

The evolution of dependence and cohesion in incipient endosymbioses

Manuscript elements: Figure 1, Figure 2, Figure 3, Figure 4, Figure 5, Table 1, Table 2, Table 3, appendices S.1 upto S.6 (including Figure S.1 to Figure S.14).

Keywords: major evolutionary transitions, symbiosis, mitochondria, microbiome, adaptive dynamics, Red King effect, obligacy, adhesion.

Manuscript type: e-article.

Abstract

Eukaryogenesis is the prototypical example of an egalitarian evolutionary transition in individuality, and endosymbiosis, more generally, is central to the origins of many complex biological systems. Why do only some symbioses undergo such a transition, and how does the host-symbiont relationship change during this process? Here, we characterise endosymbiosis by two emergent collective-level properties: host and symbiont survival as a collective (“mutual dependence”) and the level of synchronised reproduction (“reproductive cohesion”). Using adaptive dynamics, we study the evolution of the traits underlying these properties. First, by adding a carrying capacity for the collective population – a realism omitted in previous models – we find novel reasons why complete dependence or cohesion might not evolve, thus providing further theoretical support for the rarity of transitions in individuality. Second, our model suggests that asymmetries in evolutionary outcomes of hosts and symbionts can be explained by a difference in their population growth parameters, coupled with their shared fate when in a collective. Lastly, we show that during the early stages of an endosymbiosis, even if investments in dependence and cohesion are uncorrelated, mutual dependence arises faster than reproductive cohesion. Our results hence shed light on three aspects of endosymbiosis: coevolution between the host and symbiont, coevolution between dependence and cohesion, and ultimately on the opportunity of undergoing an evolutionary transition. Connecting to ecological factors, this work uncovers fundamental properties of endosymbioses, providing a clear way forward for theoretical and empirical investigations.

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 transition in individuality: initially, autonomous and unrelated entities – the host and symbiont – come together to give rise to a more complex, integrated entity (Queller, 2000; Szathmáry, 2015; Szathmáry and Smith, 1995; West et al., 2015). The origin of endosymbiosis is not only a marked evolutionary transition but also a characteristic energy transition (Lane and Martin, 2010). Arguably, access to more energy processing via the endosymbionts set the stage for further evolutionary leaps such as multicellularity and beyond (Lane, 2015).

Endosymbionts are present in most life forms, even in unicellular prokaryotes (Corsaro et al., 1999; Wujek, 1979). Mitochondria and plastids have been famously shown to be endosymbionts arising from an ancient union between Archaeabacteria and an alphaproteobacterium (Fan et al., 2020; Koonin and Yutin, 2014; Martijn et al., 2018; Sagan, 1967). Many insects, such as the sap-sucking aphids, have been co-diversifying with their *Buchnera*, *Wigglesworthia*, and *Wolbachia* endosymbionts for millions of years (Hansen and Moran, 2011; Shigenobu et al., 2000; Wu et al., 2004; Zientz et al., 2004). There are examples from diverse ecologies across the globe - methanogenic endosymbionts in anaerobic ciliates (Embley and Finlay, 1994), nitrogen-fixing endosymbionts in the diatom *Rhopalodia* (Precht et al., 2004), consortia of chemosynthetic bacteria in gutless tubeworms (Woyke et al., 2006), and cyanobacterial endosymbionts in sponges (Thacker, 2005). New kinds of endosymbiotic associations continue being discovered, such as denitrifying endosymbionts in anaerobic ciliates (Graf et al., 2021). Despite this widespread prevalence, endosymbiosis and egalitarian evolutionary transitions are understudied relative to their fraternal counterparts, such as the evolution of multicellularity or eusociality (for which there exist many review or book-length treatments, e.g. Herron et al. (2022); Hölldobler and Wilson (2008)). However, this point of view is slowly changing with the recent emergence of more

studies on symbiosis (Estrela et al., 2016; Law and Dieckmann, 1998; Nguyen and van Baalen, 2020; Patel and West, 2022; Zachar and Szathmary, 2017; Zachar et al., 2018).

In this work, we are interested in the relationship between the host and symbiont and how it evolves throughout an evolutionary transition in individuality. By a transition in individuality, we mean the emergence – from individuals of one or more species that can undergo evolution – of a higher-level entity that can itself undergo Darwinian evolution. Transitions where the lower-level individuals making up a collective are unrelated are called “egalitarian” (Maynard Smith and Szathmary, 1995). Some symbioses have undergone such a transition, whereas others have not. For example, proto-mitochondria and their ancestral hosts underwent a transition to form the modern eukaryotic cell (Sagan, 1967). *Buchnera* endosymbionts are obligately dependent on their aphid hosts, vertically transmitted, and have small genomes (Bennett and Moran, 2015). On the other hand, while the tubeworm *Riftia* is obligately dependent on its symbiont *Endoriftia*, the latter has a free-living stage, and is transmitted horizontally (Nussbaumer et al., 2006). Similarly, the bobtail squid *Euprymna 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 host’s and symbiont’s life-history traits? How do the properties of the collective co-evolve, and is there a difference between host and symbiont evolutionary trajectories? Such questions were introduced in the works of Law and Dieckmann (1998) and Nguyen and van Baalen (2020). These models emphasized different aspects of symbioses: Law and Dieckmann (1998) considered the evolution of an exploitative symbiont and showed that even unidirectional resource transfer can lead to an evolutionary transition. Nguyen and van Baalen (2020) considered a tradeoff between independent reproduction and host encounter rate, and showed that this can lead to evolutionary equilibria where the symbiont is facultatively dependent on its host. However, certain assumptions were made for analytical tractability, such as exponential population growth without any regulation for hosts, symbionts, or collectives (Law and Dieckmann, 1998) and ignoring the evolution of host traits, and letting collectives grow exponentially (Nguyen and van Baalen, 2020). In this study, we,

too, make a new simplifying assumption (see the Model section, parameter d), but it allows us to relax many others and, in doing so, ask a variety of different questions about host-symbiont co-evolution with reasonable population dynamics.

To begin, we define a host-symbiont collective to be an obligate endosymbiosis if it exhibits three properties: (i) intracellular location of the symbiont; (ii) at least one of the host or symbiont is obligately dependent on this interaction; and (iii) the collective can reproduce synchronously, i.e. 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, which is the target of our notion of “reproductive cohesion”, is similar to, but stronger than, vertical symbiont transmission: we also include the requirement that the replication of the two partners is coupled, such as the coordination of mitochondrial fission and segregation with the cell cycle (although this is more complicated, see Mishra and Chan (2014)). Importantly, synchronised 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, our definition excludes gut microbiomes as there is no intracellular location or synchronised reproduction.

To study the conditions for the evolution of obligate endosymbiosis, we use evolutionary invasion analysis from the adaptive dynamics framework (Dieckmann and Law, 1996; Geritz et al., 1998; Metz et al., 1992; Otto and Day, 2007). This is a framework to study the long-term evolution by natural selection of traits that affect their bearer’s ecological interactions. The evolutionary fate of new rare mutants is studied by determining if they can invade and fix in the population in which they arise. This analysis assumes the separation of ecological and evolutionary timescales: the ecological processes dictating the fate of a mutant (“natural selection”) take place much faster than the timescale on which new host/symbiont mutants arise.

An eco-evolutionary model

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

Inspired by [Estrela et al. \(2016\)](#), we characterise a symbiosis by two emergent, collective-level quantities: the degree of host-symbiont mutual dependence and the degree of their reproductive cohesion (see Fig. 1). Mutual dependence is an aggregative measure of how well the two partners grow when free-living instead of when they are part of a collective; reproductive cohesion measures how often they reproduce synchronously instead of individually. The evolution of symbioses can thus be visualised as taking place in the plane of these two collective-level quantities. We say that an evolutionary transition in individuality occurs when there is complete mutual dependence and reproductive cohesion - the constituent individuals cannot live or reproduce without the other. We aim to understand the evolutionary trajectories in this dependence-cohesion plane. To study the joint evolution of the traits underlying dependence and cohesion, we use the method of evolutionary invasion analysis. The main object of interest is the invasion fitness of a mutant in the environment generated by a resident population. It is assumed in this framework 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](#)).

We formalise the ecological, i.e. population dynamics over short, mutation-free timescales as follows: Let f_H, f_S, f_C be the growth rates of the host, symbiont, and collective populations, respectively. Further, suppose the host and symbiont associate with and dissociate from each other at rates a and d , flowing into and out of the collective population. See Fig. 1(c) for a

