction is beneficial as long as $d > f_C$ and becomes detrimental to both species when the collective growth rate f_C is "too high" - larger than d. Reinterpreting this cross-term $kx_H x_S$ as arising from a Lotka-Volterra-type interaction, one arrives at the conclusion that the formation of a collective, in this "limit", is equivalent to a symmetric mutualism as long as the dissociation is strong enough. If it is too weak, it goes from being a mutualism to being an interaction that is negatively affecting the growth of both parties. That it is symmetric in either case reflects the fact that with every collective reproduction event, there is exactly one new host and one new symbiont. This shows that collective reproduction can indeed be thought of as aligning reproductive interests of the host and symbiont. In other words, forming the collective benefits the host and symbiont. This is because when the collective reproduces, there is necessarily a new copy of both the host and symbiont. Like in a mutualism, the investment of one of the types (H or S) in collective reproduction therefore helps the other.

A.3 Dissociation is a course-graining of more realistic processes

Here, we detail how our model is simpler in its description of the processes that generate free-living individuals H and S from host-symbiont collectives C than previous studies (Law and Dieckmann, 1998; Nguyen and Baalen, 2020). In the exponential model of population dynamics above, the parameter d is the rate constant of the reaction $C \to H + S$. We will describe in this section how this reaction is a coarse-graining of two processes: asynchronous death of either host or symbiont while in a collective, and death of the host/symbiont in a collective. Note that both of these processes give rise, given a collective individual, to one new free-living individual.

More explicitly, let b_H be the rate at which a host that is part of collective gives birth aysnchronously to a free-living host, and similarly b_S. Then let d_H be the rate at which hosts that are part of a collective die. Note that the latter process gives rise to a free-living symbiont and not a host. Let d_S be the rate of the analogous process giving rise to free-living hosts. Incorporating these rates instead of dissociation gives rise to the system of equations

$$\dot{x}_H = f_H x_H \left(1 - \frac{x_H}{K_H}\right) - a x_H x_S + (d_S + b_H) x_C, \tag{13a}$$

$$\dot{x}_S = f_S x_S \left(1 - \frac{x_S}{K_S}\right) - a x_H x_S + (d_H + b_S) x_C, \tag{13b}$$

$$\dot{x}_C = f_C x_C + a x_H x_S - (d_H + d_S) x_C. \tag{13c}$$

Note that we are not changing anything related to the process described by the rate of successful collective-forming encounter rate a. We recover the above model if we assume that

$d_S = b_H = d_H = b_S$ and they are equal to some constant $d/2$. Biologically, this implies that the rate of dying while part of the collective is identical for the host and symbiont, and that it is equal also to the asynchronous birth rate. In the most general model, one would consider d_H and d_S as decreasing functions of the obligacy Ω_H and Ω_S respectively. In other words, the higher the dependence of the host on the collective, one would expect that it dies less frequently when it is part of the collective. Secondly, the higher the adhesion, the lower the rate of asynchronous birth. Abstracting all these processes into one reaction gives us analytical power, admittedly at the cost of detail. There will be no discussion here of the consequences of this assumption - this is partly treated in the Conclusions section.

A.4 Feasibility and linear stability of the internal fixed point

The exponential model of population dynamics has four fixed points for the population densities, three of which are trivial. These trivial fixed points are located on the boundary at $(0,0,0), (K_H,0,0)$ and $(0,K_S,0)$. The internal fixed point of the exponential model is given by

$$x_H^* = \frac{K_H f_S(f_C - d)(aK_S f_C + (d - f_C)f_H)}{a^2 K_H K_S f_C^2 - f_H f_S (d - f_C)^2}\,, \tag{14a}$$

$$x_S^* = \frac{K_S f_H(f_C - d)(aK_H f_C + (d - f_C)f_S)}{a^2 K_H K_S f_C^2 - f_H f_S (d - f_C)^2}\,, \tag{14b}$$

$$x_C^* = \frac{a f_S f_H K_S K_H (d - f_C)(aK_S f_C + (d - f_C)f_H)(aK_H f_C + (d - f_C)f_S)}{[a^2 K_H K_S f_C^2 - f_H f_S (d - f_C)^2]^2}\,. \tag{14c}$$

To establish the condition for feasibility, first notice that any fixed point must satisfy

$$x_C^* = \frac{a}{d - f_C} x_H^* x_S^*\,, \tag{15}$$

which arises from setting the dynamical equation for x_C to zero. First, notice that neither of x_H^* and x_S^* can be negative since we are searching for feasible equilibria. Therefore, we require $d > f_C$ since otherwise x_C^* cannot be positive. The expressions for the host and symbiont population size have a similar form - they are after all symmetric in their attributes. Notice that the numerators of x_H^* and x_S^* are negative whenever $d > f_C$. Therefore the expressions x_H^*, x_S^* on the whole are positive whenever the denominator is also (strictly) negative. Then notice that they have the same denominator. The negativity of the denominator, upon rearranging to solve for d, exactly gives rise to the feasibility bound

$$d > f_C \left(1 + a\sqrt{\frac{K_H K_S}{f_H f_S}}\right)\,. \tag{16}$$

Notice that this lower bound on d is strictly larger than f_C. The next step is to find conditions for stability of the equilibrium. The Jacobian of this system is given by

$$J = \begin{pmatrix} f_H(1 - \frac{2x_H}{K_H}) - ax_S & -ax_H & d \\ -ax_S & f_S(1 - \frac{2x_S}{K_S}) - ax_H & d \\ ax_S & ax_H & f_{HS} - d \end{pmatrix}\,. \tag{17}$$

In the case of the trivial fixed point $(0, 0, 0)$, the Jacobian above becomes a diagonal matrix. The eigenvalues of a diagonal matrix are exactly the elements on its diagonal, and two of these elements are f_H and f_S. We assume that these quantities are always positive in the entirety of this work, so this fixed point is never stable. It is not as easy for the other fixed points: the roots computed directly from the general characteristic equation are required and difficult to analyse, and we therefore resort to the Routh-Hurwitz criteria (Edelstein-Keshet, 2005). This is a set of criteria on the coefficients in the characteristic equation of J. Specifically, let the characteristic equation be given by

$$\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0\,.$$

These numbers are computable directly from J, and some of them are commonly known. For example, a_1 above is the negative of the trace of J and a_3 is the negative of its determinant. The Routh-Hurwitz criteria state that all eigenvalues of J i.e., all roots of the above polynomial, have negative real part if and only if

$$a_1 > 0\,, \tag{18}$$

$$a_3 > 0\,, \tag{19}$$

$$a_1 a_2 > a_3\,. \tag{20}$$

We shall not give the expressions for these coefficients a_i here, they can be found in the associated Mathematica notebook.