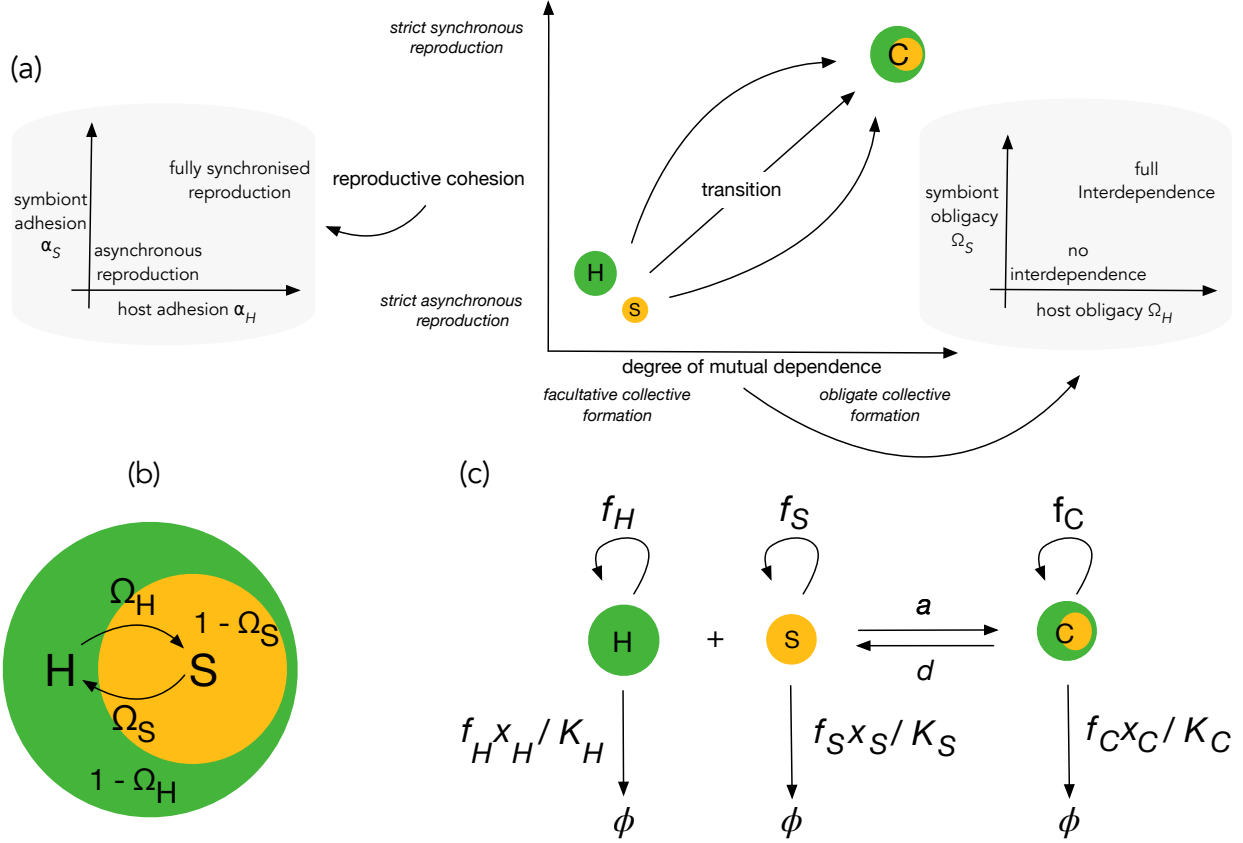


Figure 1: **(a)** We conceptualise the trajectory of an evolutionary transition in the plane of two quantities - the reproductive cohesion of the lower-level individuals and the degree of mutual dependence between them. The adhesions α_H and α_S of the host and symbiont respectively control their investment in reproductive cohesion, i.e., the degree of synchronised reproduction. The obligacy Ω_i describes the investment of type i in growth as a collective. In this study, we study the co-evolutionary trajectory of the adhesions with the obligacy and between the host and symbiont. **(b)** On a microscopic, physiological scale, it is helpful 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 and is most concretely connected to the magnitude of synchronised vs. asynchronised reproduction of species i . **(c)** The flows of the population dynamical model. All populations have a logistic growth rate corresponding e.g. to an intrinsic birth rate f_i and a density-dependent death rate $f_i x_i / K_i$. The host and symbiont associate and dissociate with rates a and d respectively.

graphical representation of all possible events. The dynamics are given by:

$$\begin{aligned}\dot{x}_H &= f_H x_H \left(1 - \frac{x_H}{K_H}\right) - a x_H x_S + d x_C, \\ \dot{x}_S &= f_S x_S \left(1 - \frac{x_S}{K_S}\right) - a x_H x_S + d x_C, \\ \dot{x}_C &= f_C x_C \left(1 - \frac{x_C}{K_C}\right) + a x_H x_S - d x_C\end{aligned}\tag{1}$$

where K_i is the carrying capacity of population i . These dynamics are our default model, which we study extensively. To achieve some analytical results, we will also study this model with an infinite collective carrying capacity ($K_C \rightarrow \infty$). We also explore numerically different extensions, which we specify in the respective results sections.

The parameters of the population dynamical model above depend on the traits underlying dependence and cohesion. Consider first mutual dependence: we introduce the traits Ω_H and Ω_S , referred to hereafter as the “obligacy” of the host and symbiont, respectively (Fig. 1a). These are dimensionless numbers in $[0, 1]$ and denote the degree of dependence of the host and symbiont on the formation of the collectives. These traits formalise a tradeoff between individual (f_H, f_S) and collective reproduction rate f_C . 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, e.g. that the growth rates of the host population in isolation do not depend on the symbiont’s investment (and vice versa).

With regards to reproductive cohesion, we introduce the traits α_H and α_S , hereafter referred to as the “adhesion” of the host and symbiont, respectively. These are also dimensionless in $[0, 1]$, and a higher adhesion denotes a higher association rate a , lower dissociation d , and a higher propensity of synchronised birth f_C . These traits induce a tradeoff between processes favouring the formation of the collective and those favouring the individuals. The mathematical translation of these statements is in terms of the partial derivatives of f_i, a, d along the traits and is stated precisely in section §S.1.1 of the Supplementary Information (hereafter “SI”). The mapping of

151 traits to population dynamical parameters that we will use for most of this work is

$$\begin{aligned}
f_H(\Omega_H, \alpha_H, \Omega_S, \alpha_S) &= r_H(1 - \Omega_H), \\
f_S(\Omega_H, \alpha_H, \Omega_S, \alpha_S) &= r_S(1 - \Omega_S), \\
f_C(\Omega_H, \alpha_H, \Omega_S, \alpha_S) &= r_C \Omega_H \Omega_S \alpha_H \alpha_S, \\
d(\Omega_H, \alpha_H, \Omega_S, \alpha_S) &= d_0(1 - \alpha_H \alpha_S), \\
a(\Omega_H, \alpha_H, \Omega_S, \alpha_S) &= a_0 \in \mathbb{R}_{\geq 0}.
\end{aligned} \tag{2}$$

152 The parameters in our model are summarised in Table (1). In the Results section, we present
153 analyses of a set of interconnected versions of the model set up thus far. We first study the
154 evolution of dependence only – letting obligacies Ω_i evolve while the adhesions α_i stay constant
155 – and then vice versa i.e. cohesion only. We then introduce a simplified version of the population
156 dynamical eqs. (1) where the collective can grow exponentially; here we gain some analytical
157 insights and study the joint evolution of obligacies and adhesions.

158 We assume a growth tradeoff: that any investment in independent reproduction from either
159 host or symbiont comes at a cost to their contribution to the collective growth rate (and vice
160 versa). Notice that in Eqs. (2), host independent growth rate f_H is maximum at $\Omega_H = 0$,
161 decreases with increasing Ω_H , and is zero at $\Omega_H = 1$ (and identical for the symbiont). This
162 tradeoff is inspired by the consideration that the selective pressures while in a collective and those
163 while free-living are sufficiently contrasting that mutations have antagonistic effects in these
164 two niches. For example, endosymbiont adaptation can be influenced by external pH levels or
165 ambient amino acid availability, which are plausibly vastly different inside and outside the host.
166 Similar statements about nutritional availability or pathogen defense can also be made for the
167 pressures on host adaptation. Such tradeoffs have been observed in the squid-*Vibrio* symbiosis
168 (Cohen et al., 2020; Pankey et al., 2017), and in some experiments with *Bradyrhizobium* symbionts
169 of plants (Sachs et al., 2011) (see Brockhurst et al. (2024) for a more extensive discussion of
170 growth tradeoffs). We are in this study interested primarily in mutualistic endosymbiosis, where
171 increased investment from either partner increases collective growth rate. The tradeoff above

is a case where investment into collective growth is a priori difficult to emerge because of the linked cost to free-living growth. It is possible that only the symbiont experiences such a tradeoff or that there are mutations that only advantage the collective or disadvantage the free-living host/symbiont; we do not consider these in our model.

To differentiate between the Ω and α trait pairs, it is useful to focus on an example such as a fig-wasp mutualism (reviewed in [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. Therefore, the two species are highly dependent on each other, but they do not physically reproduce as a unit. Hence, here the obligacies Ω_i are high, but the adhesions α_i are not.

Evolutionary invasion analysis. To compute an arbitrary mutant’s invasion fitness, we suppose a resident population is at ecological equilibrium, i.e., a stable steady state of the population dynamics (in our case Eq. (1)). The abundances at this steady state are computed numerically, unless mentioned otherwise. It is assumed that a mutant then arises, with trait value drawn from a symmetric probability distribution centred at the resident trait. The fate of this mutant is decided via its invasion fitness, i.e. the growth rate of a small number of mutants in the resident population ([Metz et al., 1992](#)). The mutant can invade if its invasion fitness is positive; if it is negative, the mutant goes to extinction ([Dercole and Rinaldi, 2008](#); [Dieckmann and Law, 1996](#)). This quantity is traditionally defined as the largest eigenvalue of the Jacobian of Eq. (1) when augmented for the presence of a mutant type. However, in our case, this quantity is not amenable to mathematical analysis, so we use the next-generation theorem, which gives an alternate characterisation of the same number ([Hurford et al., 2010](#); [van den Driessche and Watmough, 2002](#)). We then use the canonical equation of adaptive dynamics ([Dieckmann and Law, 1996](#)) to study the macroscopic behaviour of long-term evolutionary trajectories (SI §S.4).

Parameters	Description	Possible values	Comments
Ω_H, Ω_S	Host and symbiont obligations	$[0,1]$	initially set to $(0,0)$
α_H, α_S	Host and symbiont adhesions	$[0,1]$	initially set to $(0,0)$
x_H, x_S, x_C	Population densities of host, symbiont, and collectives	$[0, \infty)$	equilibrium densities depend on trait values
f_H, f_S, f_C	Growth rates of H, S, and C populations	$[0, \infty)$	usually set such that $f_S > f_C > f_H$
K_H, K_S, K_C	Carrying capacities of H, S, and C populations	$[0, \infty)$	usually set such that $K_S > K_C > K_H$
a, d	Association and dissociation rates	$[0, \infty)$	initial $d > \text{initial } a$
d_H, d_S	Within-collective mortality rates of host and symbionts	$[0, \infty)$	assumed to not depend on any traits
b_H, b_S	Within-collective birth rates of host and symbionts	$[0, \infty)$	assumed to not depend on any traits
r_H, r_S, r_C	Maximum values of f_H, f_S, f_C	$[0, \infty)$	Only relevant when Ω_i and/or α_i are evolving
d_0	Maximum value of d	$[0, \infty)$	Only relevant when α_i are evolving

Table 1: A list of and associated information regarding all parameters named in our model.