For the two one-species equilibria $(K_H, 0, 0)$ and $(0, K_S, 0)$, the criteria can be verified easily: as long as the focal fixed point is feasible, these equilibria are unstable - one of the criteria above (20) fails since $d > f_C$. In particular, the feasibility bound being satisfied implies that $a_3 < 0$ in these two cases. Further, there is a parameter range where none of the fixed points are stable: this is exactly when

$$f_C < d < f_C\left(1 + a\sqrt{\frac{K_H K_S}{f_H f_S}}\right)\,.$$

This is when the feasibility bound is violated (which is the right side of the above chain of inequalities), but not as strongly as to make $f_C > d$ (left side of chain). Therefore, the internal fixed point is rendered unstable, as well as the trivial fixed points.

For the internal equilibrium, the picture is much more involved. We cannot say much from the coefficients a_i themselves, but it can be shown that the subcriterion $a_3 > 0$ exactly corresponds, upon some rearrangement, to the inequality that we have termed the feasibility bound (2). This proves that feasibility is necessary for stability.

Further analytics is difficult, and we therefore resort to numerically computing the maximum real part of the eigenvalues of J across a biologically reasonable range of parameter space. The results of this exercise is shown in Figure 5. This shows that the internal fixed point of the exponential model is, in fact, stable in the region of parameter space that we are interested in, and therefore we proceed with further analysis.

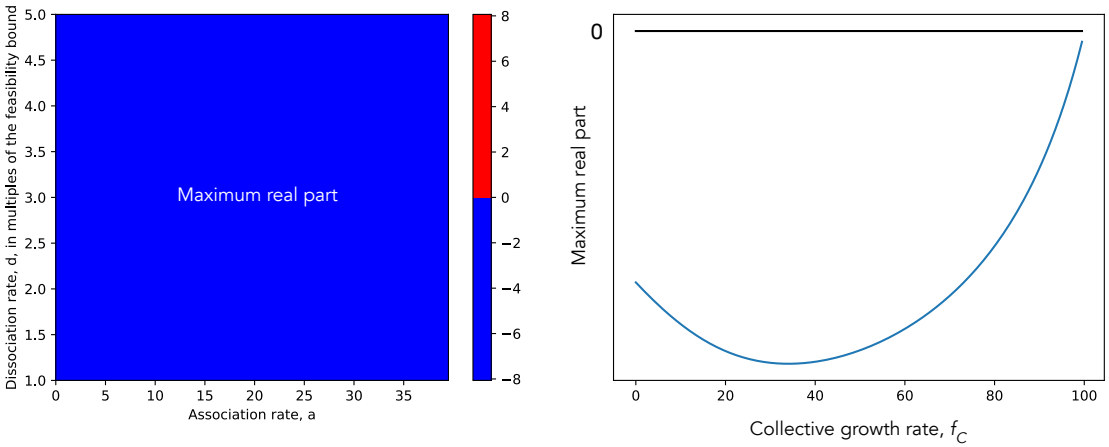

Figure 5: **The internal fixed point is stable.** Here we perform a numerical evaluation of the stability condition of the internal fixed point of the exponential model. In particular, we find and plot in two cases, the maximum real part of the eigenvalues of J, the Jacobian of the flow (1). **Common parameter values for both panels.** $K_H = 100, K_S = 200, f_H = 8, f_S = 20$. **(a)** Red if unstable, blue if stable. Since we can show already that the feasibility bound is necessary for stability (see main text of Appendix A.4), we wish to only check stability when the feasibility bound is already satisfied. This panel is therefore a little unorthodox –the x-axis is the usual, but the y-axis is in multiples above the feasibility bound for the respective x value. Let $d_f(a) = f_C(1 + a\frac{K_H K_S}{f_H f_S})$ be the equalising value of d at the feasibility bound when all other parameters are kept fixed. Then in the heatmap, for a given value of the horizontal variable a, the vertical strip at this horizontal position ranges from $[d_f(a), 5 \times d_f(a)]$. This panel therefore shows that the fixed point is stable for a large range of d values once the feasibility bound is crossed. **Parameter values.** $f_C = 10, a \in [0, 40], d \in [d_f(a), 5 \times d_f(a)]$ **(b)** This shows that the fixed point is stable across a range of f_C values. **Parameter values.** $a = 0.1, f_C \in [0, 5f_S = 100], d = a + f_{C,\max}(1 + a\sqrt{\frac{C_H C_S}{f_H f_S}})$ to maintain feasibility.

A.5 Application of the next-generation theorem in deriving the invasion fitness

In this section, we detail the derivation of the invasion criterion (3). The intuitive picture is as follows: the resident population is at an attractor of the population dynamics, and then a mutant arises. We want to understand how the fate of a mutant depends on its trait value. It can be shown (Dercole and Rinaldi, 2008) that if the presence of a small number of mutant individuals destabilises the resident equilibrium, the mutant will fix in the population and hence replace the resident. Therefore, we must study the dominant eigenvalue of the Jacobian coming from the set of equations (11). For a mutant host, this Jacobian (when the mutant is rare i.e., $y_H = 0$) takes the form

$$J = \begin{pmatrix} f_H(1-\frac{2x_H}{K_H}) - ax_S & -ax_H & d & -\frac{f_H x_H}{K_H} & 0 \\ -ax_S & f_S(1-\frac{2x_S}{K_S}) - ax_H & d & -\tilde{a}x_S & \tilde{d} \\ ax_S & ax_H & f_{HS} - d & 0 & 0 \\ 0 & 0 & 0 & \tilde{f}_H\left(1-\frac{x_H}{K_H}\right) & \tilde{d} \\ 0 & 0 & 0 & \tilde{a}x_S & \tilde{f}_C - \tilde{d} \end{pmatrix}. \tag{21}$$

Due to the block-upper-triangular form of J, its eigenvalues are given entirely by the eigenvalues of only its diagonal blocks. Note that the block-triangular form is generic for Jacobians derived from such an invasion analysis. The core assumption that makes this true is that the mutant being considered arose just once - the resident does not continuously make such mutants. This in turn arises from the separation of ecological and evolutionary timescales.

We know the top-left 3×3 block has all eigenvalues with negative real part since we began with a stable population dynamical equilibrium. Therefore, we need only check the dominant eigenvalue of the submatrix corresponding to the mutant i.e., the bottom right 2×2 block. We want to search for the condition for instability of the equilibrium i.e. when the dominant eigenvalue has real part larger than 0. The eigenvalues of a 2×2 matrix A can be easily computed as the roots of $det(A - \lambda I) = 0$. However, this quantity is not analytically tractable in our case. We therefore turn to other methods to quantify (in)stability of the resident. In particular, we use the next-generation theorem, which has roots in mathematical epidemiology, but is generally applicable (Hurford et al., 2010). This result gives, under some conditions, an alternate characterisation of the standard stability condition of all eigenvalues having negative real parts. This substitute is obtained by calculating the eigenvalues of a matrix arising from a particular kind of decomposition of the Jacobian of our system - more generally, a decomposition of the matrix corresponding to a system of first-order linear differential equations. Concretely, suppose we have a linear system of ODEs given by

$$\dot{\mathbf{x}} = A\mathbf{x}\,,$$

and the matrix A can be decomposed in terms of two matrices F and V as $A = F - V$, with some conditions on F and V. Now define $s(M)$ to be the maximum real part of all the eigenvalues of a matrix M, and $\rho(M)$ to be the maximum modulus of all the eigenvalues. The next-generation theorem (van den Driessche and Watmough, 2002) states that $s(A) = s(F - V)$

has the same sign as $\rho(F.V^{-1}) - 1$. The quantity $\rho(F.V^{-1})$ is exactly the R_0 value usually associated to the spread of virulent pathogens, and describes the average number of offspring one mutant individual sires. If it is larger than 1, then the mutant spreads. Therefore, if we can find a satisfactory decomposition $J = F - V$ of J in Equation (21), the fixed point in question is stable if and only if $\rho(F.V^{-1}) - 1 > 0$. In many cases, there is a recipe for finding the matrices F and V stemming loosely from the intuition that F is the matrix of rates of appearance of j individuals per individual of type i, and V the matrix of disappearance rates, where appearance and disappearance include both birth-death and migration between different compartments. Recall here that we are studying the Jacobian because it controls the first-order i.e., linearized behaviour of the dynamical system around a fixed point. This alternative method is useful because the latter quantity is sometimes more analytically tractable and interpretable than $s(J)$, which is traditionally defined as the invasion fitness. But it is identical to use them interchangeably since we only care about positivity and the next-generation theorem shows that the traditional invasion fitness and $\rho(F.V^{-1}) - 1$ have the same sign.

For our purposes, we use the decomposition

$$F = \begin{bmatrix} \tilde{f}_H & \tilde{d} \\ 0 & 0 \end{bmatrix}, V = \begin{bmatrix} \tilde{a}x_S + \frac{\tilde{f}_H(x_H+2y_H)}{K_H} & 0 \\ -\tilde{a}x_S & \tilde{d} - \tilde{f}_C \end{bmatrix}, V^{-1} = \begin{bmatrix} \frac{K_H}{aK_H\tilde{f}_S+\tilde{f}_H(x_H+2y_H)} & 0 \\ \frac{ax_SK_H}{(\tilde{d}-\tilde{f}_C)(aK_H\tilde{f}_S+\tilde{f}_H(x_H+2y_H))} & \frac{1}{\tilde{d}-\tilde{f}_C} \end{bmatrix}$$

which satisfies all the requirements – all entries of F are positive, all eigenvalues of $-V$ are negative, and all elements of V^{-1} are negative. We also require that $\tilde{f}_C < \tilde{d}$ in the mutant host case just as has been shown to be necessary for existence of a feasible resident equilibrium. We now restrict ourselves to finding the condition under which $\rho(F.V^{-1}) - 1 > 0$. This directly leads to the invasion criterion (3). See the associated Mathematica notebook for more details.

A.6 Long-term evolutionary dynamics via the canonical equation

Suppose, for simplicity, that the mutant has obligacy $\widetilde{\Omega}_H$ and the resident has obligacy Ω_i. The invasion fitness is given by

$$s_i(\widetilde{\Omega}_i, \Omega_i) = \tilde{f}_i \left(1 - \frac{d}{\tilde{f}_C}\right) - f_i \left(1 - \frac{d}{f_C}\right).$$

The invasion fitness determines the fate of a given mutant, and therefore induces a certain path for the evolutionary dynamics of the trait value. This path is given by the canonical equation of adaptive dynamics (Dieckmann and Law, 1996), which is an ordinary differential equation describing the macroevolutionary change, and the obligacy of species i varies as

$$\frac{d\Omega_i}{dt} = \frac{1}{2}\mu_i\nu_i\, x_i^*(\Omega_H, \Omega_S)\frac{\partial s_i(\widetilde{\Omega}_i, \Omega_i)}{\partial \widetilde{\Omega}_i}, \tag{22}$$

where μ_i, ν_i are respectively the mutation rate and the variance of the mutation step, x_i^* is the equilibrium population size, and partial derivative is the fitness gradient. By the constraints (8) and the fact that $d > f_C$, one sees that the fitness gradient is uniformly positive.

A similar analysis can be accomplished when a mutant has both different obligacy and different adhesion values. Again tilde-d quantities are associated to the mutant. The invasion fitness is given by

$$s((\widetilde{\Omega}_i, \tilde{\sigma}_i), (\Omega_i, \sigma_i)) = \frac{\tilde{f}_i}{\tilde{a}} \left(1 - \frac{\tilde{d}}{\tilde{f}_C}\right) - \frac{f_i}{a} \left(1 - \frac{d}{f_C}\right) .$$

We write now the canonical equation for a multi (here, two)-dimensional trait:

$$\frac{d(\Omega_i, \alpha_i)}{dt} = \frac{1}{2}\mu_i \;\; x_i^*(\Omega_H, \alpha_H, \Omega_S, \alpha_S)\mathbf{M_i} \cdot \nabla_i s_i((\widetilde{\Omega}_i, \tilde{\alpha}_i), (\Omega_i, \alpha_i)) . \tag{23}$$

Here μ_i is the mutation rate for type i, x_i^* is its equilibrium population size at the given trait values, and $\mathbf{M_i}$ is the variance-covariance matrix of the traits of type i which encodes correlations between these traits. s_i is the invasion fitness for type i, and the gradient is taken with respect to all of its traits. This more general equation is similar in spirit to the earlier equation (22), with changes accounting for the now arbitrary (in $\mathbb{N}$) dimension of the trait.

To show that the traits increase, it is sufficient to show that all the entries of $\nabla_i s((\widetilde{\Omega}_i, \tilde{\alpha}_i), (\Omega_i, \alpha_i))$ are positive. Under the constraints on the partial derivatives that we set up initially, and some additional similar constraints on the mapping from traits to ecological parameters, it is straightforward to show this.

A.7 The impossibility of branching when the invasion fitness is separable

Definition 1 (Separable invasion fitness)**.** *Let $s(y, x)$ be the invasion fitness of a mutant with trait y in the background of a resident population of trait value x, with traits taking value in some set $\mathcal{T}$. We say that $s(\cdot, \cdot)$ is separable if there exists a function $\mathcal{F} : \mathcal{T} \to \mathbb{R}_{\geq 0}$ such that one can write $s(y, x) = \mathcal{F}(y) - \mathcal{F}(x)$.*

Recall that a mutant invades when its invasion fitness is positive. Therefore, if the invasion fitness is separable, a mutant with trait y invades exactly when $\mathcal{F}(y) > \mathcal{F}(x)$. We notice, then, that an invasion fitness is separable when there is some function $\mathcal{F}(\cdot)$ that acts like a fitness - if y has a higher value of $\mathcal{F}$, then it invades, and if it has a lower value of $\mathcal{F}$ it goes to extinction. Evolutionary trajectories hence maximise this quantity, and therefore in this sense the fitness landscape does not change. If the invasion fitness is separable, one can construct a fitness landscape $\mathcal{F}$ on the trait space such that all permitted trajectories climb up this landscape. Now we prove the observation that we made in the main text regarding the invasion criterion.