Results and Discussion

In the first subsection, we investigate the case where obligacies evolve independently of the adhesions and vice versa. In the second, we analyse an extension of the model that includes within-collective birth and death rates. Then, in the third subsection, we study the relative importance of the collective population's growth rate and carrying capacity in deciding the course of evolution. Finally, in the last subsection, we study obligacy and adhesion coevolution in a model more amenable to analytics, which we derive by assuming an infinitely large collective carrying capacity in Eqs. (1).

It is challenging to analytically solve the system of Eqs. (1) to determine the fixed points and their stability. However, we determined computationally that for a wide range of parameter values, the dynamics converges to a stable fixed point over the evolutionary change in our traits of interest (see SI §S.2.1). Moreover, the values of the equilibrium population sizes increase or decrease gradually with the trait values, suggesting that there is a single internal fixed point. We will assume throughout that $K_S > K_H$ and $r_S > r_H$, i.e. the symbiont population has a higher carrying capacity and reaches it faster than the host. Consequently – on average over a given duration of time – more symbiont mutants arise than host mutants, and the symbiont trait hence has a higher rate of evolution (assuming, as we do, that the rate of mutation and the variance in the mutant trait distribution are the same for host and symbiont traits; see Eq. 4.12 in Dieckmann and Law (1996)).

Independent evolution of obligacies and adhesions

We begin by understanding the evolution of the obligacies when adhesions are kept constant, and vice versa. The mapping of obligacies to population dynamical parameters we use in Fig. 2,

panels (a,b,c) is:

$$\begin{aligned}
f_H(\Omega_H, \alpha_H, \Omega_S, \alpha_S) &= r_H(1 - \Omega_H), \\
f_S(\Omega_H, \alpha_H, \Omega_S, \alpha_S) &= r_S(1 - \Omega_S), \\
f_C(\Omega_H, \alpha_H, \Omega_S, \alpha_S) &= r_C\Omega_H\Omega_S,
\end{aligned} \tag{3}$$

where we assume that the adhesions α_i are not under selection. The constants r_H, r_S, r_C set the scale of the parameters f_i ; concretely, they can be understood as the intrinsic growth rates of the populations at either $\Omega_i = 0$ (for f_H and f_S), or when both obligacies are equal to 1 (for f_C). In panels (d,e,f) of Fig. 2, we assume obligacies to be constant and model the ecological consequences of a different adhesion via:

$$\begin{aligned}
f_C(\Omega_H, \alpha_H, \Omega_S, \alpha_S) &= r_C\alpha_H\alpha_S \\
d(\Omega_H, \alpha_H, \Omega_S, \alpha_S) &= d_0(1 - \alpha_H\alpha_S) \\
a(\Omega_H, \alpha_H, \Omega_S, \alpha_S) &= a_0 \in \mathbb{R}_{\geq 0}
\end{aligned} \tag{4}$$

The constants r_C, d_0, a_0 are again the maximum values of the respective functions, and set the scale of variation.

Results are presented in Fig. 2. First consider the evolutionary trajectory of the obligacies $(\Omega_H(t), \Omega_S(t))$: the symbiont obligacy Ω_S reaches 1 first (denote this time by $t_{\Omega_S}^*$), and after this the selective pressure on Ω_H disappears, leading to stagnation in its value (Fig. 2a,b). This lack of selection after $\Omega_S = 1$ occurs because when $f_S = 0$ (no independent symbiont reproduction), the host and collective populations equilibrate to their carrying capacities irrespective of Ω_H ; the symbiont population is sustained only through dissociation (see Eq. (1)). However, the symbiont has a higher equilibrium population size when $\Omega_S = 1$ than when $\Omega_S < 1$ due to the “backflow” from the dissociation of the collectives (see SI §S.2.1, or from Eq. (1)). The converse is valid for the host – its population size is not at the maximum possible over the trait space, while $\Omega_H < 1$. This is a counterintuitive result: the symbiont foregoes its ability to reproduce independently and has a higher equilibrium population size than its carrying capacity; the host now has a population

size exactly equal to its carrying capacity (when it could be higher, i.e. at $\Omega_H = 1$), but retains its ability to reproduce independently (Fig. 2c). Indeed, it is better (in terms of abundance) for both the host and symbiont to give up independent reproduction and be sustained only through dissociation, but the faster-evolving population can do this first. The above argument can also be made analytically via the computation of the basic reproductive number of a mutant $R_0(\tilde{\Omega}_H, \Omega_H, \Omega_S)$, which denotes – roughly – the number of offspring left by a mutant host with obligacy $\tilde{\Omega}_H$ that arises in a resident population of hosts with obligacy Ω_H and symbionts with obligacy Ω_S (derived using the next-generation theorem, see SI §S.4.1). This quantity is given by

$$R_0(\tilde{\Omega}_H, \Omega_H, \Omega_S) = \frac{\tilde{f}_H + \frac{dax_S^*}{(d-\tilde{f}_C)+\tilde{f}_C\frac{x_C^*}{K_C}}}{\tilde{f}_H\frac{x_H^*}{K_H} + ax_S^*} \quad (5)$$

where a star denotes that the value at equilibrium must be used. The R_0 is not maximally informative in our case since we cannot analytically solve for the equilibrium population abundances x_i^* . Nevertheless, it can be shown (see §S.4.1) that $R_0(\tilde{\Omega}_H, \Omega_H, 1) = 1$ for any combination $(\tilde{\Omega}_H, \Omega_H) \in (0, 1]^2$ i.e. there is no fitness difference between any Ω_H mutants that arise in the background of $\Omega_S = 1$. A consequence of this result that cannot be captured by adaptive dynamics is that this model predicts the accumulation of neutral genetic variation in the host population at loci coding for obligacy.

Now consider the evolutionary trajectory of adhesions, presented in Figure 2d,e. Symbiont adhesion α_S reaches 1 first while α_H is still smaller than 1, and then host adhesion increases until an evolutionarily stable strategy (ESS) is reached. Trajectories then drift along a curve near (1,1). This ridge connects the states $(\Omega_H < 1, \Omega_S = 1)$ and $(\Omega_H = 1, \Omega_S < 1)$, and it is possible that a small proportion of trajectories eventually reaches $\Omega_H = 1$. The drift in Fig. 2d,e,f is what our model predicts of biological populations, however its source in this figure is different: accepting false positives when deciding the fate of mutants based on $R_0 > 1$, which arise due to floating point errors that we cannot entirely remove. All neutral mutations would have $R_0 = 1$, whereas we accept values only larger than a value that is infinitesimally larger than one to exclude such errors (see associated scripts for exact details).

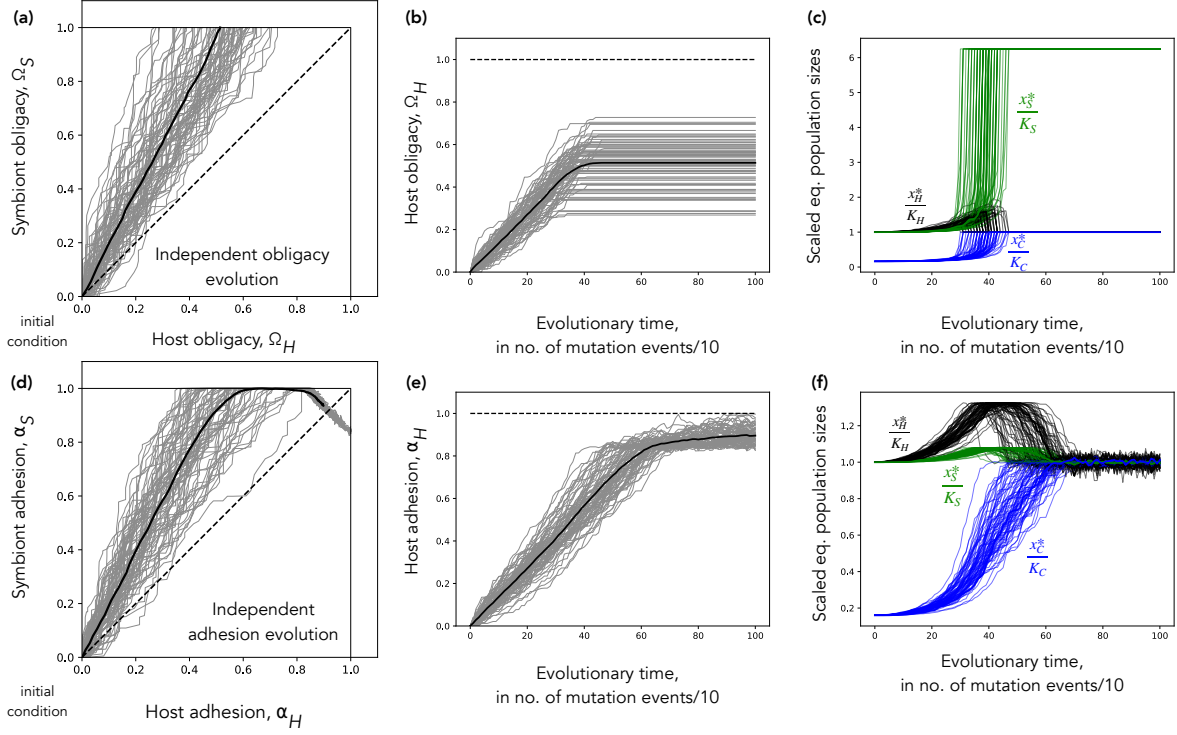


Figure 2: Evolutionary dynamics of host and symbiont obligacies. All panels contain results of 75 independent stochastic simulations. **(a)** Symbiont obligacy evolves to 1, and then host obligacy is no longer under selection (see main text for an explanation). **(b)** Symbiont adhesion evolves to 1, and then host adhesion increases until a certain value. The traits together then vary along a neutral ridge around (1,1) (see main text for an explanation). **(c,d)** These panels show the evolutionary trajectory of the host trait, since it is here that the nontrivial outcomes take place. Ω_H plateaus at different values based on when the value of Ω_S reaches one. α_H increases up to a high value and stays close to it. **Parameter values.** Common for all: $K_H = 100, K_S = 200, K_C = 250, a_0 = 0.1$; for **(a)**: $r_H = 8, r_S = 20, r_C = 10, d = 50.0$; for **(b)** and **(c)**: $f_H = 8, f_S = 20, r_C = 10, d = 50$.