Theorem 1. *Suppose $s(y, x)$ is the invasion fitness describing the evolutionary dynamics of a given system, and that it is separable. Then a given strategy x^* is uninvadable if and only if it is convergence stable.*

Proof. First we define the function

$$G(x) = \left.\frac{\partial s(y, x)}{\partial y}\right|_{y=x}$$

which is the fitness gradient for an invading mutant of trait value y against a resident with trait x. We shall use the fact that for any (C^2) function,

$$\frac{dG(x)}{dx} = \frac{\partial^2 s}{\partial x \partial y} + \frac{\partial^2 s}{\partial y^2}$$

But since s is separable, there is a function $\mathcal{F}$ such that $s(y, x) = \mathcal{F}(y) - \mathcal{F}(x)$. This implies that $\frac{\partial s}{\partial y} = \frac{\partial \mathcal{F}(y)}{\partial y}$ since $\mathcal{F}(x)$ does not depend on y, and then $\frac{\partial^2 s}{\partial x \partial y} = \frac{\partial}{\partial x}\mathcal{F}(y)$ which is obviously zero. So we have shown that

$$\frac{dG(x)}{dx} = \frac{\partial^2 s}{\partial y^2}$$

The left-hand side determines convergence stability and the right-hand side determines (un)invadability, and since they are equal, these notions cannot diverge! □

Recall that a singular strategy is called a branch point i.e., a point at which evolutionary branching occurs, if it is convergence stable but invadable. A corollary of the above result, then, is that separable invasion fitnesses do not admit branch points - evolutionary branching is impossible. It is perhaps useful to think in terms of the contrapositive: if evolutionary branching is possible, then the invasion fitness is *not* separable.

This is because for a singular strategy to be a branch point, it must by definition be convergence stable but invadable by similar mutants (Geritz et al., 1998). This is possible in general because the conditions to reach a maximum (convergence stability) and to stay there (invadability) are not the same. However, if one can identify a quantity that uniformly increases along all permissible evolutionary trajectories, these notions cannot be different. The reason why is simple and goes as follows. When such a quantity exists, uninvadable points are all convergence stable since they can be reached by trajectories that maximise this quantity. Conversely, convergence stable singular strategies are always uninvadable since any trajectory that converges there must maximise this quantity.

A.8 Independent evolution of the adhesion

As stated in the main text, there is a similar invasion criterion for adhesion evolution. A mutant with trait value $\tilde{\alpha}_i$ invades in a resident population of trait α_i if

$$\frac{\tilde{f}_C}{\tilde{d}} > \frac{f_C}{d} \,. \tag{24}$$

Relatedly, the invasion fitness of the mutant is given by

$$s(\tilde{\alpha}_i, \alpha_i) = \frac{\tilde{f}_C}{\tilde{d}} - \frac{f_C}{d} \,. \tag{25}$$

The canonical equation for adhesion evolution thus reads

$$\frac{d\alpha_i}{dt} = \frac{1}{2}\mu_i \nu_i \, x_i^*(\alpha_H, \alpha_S) \cdot \left[\frac{1}{\tilde{d}} \frac{\partial \tilde{f}_C}{\partial \tilde{\alpha}_i} - \frac{\tilde{f}_C}{\tilde{d}^2} \frac{\partial \tilde{d}}{\partial \tilde{\alpha}_i} \right] \Bigg|_{\tilde{\alpha}_i = \alpha_i} \,. \tag{26}$$

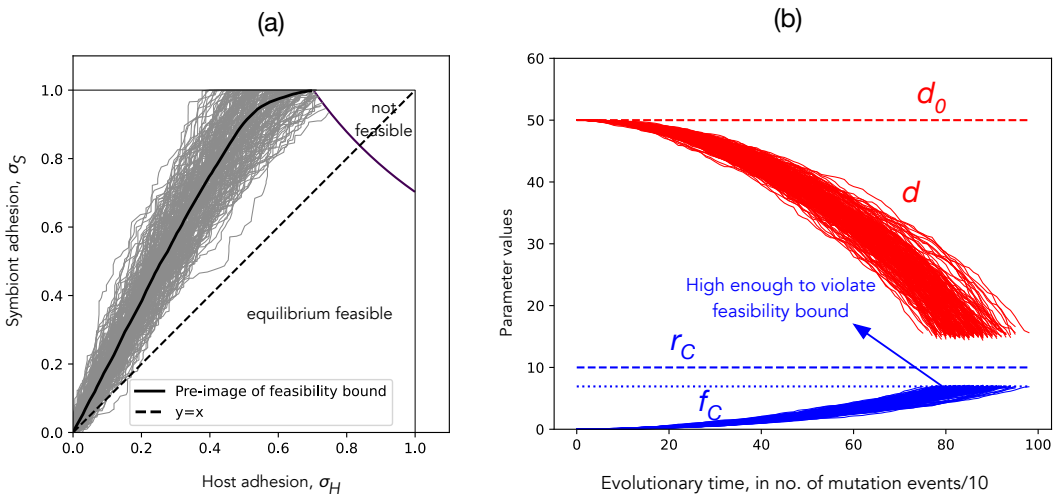

Figure 6: **Independent evolutionary trajectories of adhesion. (a)** For the mapping in (27), we simulated 100 independent evolutionary trajectories. The trajectories monotonically increase with α_S increasing faster, and then stop when the fixed point is no longer stable. This behaviour is consistent with the analytical predictions made associated with Equation (26). In the unstable region, the collective population density increases to infinity and, due to dissociation, maintains a vanishingly minute population of independent hosts and symbionts. **(b)** Using the mapping (27), we can also visualise the dynamics of the growth rates. We observe that the independent growth rates decrease, and the collective growth rate increases until it becomes high enough that the feasibility bound is violated. **Parameter values**. $f_H = 8, f_S = 30, r_C = 10, K_H = 100, K_S = 200, a = 0.1, d_0 = 50$.

Since f_C increases with adhesion and d decreases, the fitness gradient is uniformly positive. All the coefficients before it are positive as well, and so we have shown that the adhesions must also monotonically increase.

Like the obligacies, adhesion evolution is also impeded by the feasibility bound. This is shown in Figure 6, for the mapping from traits to ecological parameters given by

$$f_C(\alpha_H, \alpha_S) = r_C \alpha_H \alpha_S \,, \tag{27a}$$

$$d(\alpha_H, \alpha_S) = d_0(1 - \alpha_H \alpha_S) \,. \tag{27b}$$

Notice that the qualitative behaviour of the adhesions is identical to that of the obligacies, as we have remarked and explained in the main text.

A.9 Trajectories in the dependence-cohesion plane

In this section, we provide the supporting information for the result that mutual dependence evolves faster than reproductive cohesion. We will first provide an explanation - in terms of the model and then biologically - for why it is adaptive to invest more in obligacy than adhesion. Second, we will show that the bias towards mutual dependence is robust to other minor choices that we made in arriving there.