Nevertheless, there are two facts to explain: that the traits increase and that there is a host-symbiont asymmetry in outcomes. Our observations can be explained by the variation in equilibrium population sizes as traits change (see SI §S.2.1 for heatmaps). Notably, (i) they increase along the adhesions α_i (and hence the traits themselves also increase); and (ii) they do so non-monotonically, with a local maximum in $x_i^*(\alpha_H, \alpha_S)$ occurring over a ridge in the interior of the $\alpha_H - \alpha_S$ space that surrounds the top right corner (1,1) (see Fig. §S.2.1). The first consequence is that any trajectory (α_H, α_S) must pass over this ridge to reach (1,1). The symbiont adhesion can reach 1 before encountering this bump because of its faster rate of change, and the trajectory has to overcome it then while α_H is still less than 1. This is the source of the asymmetry: If both species have the same r and K , then the landscape of x_i^* is the same, but now the trajectory (α_H, α_S) hits the ridge when both adhesions are approximately equal and less than 1 (see SI §S.3.1). We suggest that the inflection point of this local maximum corresponds to the ESS, since decreasing population size (even locally) is disadvantageous, and such a mutant could not invade.

More proximately, an increased adhesion leads to modifications in two forces acting on host/symbiont population size via the term dx_C : (1) a decrease because fewer dissociations occur, but also (2) an increase because an increased adhesion increases f_C , and hence the collective population size. Since we begin with the intuitive scenario of high dissociation rate and zero synchronised growth, an initial increase in adhesion is beneficial, but this benefit is only up to a point, after which the decrease of d overpowers the increase due to f_C . This switch is the explanation for the presence of the intermediate ESS and the ridges in population size curves. This demonstrates that it is essential to consider not only intuitive parameters like f_H or f_S but also population size dynamics which can strongly influence invasion fitness in this context.

The effect of within-collective mortality and reproduction

In our model, we have assumed so far that dissociation is perfect: all events producing free-living individuals from a collective give rise exactly to one host and one symbiont. However, it

is likely that there is differential mortality for hosts and symbionts while they are in a collective or during dissociation. This would give rise to free-living hosts when the symbiont dies (at some rate d_S), and free-living symbionts with analogous rate d_H . Further, hosts and symbionts, while in collective, can also give birth to free-living hosts or symbionts. Suppose this takes place at rates b_H and b_S , respectively. The population dynamics is then described by

$$\begin{aligned}\dot{x}_H &= f_H x_H \left(1 - \frac{x_H}{K_H}\right) - a x_H x_S + d x_C + (d_S + b_H) x_C, \\ \dot{x}_S &= f_S x_S \left(1 - \frac{x_S}{K_S}\right) - a x_H x_S + d x_C + (d_H + b_S) x_C, \\ \dot{x}_C &= f_C x_C \left(1 - \frac{x_C}{K_C}\right) + a x_H x_S - d x_C - (d_S + d_H) x_C.\end{aligned}\tag{6}$$

The previous section then corresponds to e.g. setting the rates $b_i = 0$ and $d_i = 0$ $i \in \{H, S\}$. The correspondence is slightly more general; one obtains Eqs. (1) whenever $b_i = d_j$ for all combinations of $i, j \in \{H, S\}$; with the caveat that the corresponding value of the dissociation rate would be higher, i.e. $d + b_i + d_j$.

We consider two scenarios: (i) differential mortality while in collective, but no within-collective reproduction ($d_H = d_S \neq 0, b_i = 0$ $i \in \{H, S\}$; Fig. (3) right), and (ii) within-collective H (or S) can give birth to free-living H (or S) but within-collective mortality is absent ($b_H = b_S \neq 0, d_i = 0$ $i \in \{H, S\}$; Fig. (3) left). For the sake of brevity, we do not consider the cases where $b_H \neq b_S$ or $d_H \neq d_S$, or where the within-collective birth and death rates are both greater than zero. However, the latter is not modelling a completely new phenomenon since d_i and b_j appear together as an aggregate coefficient of the collective abundance x_C in Eq. (6). First the adhesion trajectories (Fig. 3 right): there are no qualitative changes compared to the case when $d_i = b_i = 0$, the prediction that the adhesions increase and then hit a ridge of evolutionarily stable strategies is robust to the addition of the above within-collective rates. At higher shared values of d_i , the adhesions do not increase from initial values (see SI §S.3.4).

The evolution of obligacies, however, displays departures both when $d_i > 0$ or $b_i > 0$. When hosts or symbionts can die while in collective but cannot give birth, the obligacies do not initially increase from (0,0) at all. However, if the obligacies are at a high enough initial value to begin

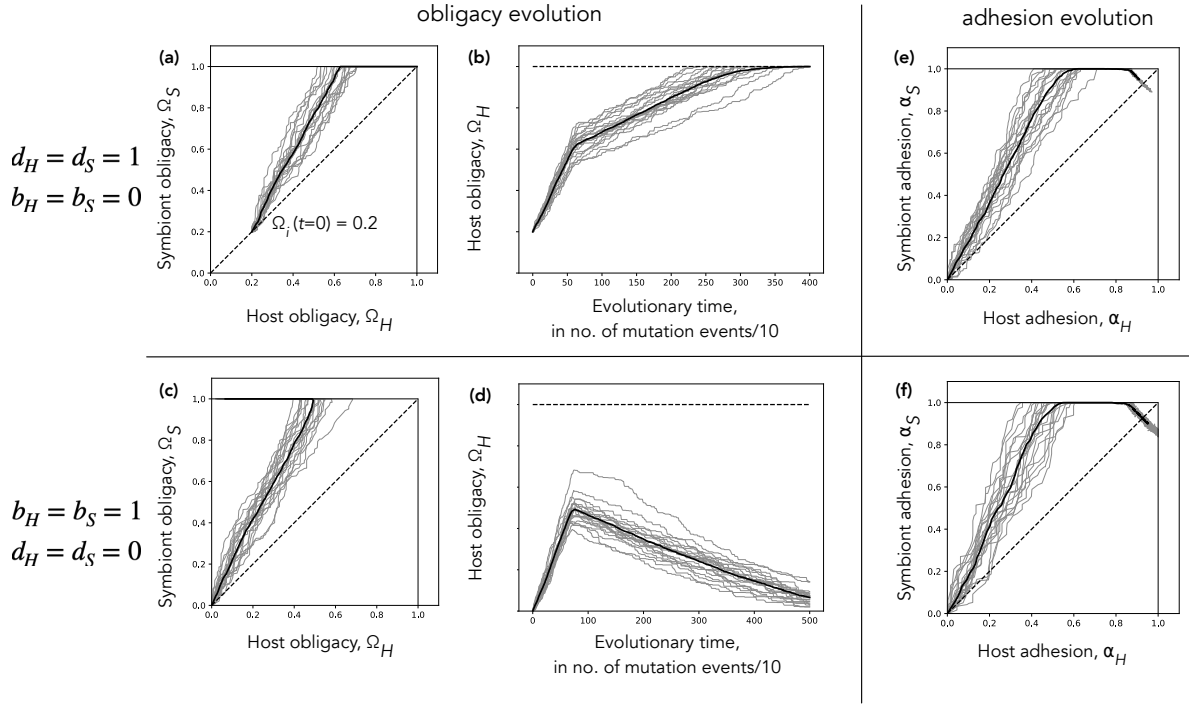


Figure 3: **Evolutionary dynamics of obligacies and adhesions when b_i and d_i are nonzero.**

All panels contain results of 20 independent stochastic simulations. **(a,b)** When $d_H = d_S > 0$, the initial obligacies must be above a threshold for them to further increase. Symbiont obligacy increases faster, and when Ω_S hits 1, host obligacy increases further upto 1. **(c,d)** When $b_H = b_S > 0$, dependence directly increases over time without a threshold minimum value that is necessary; when Ω_S hits 1, host obligacy now decreases to zero. **(e,f)** Qualitatively identical to adhesion evolution presented in Figure 2. **Parameter values.** Common for all: $K_H = 100, K_S = 200, K_C = 250, a_0 = 0.1$. Panels **(a,b,e)**: $d_H = d_S = 1, b_H = b_S = 0$. Panels **(c,d,f)**: $d_H = d_S = 0, b_H = b_S = 1$. Panels **(a,b,c,d)**: $r_H = 8, r_S = 20, r_C = 10, d = 50.0$. Panels **(e,f)**: $f_H = 8, f_S = 20, r_C = 10, d_0 = 50$.

with, the obligacies then increase as before with the symbiont obligacy Ω_S reaching 1 first. This threshold initial value increases with the shared value of d_H and d_S , but above this value the dynamics are qualitatively identical (see SI §S.3.4). There is another important difference: once $\Omega_S = 1$, the selective pressure on Ω_H does not disappear, instead now the host obligacy also increases to 1, leading to full mutual dependence. At the state $\Omega_H = \Omega_S = 1$, the equilibrium population abundances remain at steady, non-zero values in our case of $d_S = d_H$.

When hosts and symbionts can give birth while in collective but cannot die, the outcome is again different: the obligacies directly increase from (0,0) without needing a threshold dependence, but once symbiont obligacy reaches 1, the host is under selection to now decrease obligacy. This leads to a non-monotonic route to a one-sided mutualism, where the symbiont is completely dependent and the host not at all. In the ultimate state of $\Omega_H = 0, \Omega_S = 1$, there is no fixed point possible for Eqs. (6); the symbiont goes extinct, while the host and collective populations blow up due to the birth rates. At higher shared values of b_i , the obligacies still increase, but host-symbiont symmetry in obligacies is reduced (see SI §S.3.4).

To explain the results above, we again appeal to the population sizes at equilibrium. It can be shown that, if $d_S + b_H = d_H + b_S$ (which is true in Fig. 3 by the parameters we have set), once $\Omega_S = 1$ (implying $f_S = 0$), the equilibrium abundances calculated from Eqs. (6) are given by the equations

$$\begin{aligned} x_H^* &= K_H, \\ x_C^* &= \left(1 + \frac{(b_S - d_S)}{f_C}\right) K_C, \\ x_S^* &= \frac{(d + d_H + b_S)x_C^*}{ax_H^*}. \end{aligned} \tag{7}$$

Dependence on Ω_i here enters through the f_C in x_C^* : when $b_S > d_S$, an increased host obligacy decreases collective abundance at equilibrium x_C^* , and when $b_S < d_S$, an increased host obligacy *increases* x_C^* . More generally, the form of x_C^* reflects that the forces driving the collective population away from its carrying capacity are b_S -produced free-living symbionts which form collectives via host association, and collectives leaving the compartment at rate d_S to form hosts.

This suggests that even if the rates b_i and d_i are unequal, a case we have not tested here, the difference $b_S - d_S$ plays a central, if not solitary, role in deciding the outcome.

The inclusion of births and deaths of hosts and symbionts while in collective therefore affects two properties of the trajectory: initial increase of the obligacies, and what happens when $\Omega_S = 1$ is reached (see Table (2) for a summary of these effects). Including trait-dependence onto these within-collective rates would further complicate the picture since their values would become dynamic, and combinations of the above arguments might become necessary. We do not consider this possibility here, but some possible consequences are discussed in the Conclusions section. Having clarified the effects of the parameters b_i and d_i , we shall set them again to zero for the remaining sections to emphasize other predictions of this model.

Collective growth rate and collective carrying capacity have different effects on evolutionary outcomes

We have shown that dependence and cohesion can in fact increase under our assumptions, and that associating with a partner to form a collective is thus beneficial. However, this benefit is in terms of two independent (at least in our model) parameters - the collective population's growth rate r_C and its carrying capacity K_C . In this section, we are interested in which of these parameters is more important in determining the evolutionary outcome. We computed evolutionary trajectories for high and low values of the parameters r_C and K_C each - $r_C \in \{1, 40\}$ and $K_C \in \{10, 500\}$. At the low value of K_C , both obligacies and adhesions remained at their initial condition (0,0); at the high value of K_C they evolved away from (0,0). The outcome was not influenced by the value of r_C . This shows that it is more important that the collective's equilibrium population size is high, no matter how much the mutant and resident differ in how fast they reach this size. We propose a cutoff for the threshold value of K_C in the following paragraphs.

Nevertheless, once K_C is high enough, its precise value can influence the ultimate values of adhesion, but not obligacy. Fig. 4 presents the results of evaluating the evolutionary trajectories

Scenario	Parameters	Consequences on $\Omega_i(t)$	Consequences on $\alpha_i(t)$
Perfect collective dissociation, no within-collective births or deaths	$d_H = 0, d_S = 0$ $b_H = b_S = 0$	both increase, host obligacy stagnates once symbiont completely obligate	increase, then drift along a ridge
Within-collective mortality, but no births	$d_H = d_S > 0,$ $b_H = 0, b_S = 0$	both increase if initial obligacy high enough, ultimately both host and symbiont both become obligately dependent	increase, then drift along a ridge
Within-collective hosts can give birth to new hosts (similarly symbionts), but no within-collective deaths	$b_H = b_S > 0$ $d_H = 0, d_S = 0$	both increase without a threshold, ultimately symbiont becomes fully obligately whereas host completely independent	increase, then drift along a ridge
Within-collective birth and death rates are all equal	$b_H = b_S = d_H = d_S$	equivalent to setting $b_i = d_i = 0$, and increasing d , so there is no formal difference to the first row	

Table 2: A summary of the different evolutionary outcomes for Ω_i and α_i depending on the rates of within-collective mortality and reproduction.

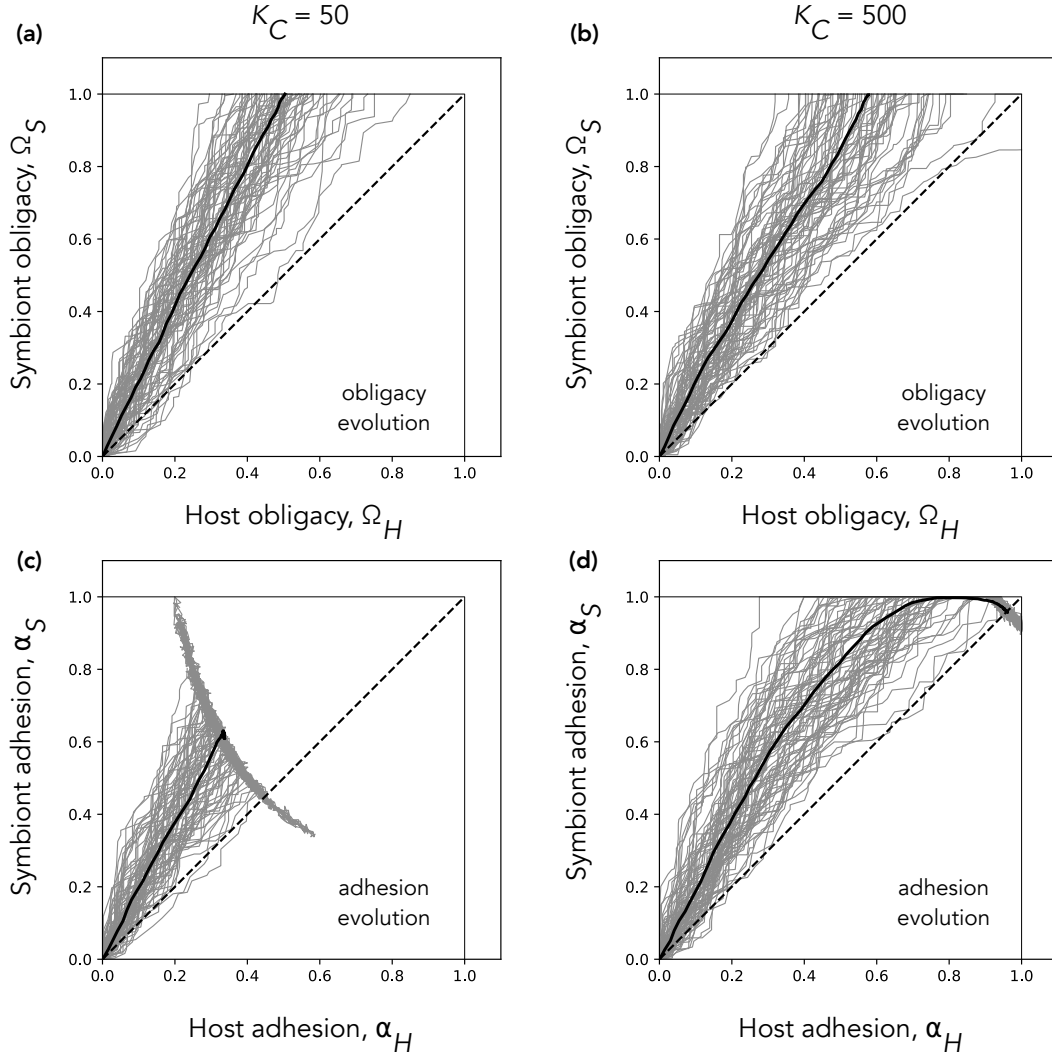


Figure 4: **Collective carrying capacity qualitatively and quantitatively affects ultimate trait values.** Each panel contains the results of 75 independent simulation runs (in grey), with the average trajectory in solid black. Both values of $K_C = 50, 500$ in this figure are high enough to allow traits to increase from $(0,0)$. $K_C = 10$ was not. When K_C is high enough, obligacies increase such that symbiont obligacy evolves to 1 and host obligacy subsequently is no longer under selection. Ultimate values of adhesion are affected by the precise values of K_C . The question of how high K_C must be is addressed in the main text. This outcome did not change between simulations run under low and high values of $r_C = 1$ and $r_C = 40$ respectively. **Parameter values.** $a_0 = 0.1, d = 50, r_H = 8, r_S = 20, r_C = 40, K_H = 100, K_S = 200$; r_C and K_C vary as indicated in the figure.

of obligacies and adhesions at two values of the parameters K_C , both of which are high enough to allow traits to increase from (0,0) but quantitatively different.

Obligacy evolution is not affected by the value of K_C . This is in line with our intuition developed in the previous sections – the defining event in an increasing obligacy trajectory is the location of its arrival at the unit square’s boundary, and this cannot be affected by K_C , a parameter that does not bias trajectories preferentially towards hosts or symbionts. The bias $\Omega_S > \Omega_H$ does indeed persist, induced by the difference in host and symbiont population growth parameters – something we did not manipulate for these tests.

Adhesion trajectories, however, are affected by the precise value of K_C . This is because the important event here is the trajectory’s arrival at the neutral ridge of ESSs described in the previous section. The presence of this ridge is explained by the non-monotonically varying benefit that dissociation of collectives provides to independent host and symbiont populations. Therefore it is intuitive that the ridge’s location is affected by K_C , a parameter that strongly sets the equilibrium abundance of collectives. At lower K_C , the relative number of collectives is lower and hence dissociation becomes non-beneficial at relatively lower values of the adhesions and hence relatively higher values of the effective dissociation rate $d_0(1 - \alpha_H\alpha_S)$.

Note that the ridge when $K_C = 50$ also prevents the adhesions from reaching maximum values even for the symbiont (despite the bias $\alpha_S > \alpha_H$ persisting for identical reasons as above). When the carrying capacity K_C is very high (see SI §S.3.3), this ridge is extremely close to (1,1) and the adhesions effectively reach their maximum values. The slight movements around $\Omega_H(t_{\Omega_S}^*)$ and the drift of adhesions along the ridge are artefacts of floating-point errors as before and can be ignored; we refrain from finding an involved artificial way to remove these since their origin is provably clear (see previous sections).