For providing an explanation, consider the trait-mapping that we have utilised in the main text, given below.

$$f_H(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_H(1 - \Omega_H) \tag{28a}$$
$$f_S(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_S(1 - \Omega_S) \tag{28b}$$
$$f_C(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_C\Omega_H\Omega_S\alpha_H\alpha_S \tag{28c}$$
$$d(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = d_0(1 - \alpha_H\alpha_S) \tag{28d}$$
$$a(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = a \in \mathbb{R}_{\geq 0} \tag{28e}$$

We claimed in the main text that since a change in α_H has functional effect that depends also on α_S (and vice versa), it has lesser adaptive benefit than a change in Ω_i. To show this, notice that two important choices were made in arriving at the above trait mapping – first that the independent growth of the host and symbiont depends only on their own obligacy; second, the dissociation rate depends on both α_H and α_S. However, in general, the independent growth rates can depend on either (i) only own obligacy or (ii) both obligacies; the dissociation rate can depend on (I) only one adhesion and (II) both adhesions. The "default" map that was motivated by biological reasoning corresponds to the combination i-II above. We shall below give the trait mapping of all other combinations, with each followed by the induced evolutionary trajectories of the traits Ω_i and α_i.

There is another note to be made here: when we make the association and dissociation rates dependent only on one of α_H and α_S, we pick the symbiont trait without loss of generality. This does not have a qualitative effect, and the exact effect will be made clear when necessary. Further, we fix the effect of the obligacies and adhesions on f_C; this does not change because our mental model of an endosymbiosis still stays the same - increasing both of these traits improves collective growth rate. The association rate a will remain a constant here like the rest of this work.

In all of the below cases, there will be some commonalities - all four traits always increase, and within each trait pair the symbiont trait increases faster. This is due to the fact that exponential growth of the collective and the difference in parameters are not changed over the course of this exercise. The traits increase due to the exponential growth of the collective and the fact that f_C can be increased by increased by any one of the four traits. The symbiont trait increases faster than its corresponding host trait due to the different generation times and carrying capacities.

Now, consider the combination i-I, with results in Figure (7). Here, the trait mapping takes the form

$$f_H(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_H(1 - \Omega_H) \tag{29a}$$
$$f_S(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_S(1 - \Omega_S) \tag{29b}$$
$$f_C(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_C\Omega_H\Omega_S\alpha_H\alpha_S \tag{29c}$$
$$d(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = d_0(1 - \alpha_S) \tag{29d}$$
$$a(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = a \in \mathbb{R}_{\geq 0} \tag{29e}$$

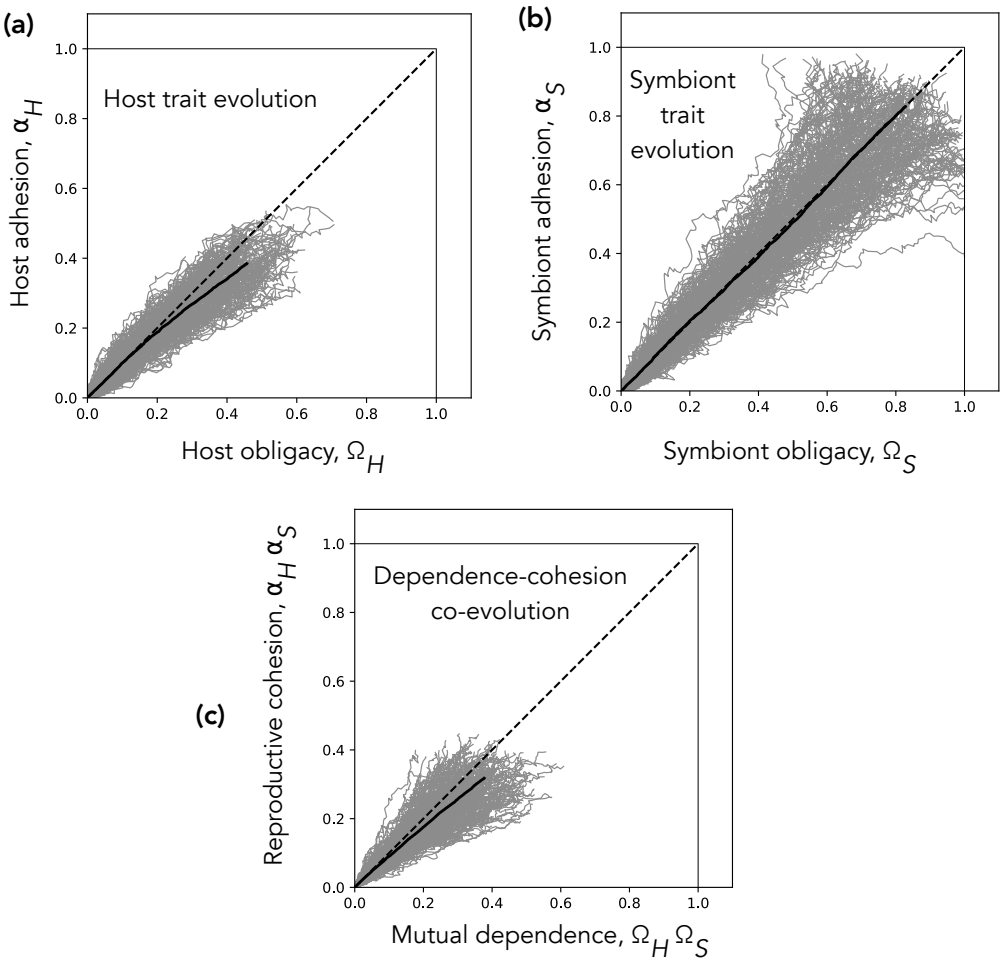

Figure 7: Independent growth rate depends only on own obligacy, dissociation rate depends only on one adhesion.

Here there is a slight bias towards mutual dependence which can be understood by looking at the evolution of the host and symbiont in their own trait space. There is a benefit for both host and symbiont to increase obligacy, and there is the *same* kind of benefit for the symbiont to increase its adhesion since $d = d_0(1 - \alpha_S)$. There is a smaller benefit for the host to increase its adhesion because an increased host adhesion only increases f_C but doesn't decrease d. Therefore, Ω_H evolves faster than α_H; however, Ω_S and α_S evolve at the same rate. If instead $d = d_0(1 - \alpha_H)$, there is no bias between Ω_H and α_H whereas Ω_S evolves faster than α_S. This leads to a slight bias towards mutual dependence. This bias is a little artificial in the following sense: the bias arises only because we are forced to make a choice between α_H and α_S in the form of d, and would not exist if this choice was not thrust upon us.