The $r - K$ difference in importance also is reflected in the variation of the equilibrium population sizes as traits evolve (SI §S.2.1). The pattern of variation supports the claims we have made so far – it is not qualitatively affected by the value of r_C , but is so affected by changes in K_C . The ridge in population sizes can also be shown to be closer to (1,1) for higher K_C . This agrees with

the intuition that the scale of the equilibrium population size x_C^* is set more strongly by K_C , with a smaller spread around it that is determined by other parameters. The maximum population size attained by the collective is its carrying capacity, and this takes place when at least one of $(\Omega_H, \Omega_S) = 1$, and when both $\alpha_H = \alpha_S = 1$.

Our analysis suggests that the pattern of variation switches depending on which quantity is bigger - dK_C or aK_HK_S , i.e. outflow or inflow into the population i when $\Omega_i = 1$. In particular, when $\Omega_H = \Omega_S = 0$, population sizes equilibrate to $x_H^* = K_H, x_S^* = K_S$; when $\Omega_i = 1$ for $i = \{H, S\}$, then $x_i^* = \frac{dK_C}{aK_j}$ where j is the other species. This suggests a sufficient condition on how high K_C must be: if $K_C > \frac{aK_HK_S}{d}$, hosts and symbiont populations are incentivised to increase obligacy. We tested this by varying a and d to perturb the same threshold quantity, and the results support our claim (see SI section §S.3.2). These arguments are difficult to make fully explicit given intractable analytical results – a problem that we directly address in the next section.

Does mutual dependence evolve before or after reproductive cohesion?

To understand the co-evolution of both obligacies and adhesions across the host and symbiont, we introduce a simpler version of Model (1). The motivation is two-fold: first, it allows for analytical tractability, enabling better understanding of our model; second, it restricts focus to the early evolutionary dynamics, since the dynamics at the boundaries is strongly dependent on the parameters b_i and d_i .

This simpler model assumes that the collective population exhibits unbounded exponential growth, i.e. that the carrying capacity is infinitely high. The “exponential model” will henceforth

407 refer to the following system of ordinary differential equations:

$$\begin{aligned}
 \dot{x}_H &= f_H x_H \left(1 - \frac{x_H}{K_H}\right) - a x_H x_S + d x_C, \\
 \dot{x}_S &= f_S x_S \left(1 - \frac{x_S}{K_S}\right) - a x_H x_S + d x_C, \\
 \dot{x}_C &= f_C x_C + a x_H x_S - d x_C.
 \end{aligned} \tag{8}$$

408 Analogously, Eq. system (1) will now be referred to as the “logistic model” when necessary. One
 409 must keep in mind that comparisons to the logistic model of the previous sections can be made
 410 only when the traits Ω_i and α_i (and hence f_C , see the mapping (2)) are of small value, since
 411 once f_C is high enough, the self-limitation term $-f_C x_C / K_C$ of the logistic growth rate becomes
 412 relevant. The formation of the collective can here be formally shown to be comparable to a
 413 mutualism between the host and symbiont, see SI §S.1.2.

414 Many statements can now be made regarding the existence, feasibility, and stability of pop-
 415 ulation dynamical equilibria. The precise statements are relegated to the SI, in §S.2.2 and §S.4.2;
 416 however the important features are that there is only one stable internal fixed point, which is
 417 feasible when the dissociation rate d is high enough to counteract the exponential growth of the
 418 collective. This is also borne out in the evolutionary dynamics of Ω_i and α_i under the exponen-
 419 tial model (see §S.4.2); the traits increase until a point, and much before either has reached its
 420 maximum value, the collective’s reproduction is so strong that no ecological equilibrium exists
 421 and the collective population blows up to infinity. Analogous to Eq. (5), an R_0 value can also
 422 be calculated here, and it is very close in functional form to the R_0 under the logistic model. In
 423 particular, it is given by

$$R_{0,exp}(\tilde{\Omega}_H, \Omega_H, \Omega_S) = \frac{\tilde{f}_H + \frac{d a x_S^*}{(d - \tilde{f}_C)}}{\tilde{f}_H \frac{x_H^*}{K_H} + a x_S^*} \tag{9}$$

424 and is exactly the result of taking $K_C \rightarrow \infty$ in the logistic R_0 (see Eq. (5)). This shows that
 425 the evolutionary dynamics captured by the exponential model are, at early trait values, truly
 426 comparable to the logistic model (1). In fact, one can go further here, and derive an exact
 427 invasion criterion – suppose a host mutant arises with obligacy $\tilde{\Omega}_H$ and adhesion $\tilde{\alpha}_H$. We show

(§S.4.2) that this mutant will invade if and only if

$$\frac{\tilde{f}_H}{\tilde{a}} \left(1 - \frac{\tilde{d}}{\tilde{f}_C} \right) > \frac{f_H}{a} \left(1 - \frac{d}{f_C} \right). \quad (10)$$

where quantities with a tilde (\sim) are associated to the mutant. Since the host and symbiont are in our model 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 underlying quantity is maximised. The existence of such an optimisation principle (Gyllenberg and Service, 2010) also means that there can never be evolutionary branching in our traits at these early times, see §S.4.4. Finally, using the constraints we set up in the Model section and §S.1.1, it can also be shown that all four traits $\Omega_H, \Omega_S, \alpha_H, \alpha_S$ must monotonically increase over evolutionary time.

To study dependence-cohesion coevolution, we again consider the mapping in Eq. (2); results of the simulations are shown in Fig. 5. We represent the degree of dependence between the host and symbiont by the product $\Omega_H \Omega_S$ of their obligacies, and the degree of reproductive cohesion by $\alpha_H \alpha_S$ ¹. These results show two important facts. First, for both the host and symbiont, it is adaptive to evolve such that $\Omega_i > \alpha_i$, i.e. being more obligate than cohesive, results from natural selection. Focusing on $\Omega_H \Omega_S$ and $\alpha_H \alpha_S$, one concludes that evolutionary trajectories are biased toward more mutual dependence than reproductive cohesion. This is a central result. It shows that, over time, we expect that a host-symbiont collective evolves such that the partners are closer to complete mutual dependence than to reproducing synchronously.

Fundamentally, this is because the functional effects of the pairs of traits are different. Recall that both trait pairs increase because of the better growth of the collective. However, there is a difference between them: independent growth of type i ($\in \{H, S\}$) is a function only of Ω_i , whereas the effect of the α_i occurs only in “coordination” with both species since dissociation

¹The multiplicative form here is connected to how they map to ecological parameters - 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 $\Omega_H + \Omega_S$.

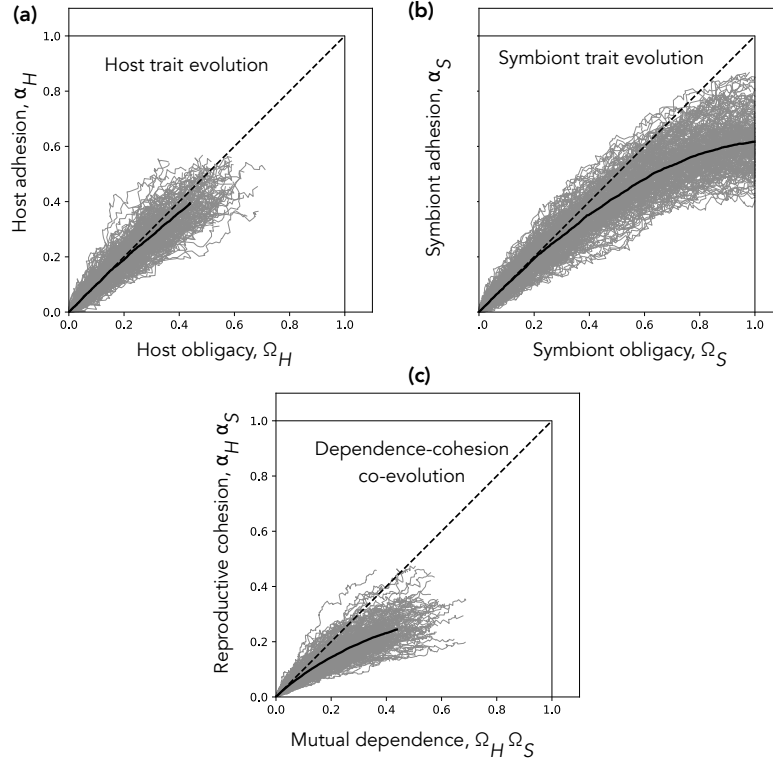


Figure 5: **Mutual dependence evolves before collective reproduction.** Here we consider the co-evolution of the four traits Ω_H, α_H and Ω_S, α_S under the exponential model (Eqs. 8). **(a, b)** Both host and symbiont traits evolve such that, on average, $\Omega_i > \alpha_i$. Notice that symbiont evolution (panel **(b)**) shows that symbiont traits evolve to much higher values than host traits. Both these traits align the host's and symbiont's reproductive traits, 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 mutual dependence evolves faster than reproductive cohesion. The trajectories stop when they hit the feasibility bound; it is nontrivial to visualise its manifestation in this plane and is hence absent. **Parameter values.** $K_H = 100, K_S = 200, K_C = 500, a_0 = 0.1, r_H = 8, r_S = 20, r_C = 10, d_0 = 50.0$.

is a function of the product of α_H and α_S (see Equation (2)). Therefore, an increase in Ω_i has a higher functional effect (in increasing the invasion fitness) than the same increase in α_i , since the latter's effect is damped by the other adhesion as well (since $\alpha_i \in [0, 1]$). See SI §S.4.5 for a more comprehensive explanation via other functional form choices.

This result is robust to the choice of parameter values, different generation times, additive (instead of multiplicative) effect of the individual traits on the collective, etc. (SI Fig. S.14). The bias persists, and mutual dependence evolves faster than reproductive cohesion. This strongly suggests that endosymbioses in nature are more likely to be mutually dependent than cohesive.