Second, consider the combination ii-I with results in Figure (8) and the following trait mapping:

$$f_H(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_H(1 - \Omega_H\Omega_S) \tag{30a}$$
$$f_S(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_S(1 - \Omega_H\Omega_S) \tag{30b}$$
$$f_C(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_C\Omega_H\Omega_S\alpha_H\alpha_S \tag{30c}$$
$$d(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = d_0(1 - \alpha_S) \tag{30d}$$
$$a(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = a \in \mathbb{R}_{\geq 0} \tag{30e}$$

Here, there is a bias towards reproductive cohesion. This is because the benefit of increasing the adhesion is now stronger than that of increasing obligacy. This is further, to be more precise, because the functional effect of increasing adhesion does not depend also on the other adhesion. There is also a slight bias towards dependence in the host because increasing adhesion for the host does not increase d, but again this is because we are forced to make the same choice between α_H and α_S for d. If instead $d = d_0(1 - \alpha_H)$, there would be a slight bias towards dependence in the symbiont and a bias towards cohesion in the host.

Recall that the default combination that we have worked through in the main text is given, in this shorthand, by i-II so we will not consider it again. Now, lastly, let us study the combination ii-II with results in Figure (9) and trait maps:

$$f_H(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_H(1 - \Omega_H\Omega_S) \tag{31a}$$
$$f_S(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_S(1 - \Omega_H\Omega_S) \tag{31b}$$
$$f_C(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_C\Omega_H\Omega_S\alpha_H\alpha_S \tag{31c}$$
$$d(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = d_0(1 - \alpha_H\alpha_S) \tag{31d}$$
$$a(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = a \in \mathbb{R}_{\geq 0} \tag{31e}$$

There is, as expected, no bias between mutual dependence and reproductive cohesion. Table 2 summarises the above results.

Now, we shall show that the bias towards mutual dependence is robust to choices that we made throughout the main text in arriving to this result. For ease of reference, we present

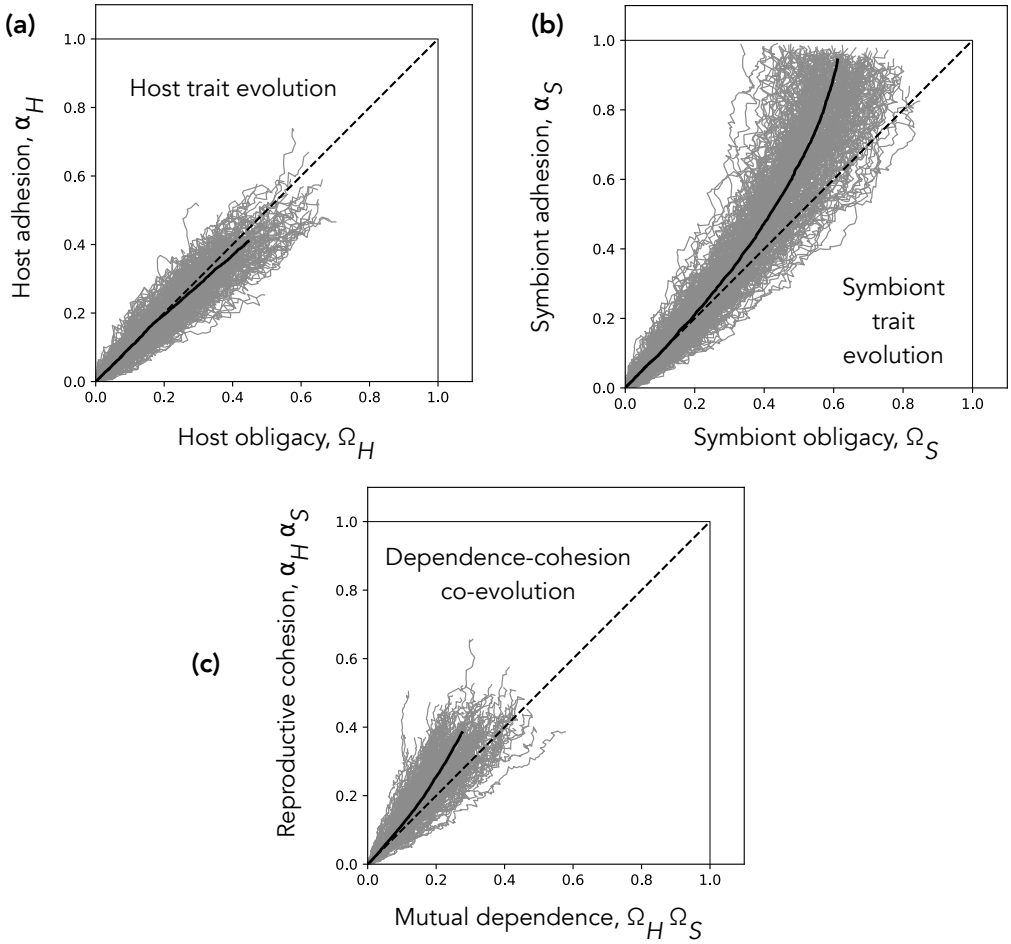

Figure 8: Independent growth rate depends on both obligacies, dissociation rate depends on only one adhesion.

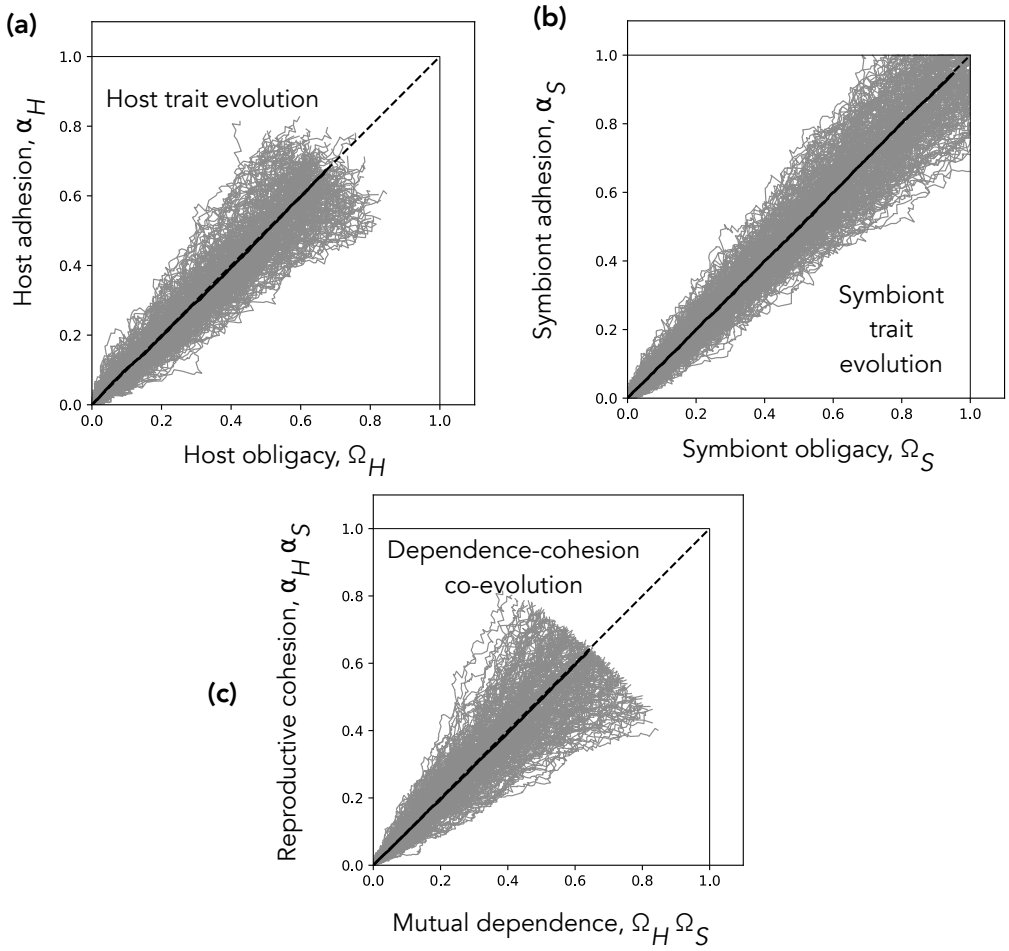

Figure 9: Independent growth rate depends on both obligacies, dissociation rate depends on both adhesions.

again the alternate trait map relevant for panel **(a)** of Figure 10

$$f_H(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_H(1 - \Omega_H) \tag{32a}$$
$$f_S(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_S(1 - \Omega_S) \tag{32b}$$
$$f_C(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = r_C \left(\frac{\Omega_H + \Omega_S}{2}\right)\left(\frac{\alpha_H + \alpha_S}{2}\right) \tag{32c}$$
$$d(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = d_0 \left(1 - \frac{\alpha_H + \alpha_S}{2}\right) \tag{32d}$$
$$a(\Omega_H, \alpha_H, \Omega_S, \alpha_S) = a \ \in \ \mathbb{R}_{\geq 0} \tag{32e}$$

B Computation of the average trajectory in stochastic simulations with runs of different lengths

In several figures of the main text, we present plots with several independent simulation runs in grey, with a mean trajectory in solid black for interpretational convenience. Here we briefly detail the computation of this average trajectory.

This section is necessitated by the fact that different evolutionary trajectories take different amounts of time – in particular, mutational steps – to reach a given final state. This implies that a mean trajectory cannot be calculated in the usual way - the mean trajectory at time t is given by the mean of the individual trajectories at time t. Take, for example, the case of the feasibility bound in the exponential model of population dynamics. All trajectories "collapse" when they reach the feasibility bound, but they take different amounts to do so.

The mean trajectory is calculated in the following way. Suppose there are N_T total trajectories (grey in the figures). Let L_i be the length of the ith trajectory, and suppose L_{max} = $max_i L_i$ is the duration (hereafter used interchangeably with length) of the longest trajectory i.e. the trajectory that takes the most time. Denote trajectory i by a function $f_i(t)$ where $t \in \{0, ..., L_i\}$ (for example, from zero to the time it takes to reach the feasibility bound). We first artificially extend all trajectories such that they are all of length L_{max}. In particular, we redefine for all i

$$f_i(t) = \begin{cases} f_i(t) & 0 \leq t \leq L_i \\ f_i(L_i) & L_i < t \leq L_{max} \end{cases}$$

Recall that t cannot be larger than L_{max} for any trajectory by definition.

All trajectories are now of the same length, and the mean trajectory $\bar{f}(t)$ (in solid black) is calculated by

$$\bar{f}(t) = \frac{1}{N_T} \sum_{i=1}^{N_T} f_i(t)$$

See the associated GitHub repository for the scripts that this is implemented in.

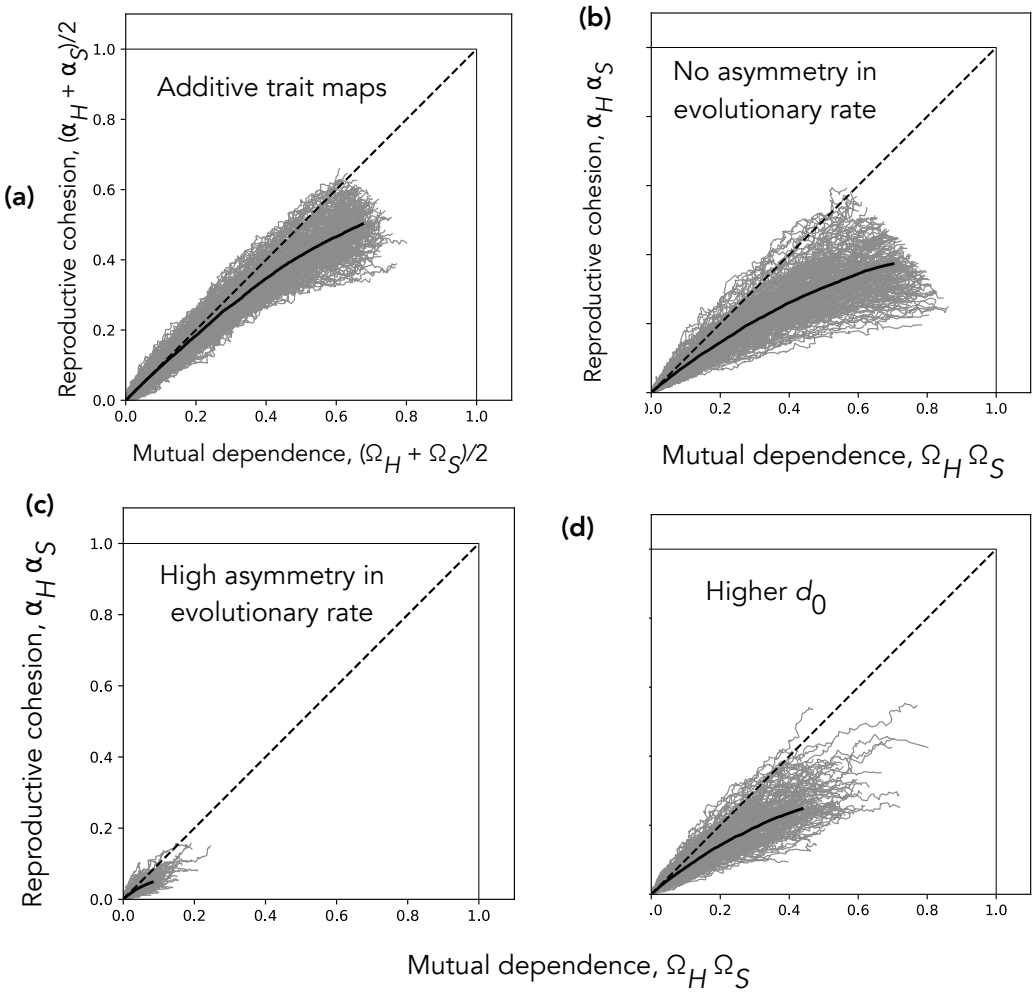

Figure 10: **The bias towards mutual dependence is remarkably robust to model choices.** We change important model choices to test the robustness of this result. None of the changes we have made affect its validity. **Base parameter values**. $f_H = 8, f_S = 20, r_C = 10, K_H = 100, K_S = 200, a = 0.1, d_0 = 50$. Changes are made to these values in the text that follows only when explicitly mentioned. **(a)** First, we change the map taking traits to ecological parameters into one where the traits have additive effects: the growth rate of the collective is proportion to the arithmetic mean of Ω_H and Ω_S, etc. The full map is presented in Equation (32). **(b)** Here, we set $r_H = r_S$ and $C_H = C_S$ while keeping all other parameters fixed. **(c)** Now we set $r_S = 10r_H$ and $C_S = 10C_H$ to understand if a much higher asymmetry in the growth rates and carrying capacities affects evolutionary trajectories. **(c)** Lastly we set $d_0 = 150$ to understand if a much higher initial dissociation rate does something. It does not, really.