Conclusions and conjectures

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

We study two significant characteristics of an endosymbiotic collective undergoing an evolutionary transition - the reproductive cohesion of the host and symbiont (affected by host and symbiont adhesion, α_H and α_S), and the level of mutual dependence between them (obligacy, Ω_H and Ω_S). Our model shows that when obligacies evolve independently, one expects the symbiont obligacy to increase to its maximum value first, which can lead to the host obligacy being under selection to either increase, decrease, or stay constant under different ecological scenarios. This might explain the diversity in dependence outcomes that we see in nature, and we identify the two parameters in our model, b_i and d_i ($i = H, S$), that control this outcome. The adhesions also increase (with symbiont adhesion reaching high values first), ultimately leading to the accumulation of neutral variation due to drift of the evolutionary trajectory along a line of evolutionarily stable strategies. The final value of the adhesions depends on the collective carrying

capacity. Second, we show that host-symbiont asymmetries in evolutionary outcomes arise due to differences in their population growth rates and carrying capacities. Lastly, we show that at early evolutionary times, the density of permitted evolutionary trajectories in the dependence-cohesion plane is not uniform: when Ω_i and α_i co-evolve, both species evolve to be more obligate than adhesive, irrespective of the rate of these traits' evolution.

Our work highlights the importance of considering differences between host and symbiont population growth parameters. That the symbiont trait (either obligacy or adhesion) increases faster than its host counterpart is due to the symbiont's larger population size and hence faster rate of evolution: an instantiation of the Red King effect (Bergstrom and Lachmann, 2003; Frean and Abraham, 2004) – such a bias must be expected when host and symbiont interests are aligned, and defection (lower Ω or α) leads to a lower payoff. In the original Red King effect, the interest-aligning mechanism is a mutualism – an ecological interaction. Our analysis shows that the collective's formation (and shared fate) is another fundamental, interest-aligning mechanism in evolutionary transitions. Further, this faster rate of evolution has qualitatively novel effects in our model: in the case of obligacy, the symbiont evolves to full obligacy first, which leads to conditions where the host obligacy can be under selection to increase, decrease, or neutrally drift. Our model predicts full dependence for the interaction partner with higher growth rate and carrying capacity, and parameter-dependent scenarios for the partner with lower growth rate and carrying capacity. There are at least two more qualitatively different routes to asymmetric investments: first, the growth tradeoffs $f_H(\Omega_H)$ and $f_S(\Omega_S)$ for host and symbiont investing in collective reproduction might be different, and might in nature take non-linear forms. The work of Estrela et al. (2015) suggests that this might have qualitative impact: in their model, obligate exploitation can be observed over mutual dependence when the growth benefit of losing a costly function is accelerating in the amount of function lost. Further, we assumed that the within-collective reproduction and mortality rates are identical for host and symbiont. We assumed this for explanatory clarity, but it is almost certainly not true, and will give rise to new layers of asymmetry. Since these parameters are central to the evolutionary outcome, it is imperative in future work to mea-

sure them experimentally, and characterise differences in their values. Developing this part of the theory might also help understand why current host-symbiont relationships are often biased towards the symbiont being much more obligate than the host. This is clearest in the evolution of tiny genomes in endosymbionts (Bennett and Moran, 2015; McCutcheon and Moran, 2011; Moran and Bennett, 2014), and the ability to go through symbiont loss and replacements in some hosts (Chong and Moran, 2018). However, reductive genome evolution has also been shown to be driven by other mechanisms such as Muller’s ratchet (Moran, 1996) or environment-induced redundancy (Husnik et al., 2013; McCutcheon and Moran, 2010; Morris et al., 2012), in addition to Red King-type effects. These latter effects are predicated on endosymbionts occurring within the hosts and not the other way around. Simple factors such as growth parameters 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 (Daubech et al., 2017; Press and Dyson, 2012).

Another main aim was to understand why an evolutionary transition does not always take place. In this context, our model suggests the importance of the within-collective birth and death rates. They control model behaviour at two important timepoints: whether the obligacies initially increase, and what happens to the host obligacy once the symbiont becomes completely obligate. If within-collective births are stronger than deaths, (i) both obligacies increase initially, and (ii) host obligacy decreases to zero once the symbiont has become obligate. If the opposite is true i.e. within-collective deaths are stronger than births, (i) the obligacies increase only above a certain threshold initial value, and (ii) the host becomes completely obligate once the symbiont has done so. Our model thus demonstrates that the evolutionary dynamics of dependence can impose strong constraints on the emergence of obligate endosymbioses. If the within-collective rates depend on the evolving traits Ω_i and α_i , it would allow for the situation where births are stronger than deaths initially (so both host and symbiont obligacies initially increase) and deaths are stronger than births later (and hence both become fully obligate). However, the rates may depend on different (combinations of) traits, and different dependence-structures of b_i, d_i

on $\Omega_i, \alpha_i, i \in \{H, S\}$ might have different consequences. The values of b_i and d_i also need not be equal in the way that we have considered. Understanding the consequences of these extensions constitutes, in our view, the next study necessary in this body of work. Lastly, full adhesions can be achieved only when the collective carrying capacity is very high; when it is lower, final adhesion values must be interpreted carefully since our model predicts drift over time and hence variability.

The prediction that dependence evolves faster than cohesion at early times can also be confronted with biological examples. Following the work of [Estrela et al. \(2016\)](#), we compiled a short, non-exhaustive list of well-studied endosymbioses (Table 3) where there is information on the level of vertical transmission (a proxy for reproductive cohesion) and the degree of mutual dependence. However, due to the qualitative nature of the data, the causal mechanics of the interactions are inconclusive. Some cases show direct connection to our theory: in the *Riftia-Endoriftia* endosymbiosis, there is high dependence but with horizontal transmission, others do not – 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. There is indirect evidence that mutual dependence is easier to evolve than reproductive cohesion, since it is widely observed empirically ([Kost et al., 2023](#)), understood well via e.g. the Black Queen effect ([Morris et al., 2012](#)), can evolve rapidly in an experiment ([D’Souza and Kost, 2016](#)), and in principle requires very few traits ([Estrela et al., 2015](#)). On the other hand, few studies have precisely estimated quantities related to the level of vertical and horizontal transmission ([Darby and Douglas, 2003](#); [Hosokawa et al., 2007](#)). Further, it is not clear if there are genetic constraints that would lead to correlations in obligacy and adhesion – a strong enough correlation could shift in a new direction the bias towards codependence that we observe. On the theoretical side, an interesting extension of our model is to understand if the bias towards dependence that we see survives to later evolutionary times, when traits are closer to their final values. This study and its limitations thus highlight the need for a tighter connection between empirical work and theory and the requisite experiment where the complete

evolutionary transition can be quantified.

Nonetheless, taken as a null model, our study shows that an evolutionary transition in individuality is far from an inevitable outcome of the existence of a higher level of selection, suggesting that one might need invoke different mechanisms to explain phenomena such as eukaryogenesis. While our model already allows for many of the scenarios that one would like to capture, many possible extensions exist that could improve the biological realism. Perhaps most importantly, we consider here the extreme case of the interaction becoming beneficial only in close proximity; one can also construct a model where interaction benefits are relevant even outside the collective. This would add Lotka-Volterra-type cross terms to the ecological dynamics captured by Eq. (1). Additionally, we set the collective's carrying capacity K_C to a constant value independent of host or symbiont traits, since we envision that the mechanism of benefit exchange underlying this endosymbiosis enables the colonisation of a new niche (broadly construed) and that this niche has an associated carrying capacity. Better or worse usage of the niche (via different trait values) does not guarantee a higher maximum occupancy, only a quicker rate of reaching it. Further, we consider that the collectives must always be formed by a fixed density of symbionts acquired together and only once, i.e. we do not consider symbiont growth inside a single host. The size of these symbiont blocks does not affect the dynamics of our model per se, and so one can think of a single symbiont inside each host without loss of generality. However, in reality, each host has a dynamically changing symbiont population, and successive symbiont acquisition events would allow e.g. decreasing symbiont population sizes to be propped up by immigration, or to be qualitatively changed by mutualism or antagonism between different symbiont strains (see [Patel and West \(2022\)](#)). Selection on the host due to the interactions with multiple symbionts might add further selection pressures that lead to a transition. Lastly, we also do not consider that there might be competition between the independent types and the collective due to niche overlap. Relaxing each of these assumptions are worthy and important directions for future work.

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

Table 3: 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), adapted from a similar table compiled by [Estrela et al. \(2016\)](#). The information present, however, is (i) qualitative, (ii) a snapshot of the current state. The current state does not carry information about the full evolutionary trajectory, and hence is insufficient to estimate the traits Ω_i and α_i . Therefore the available information does not allow us to confirm/refute the bias towards mutual dependence that our work predicts.

Acknowledgments

Data and Code Availability

All data and simulation codes for generating the figures are available on Anonymized Github (https://anonymous.4open.science/r/incipient_endosymbiosis-563C).

Literature Cited

- Bennett, G. M., and N. A. Moran. 2015. Heritable symbiosis: The advantages and perils of an evolutionary rabbit hole. *Proceedings of the National Academy of Sciences* 112:10169–10176.
- Bergstrom, C. T., and M. Lachmann. 2003. The red king effect: When the slowest runner wins the coevolutionary race. *Proceedings of the National Academy of Sciences* 100:593–598.
- Brockhurst, M. A., D. D. Cameron, and A. P. Beckerman. 2024. Fitness trade-offs and the origins of endosymbiosis. *PLOS Biology* 22:e3002580.
- Buchner, P. 1965. *Endosymbiosis of Animals with Plant Microorganisms*. Wiley.
- Chong, R. A., and N. A. Moran. 2018. Evolutionary loss and replacement of *buchnera*, the obligate endosymbiont of aphids. *The ISME Journal* 12:898–908.
- Cohen, M. L., E. V. Mashanova, S. V. Jagannathan, and W. Soto. 2020. Adaptation to pH stress by *vibrio fischeri* can affect its symbiosis with the hawaiian bobtail squid (*euprymna scolopes*). *Microbiology* 166:262–277.
- Corsaro, D., D. Venditti, M. Padula, and M. Valassina. 1999. Intracellular life. *Critical Reviews in Microbiology* 25:39–79.
- Darby, A. C., and A. E. Douglas. 2003. Elucidation of the transmission patterns of an insect-borne bacterium. *Applied and Environmental Microbiology* 69:4403–4407.
- Daubech, B., P. Remigi, G. Doin de Moura, M. Marchetti, C. Pouzet, M.-C. Auriac, C. S. Gokhale, C. Masson-Boivin, and D. Capela. 2017. Spatio-temporal control of mutualism in legumes helps spread symbiotic nitrogen fixation. *eLife* 6.
- Denison, R. F., and E. T. Kiers. 2011. Life histories of symbiotic rhizobia and mycorrhizal fungi. *Current Biology* 21:R775–R785.