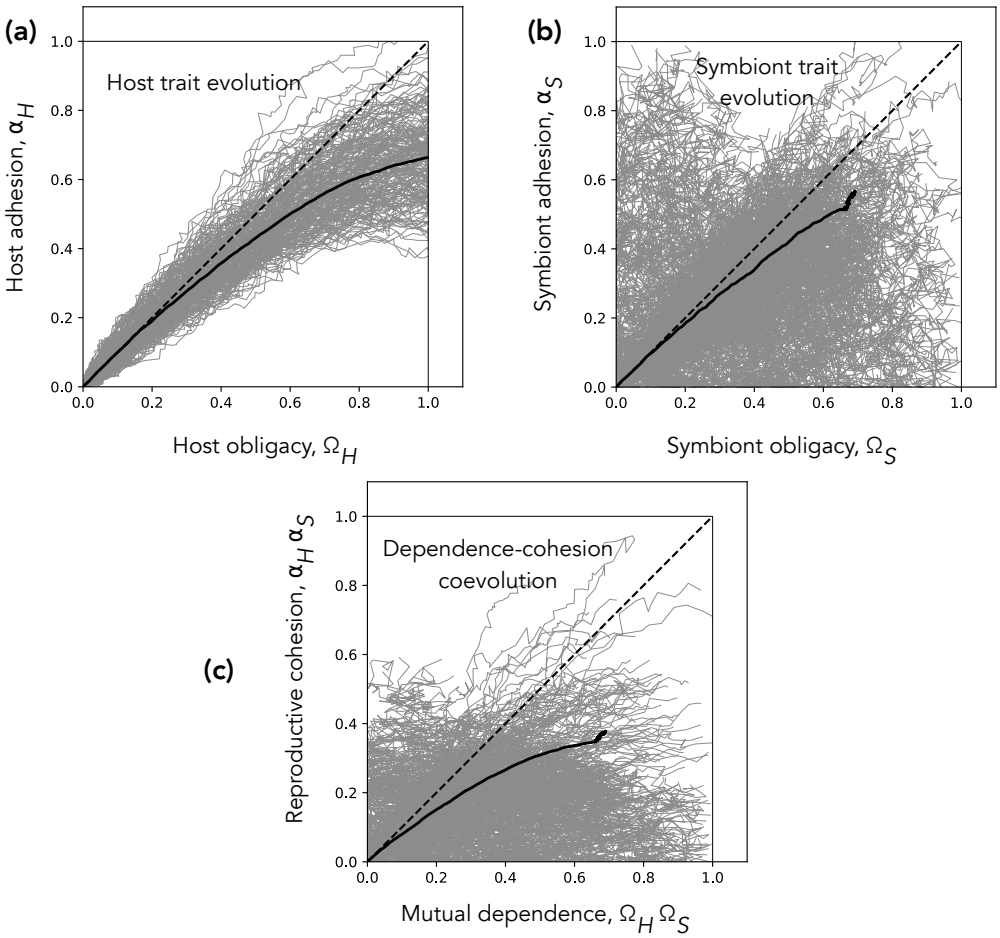

Figure 11: **Dependence-cohesion trajectories in the logistic model - the wild, wild west.** One can repeat the full exercise of finding co-evolutionary trajectories in the dependence-cohesion plane for the logistic model. This gives the above preliminary results, which are consistent - in the initial stages of evolution, at least - with the picture drawn by the exponential model. The author presents them here only to place them on record, but more analysis is required for a deeper understanding of this data.

Host-symbiont	Degree of dependence	Mode of transmission	References
Ancestral eukaryote host-protomitochondra, protoplastid	Mutually obligate	Vertical	Sagan (1967)
Aphid-*Buchnera aphidicola*	Mutually obligate	Vertical	Buchner (1965); Gündüz and Douglas (2008); Vogel and Moran (2010); Hansen and Moran (2011)
Riftia-Endoriftia	Host obligate, symbiont facultative	Horizontal	Nussbaumer et al. (2006); Felbeck and Jarchow (1998)
Dictyostelium discoideum-Burkholderia	Both facultative	Mixed	DiSalvo et al. (2015)
Squid-*Vibrio* (not intracellular)	Both facultative	Horizontal	McFall-Ngai (2014)
Pea aphid-*Regiella*	Both facultative	Vertical, horizontal possible	Scarborough et al. (2005); Gehrer and Vorburger (2012)
Legume-rhizobium (not intracellular)	Both facultative	Horizontal	Kiers et al. (2003); Denison and Kiers (2011)
Flatworm-*Riegeria*	Host obligate, symbiont dependence unclear	Vertical	Gruber-Vodicka et al. (2011)
Rhopalodia-nitrogen fixing *Cyanobacteria*	Mutually obligate	Vertical	Prechtl et al. (2004)
Cicada-*C. sulcinia/hodgkinia*	Mutually obligate	Vertical	McCutcheon et al. (2009)

Table 1: This table is a survey of symbioses for which there is nontrivial information of the degree of dependence and the mode of symbiont transmission (a proxy for reproductive cohesion). The information present, however, is insufficient since it is only (i) qualitative, (ii) a snapshot of the current state. It is impossible to exactly calculate the traits Ω_i and α_i with only qualitative information, and hence make the comparison necessary to confirm/refute the bias towards mutual dependence that this work predicts. Further, a trivial observation is that information about current state does not carry information about the full evolutionary trajectory. These facts necessitate an evolution experiment on a endosymbiosis. We note that our model considers only asexual reproduction and is phenomenological; it therefore applies exactly only to bacterial symbioses, and might not map in all cases perfectly onto the experimental system.

	i	ii
I	Slight bias	Bias towards reproductive cohesion
II	Default, bias mutual dependence	No bias

Table 2: A schematic description of trajectories in the dependence-cohesion under different kinds of mappings, with the shorthand i-ii-I-II described in Section A.9. There is a pattern in this table: when both trait pairs Ω_i and α_i combine in the same way to influence ecological parameters (i-I or ii-II), there is little to no bias. The slight bias in i-I case is an artefact, see main text for an explanation. On the other hand, when they combine in different ways (i-II or ii-I), there is a bias - towards faster evolution of dependence in one case and faster evolution of cohesion in the other. Our work arrived at case i-II from biological considerations, and we therefore make the associated prediction.