- Dercole, F., and S. Rinaldi. 2008. *Analysis of Evolutionary Processes: The Adaptive Dynamics Approach and Its Applications*. Princeton University Press.
- Dieckmann, U., and R. Law. 1996. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *Journal of Mathematical Biology* 34:579–612.
- DiSalvo, S., T. S. Haselkorn, U. Bashir, D. Jimenez, D. A. Brock, D. C. Queller, and J. E. Strassmann. 2015. Burkholderia bacteria infectiousy induce the proto-farming symbiosis of dictyostelium amoebae and food bacteria. *Proceedings of the National Academy of Sciences* 112.
- D’Souza, G., and C. Kost. 2016. Experimental evolution of metabolic dependency in bacteria. *PLOS Genetics* 12:e1006364.
- Embley, T. M., and B. J. Finlay. 1994. Systematic and morphological diversity of endosymbiotic methanogens in anaerobic ciliates. *Antonie van Leeuwenhoek* 64:261–271.
- Estrela, S., B. Kerr, and J. J. Morris. 2016. Transitions in individuality through symbiosis. *Current opinion in microbiology* 31:191–198.
- Estrela, S., J. J. Morris, and B. Kerr. 2015. Private benefits and metabolic conflicts shape the emergence of microbial interdependencies. *Environmental Microbiology* 18:1415–1427.
- Fan, L., D. Wu, V. Goremykin, J. Xiao, Y. Xu, S. Garg, C. Zhang, W. F. Martin, and R. Zhu. 2020. Phylogenetic analyses with systematic taxon sampling show that mitochondria branch within alphaproteobacteria. *Nature Ecology & Evolution* 4:1213–1219.
- Felbeck, H., and J. Jarchow. 1998. Carbon release from purified chemoautotrophic bacterial symbionts of the hydrothermal vent tubeworm riftia pachyptila. *Physiological Zoology* 71:294–302.
- Frean, M. R., and E. R. Abraham. 2004. Adaptation and enslavement in endosymbiont-host associations. *Phys. Rev. E* 69:051913.
- Gehrer, L., and C. Vorburger. 2012. Parasitoids as vectors of facultative bacterial endosymbionts in aphids. *Biology Letters* 8:613–615.

- Geritz, S. A. H., E. Kisdi, G. Meszéna, and J. A. J. Metz. 1998. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evolutionary Ecology Research* 12:35–57.
- Graf, J. S., S. Schorn, K. Kitzinger, S. Ahmerkamp, C. Woehle, B. Huettel, C. J. Schubert, M. M. M. Kuypers, and J. Milucka. 2021. Anaerobic endosymbiont generates energy for ciliate host by denitrification. *Nature* 591:445–450.
- Gruber-Vodicka, H. R., U. Dirks, N. Leisch, C. Baranyi, K. Stoecker, S. Bulgheresi, N. R. Heindl, M. Horn, C. Lott, A. Loy, M. Wagner, and J. Ott. 2011. Paracatenula, an ancient symbiosis between thiotrophic alphaproteobacteria and catenulid flatworms. *Proceedings of the National Academy of Sciences* 108:12078–12083.
- Gündüz, E. A., and A. Douglas. 2008. Symbiotic bacteria enable insect to use a nutritionally inadequate diet. *Proceedings of the Royal Society B: Biological Sciences* 276:987–991.
- Gyllenberg, M., and R. Service. 2010. Necessary and sufficient conditions for the existence of an optimisation principle in evolution. *Journal of Mathematical Biology* 62:359–369.
- Hansen, A. K., and N. A. Moran. 2011. Aphid genome expression reveals host-symbiont cooperation in the production of amino acids. *Proceedings of the National Academy of Sciences* 108:2849–2854.
- Herre, E. A., K. C. Jandér, and C. A. Machado. 2008. Evolutionary ecology of figs and their associates: Recent progress and outstanding puzzles. *Annual Review of Ecology, Evolution, and Systematics* 39:439–458.
- Herron, M., P. Conlin, and W. Ratcliff. 2022. *The Evolution of Multicellularity*. Evolutionary Cell Biology. CRC Press.
- Hölldobler, B., and E. O. Wilson. 2008. *The Superorganism: The Beauty, Elegance, and Strangeness of Insect Societies*. WW Norton.

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

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

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

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

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

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

672 Kost, C., K. R. Patil, J. Friedman, S. L. Garcia, and M. Ralser. 2023. Metabolic exchanges are
 673 ubiquitous in natural microbial communities. *Nature Microbiology* 8:2244–2252.

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

675 Lane, N., and W. Martin. 2010. The energetics of genome complexity. *Nature* 467:929–934.

676 Law, R., and U. Dieckmann. 1998. Symbiosis through exploitation and the merger of lineages in
 677 evolution. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 265:1245–
 678 1253.

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

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

683 McCutcheon, J. P., B. R. McDonald, and N. A. Moran. 2009. Convergent evolution of metabolic
684 roles in bacterial co-symbionts of insects. *Proceedings of the National Academy of Sciences*
685 106:15394–15399.

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

688 ———. 2011. Extreme genome reduction in symbiotic bacteria. *Nature Reviews Microbiology*
689 10:13–26.

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

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

694 Mishra, P., and D. C. Chan. 2014. Mitochondrial dynamics and inheritance during cell division,
695 development and disease. *Nature Reviews Molecular Cell Biology* 15:634–646.

696 Moran, N. A. 1996. Accelerated evolution and muller's ratchet in endosymbiotic bacteria. *Pro-*
697 *ceedings of the National Academy of Sciences* 93:2873–2878.

698 Moran, N. A., and G. M. Bennett. 2014. The tiniest tiny genomes. *Annual Review of Microbiology*
699 68:195–215.

700 Morris, J. J., R. E. Lenski, and E. R. Zinser. 2012. The black queen hypothesis: Evolution of
701 dependencies through adaptive gene loss. *mBio* 3.

702 Nguyen, P. L., and M. van Baalen. 2020. On the difficult evolutionary transition from the free-
703 living lifestyle to obligate symbiosis. *PLOS ONE* 15:e0235811. Publisher: Public Library of
704 Science.

- Nussbaumer, A. D., C. R. Fisher, and M. Bright. 2006. Horizontal endosymbiont transmission in hydrothermal vent tubeworms. *Nature* 441:345–348.
- Otto, S. P., and T. Day. 2007. *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution*. Princeton University Press, Princeton, New Jersey.
- Pankey, M. S., R. L. Foxall, I. M. Ster, L. A. Perry, B. M. Schuster, R. A. Donner, M. Coyle, V. S. Cooper, and C. A. Whistler. 2017. Host-selected mutations converging on a global regulator drive an adaptive leap towards symbiosis in bacteria. *eLife* 6:e24414.
- Patel, M., and S. West. 2022. Microbial warfare and the evolution of symbiosis. *Biology Letters* 18.
- Prechtel, J., C. Kneip, P. Lockhart, K. Wenderoth, and U.-G. Maier. 2004. Intracellular spheroid bodies of rhopalodia gibba have nitrogen-fixing apparatus of cyanobacterial origin. *Molecular Biology and Evolution* 21:1477–1481.
- Press, W. H., and F. J. Dyson. 2012. Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences* 109:10409–10413.
- Queller, D. C. 2000. Relatedness and the fraternal major transitions. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 355:1647–1655.
- Sachs, J. L., J. E. Russell, and A. C. Hollowell. 2011. Evolutionary instability of symbiotic function in bradyrhizobium japonicum. *PLoS ONE* 6:e26370.
- Sagan, L. 1967. On the origin of mitosing cells. *Journal of Theoretical Biology* 14:225–IN6.
- Scarborough, C. L., J. Ferrari, and H. C. J. Godfray. 2005. Aphid protected from pathogen by endosymbiont. *Science* 310:1781–1781.
- Shigenobu, S., H. Watanabe, M. Hattori, Y. Sakaki, and H. Ishikawa. 2000. Genome sequence of the endocellular bacterial symbiont of aphids buchnera sp. *APS. Nature* 407:81–86.

729 Szathmáry, E. 2015. Toward major evolutionary transitions theory 2.0. *Proceedings of the Na-*
730 *tional Academy of Sciences USA* 12:10104–10111.

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

732 Thacker, R. W. 2005. Impacts of shading on sponge-cyanobacteria symbioses: A comparison
733 between host-specific and generalist associations. *Integrative and Comparative Biology* 45:369–
734 376.

735 van den Driessche, P., and J. Watmough. 2002. Reproduction numbers and sub-threshold endemic
736 equilibria for compartmental models of disease transmission. *Mathematical Biosciences* 180:29–
737 48.

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

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

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

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

748 Wu, M., L. V. Sun, J. Vamathevan, M. Riegler, R. Deboy, J. C. Brownlie, E. A. McGraw, W. Martin,
749 C. Esser, N. Ahmadinejad, C. Wiegand, R. Madupu, M. J. Beanan, L. M. Brinkac, S. C. Daugh-
750 erty, A. S. Durkin, J. F. Kolonay, W. C. Nelson, Y. Mohamoud, P. Lee, K. Berry, M. B. Young,
751 T. Utterback, J. Weidman, W. C. Nierman, I. T. Paulsen, K. E. Nelson, H. Tettelin, S. L. O'Neill,

752 and J. A. Eisen. 2004. Phylogenomics of the reproductive parasite *wolbachia pipientis* wMel:
753 A streamlined genome overrun by mobile genetic elements. *PLoS Biology* 2:e69.

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

756 Zachar, I., and E. Szathmáry. 2017. Breath-giving cooperation: critical review of origin of mito-
757 chondria hypotheses. *Biology Direct* 12.

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

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