

Coevolutionary Arms Races and the Conditions for the Maintenance of Mutualism

Bob Week^{1,*} and Scott L. Nuismer^{1,2}

1. Program in Bioinformatics and Computational Biology, University of Idaho, Moscow, Idaho 83844; 2. Department of Biological Sciences, University of Idaho, Moscow, Idaho 83844

Submitted September 5, 2020; Accepted February 25, 2021; Electronically published May 27, 2021

Online enhancements: supplemental PDF.

ABSTRACT: Empirical evidence suggests that coevolutionary arms races between flowering plants and their pollinators can occur in wild populations. In extreme cases, trait escalation may result in evolutionary switching from mutualism to parasitism. However, theoretical approaches to studying coevolution typically assume fixed types of ecological interactions and ignore the evolution of absolute fitness. Here, we introduce a novel approach to track the evolution of absolute fitness as a framework to determine when escalatory coevolution results in a switch from mutualism to parasitism. We apply our approach to two previously studied mechanisms mediating selection as a function of phenotype. Our results demonstrate that interactions mediated by a “bigger-is-better” mechanism evolve toward parasitism. In contrast, generalizing the classical trait-matching mechanism so that the fitness of each species is optimized when trait values mismatch by a particular amount, we find theoretical support for indefinite trait exaggeration that preserves mutualistic interactions. Building on our results, we discuss the consequences of coevolutionary arms races for the maintenance of cheating. Moving beyond pairwise interactions, we consider the ramifications of coevolution in a South African pollination network for the evolution of parasitism. Future work extending our approach beyond pairwise interactions can lead to a framework for understanding the evolution of parasitism in mutualistic networks and further insights into the structure and dynamic nature of ecological communities in general.

Keywords: coevolution, mutualism, arms race, offset matching.

Introduction

Coevolution between pairs of species has long been considered an important driver of phenotypic exaggeration (Darwin 1862; Wallace 1867; Benkman et al. 2003; Brodie et al. 2005; Toju and Sota 2006; Anderson and Johnson 2007, 2009; Pauw et al. 2009; Anderson et al. 2010; Muchhala and Thomson 2010; Thompson 2014; Nuismer 2017; Nuismer

and Week 2019; Week and Nuismer 2019). In particular, coevolutionary theory predicts that antagonistic victim-exploiter interactions, such as host-parasite and predator-prey interactions, often result in coevolutionary cyclic chases (Gavrilets 1997; Gavrilets and Hastings 1998) or escalatory arms races (Nuismer et al. 2007). If an arms race goes unchecked by external sources of stabilizing selection, the interacting species are predicted to evolve ever greater defensive and offensive trait values. However, this dynamic is not unique to antagonistic interactions. Indeed, several cases of apparent trait escalation involving mutualistic interactions have been documented (Anderson and Johnson 2007, 2009; Pauw et al. 2009). Furthermore, the classification of an interaction as a mutualism or an antagonism may change as partner species evolve or as the ecological context in which the interaction occurs changes over space and time (Bronstein 1994; Gomulkiewicz et al. 2003; Nuismer et al. 1999, 2003). For example, the interactions between the pollinating seed parasite *Greya politella* and the plant it pollinates, *Lithophragma parviflorum*, changes from mutualism to parasitism as the abundance of copollinators decreases (Thompson and Pellmyr 1992). Alternatively, if one mutualistic partner outpaces the other in an evolutionary race, the carefully balanced mutual benefits may disintegrate leading to the evolutionary switching from mutualism to parasitism (Pauw et al. 2009; Jones et al. 2015).

Here, we study the disintegration of mutualisms due to escalatory coevolution. We focus on two mechanisms mediating fitness as a function of phenotype. In one model, the arms race is driven by a “bigger-is-better” (i.e., trait differences sensu Nuismer et al. 2007) mechanism where fitness increases indefinitely with trait size. In the other model, the arms race is driven by an “offset-matching” mechanism (sensu Week and Nuismer 2019), where fitness is a unimodal function peaking at some value larger than the trait of the interacting partner. Our goal is to determine whether these interaction mechanisms yield coevolutionary dynamics that preserve or destroy mutualistic interactions.

* Corresponding author; email: bobweek@gmail.com.

ORCID: Week, <https://orcid.org/0000-0001-7687-4757>; Nuismer, <https://orcid.org/0000-0001-9817-0056>.

Am. Nat. 2021. Vol. 198, pp. 195–205. © 2021 by The University of Chicago. 0003-0147/2021/19802-60172\$15.00. All rights reserved.
DOI: 10.1086/714274

Although coevolutionary models of victim-exploiter interactions often predict indefinite trait escalation, most coevolutionary models of mutualistic interactions rarely, if ever, predict coevolutionary arms races and sustained escalation of traits (Nuismer 2017). A notable exception is Kiestler et al. (1984), where it was demonstrated that sexual selection in pollinator populations can lead to escalation of pollinator preference. When plant traits evolve to track pollinator preference, sexual selection in the pollinator population may then indirectly lead to the evolutionary escalation of plant traits. However, the escalatory dynamics described in this situation are driven by sexual selection within species and not coevolutionary selection between species. Furthermore, Kiestler et al. (1984) concluded that in the absence of sexual selection, coevolution in mutualisms would lead to the stable convergence of trait values. Recent coevolutionary models focused on dynamics driven solely by interspecific interactions continue this tradition of modeling the phenotypic interface of mutualisms with a trait-matching mechanism (Guimarães et al. 2017; Nuismer 2017; Medeiros et al. 2018; de Andreazzi et al. 2019). Hence, it is generally assumed that coevolving mutualists do not engage in cycles or arms races.

While this dichotomy between antagonistic and mutualistic interactions is generally accepted, striking counterexamples exist in a range of empirically well-studied mutualistic interactions characterized by extremely exaggerated traits (Anderson and Johnson 2007, 2009; Pauw et al. 2009; Muchhala and Thomson 2010). For instance, in the interaction between the long-proboscid fly *Prosoeca ganglbaueri* and the flowering plant *Zaluzianskya microsiphon*, Anderson and Johnson (2007) found extreme trait exaggeration, including proboscis lengths of up to 50 mm and floral tube depths of up to 55 mm. Furthermore, these authors discovered significant spatial correlations in mean proboscis length and floral tube depth, suggesting a role for coevolution in explaining patterns of phenotypic exaggeration. Similarly, Pauw et al. (2009) found significant spatial correlations between another long-proboscid fly, *Moegistorhynchus longirostris*, and a long-tubed flower, *Lapeirousia anceps*. Using a clever experimental design, Pauw et al. (2009) also demonstrated evidence for reciprocal directional selection caused by this interaction for longer floral tubes and pollinator proboscises. Hence, evolutionary exaggeration in these systems may be explained by interspecific coevolutionary selection instead of intraspecific sexual selection.

These examples suggest that coevolutionary trait exaggeration in mutualisms is more widespread than previously thought. However, if trait escalation of one partner outpaces that of the other, the interaction may cease to be beneficial for the second partner. That is, the mutualism may collapse into a parasitism (Jones et al. 2015). For instance, if the fly mouthpart evolves to exceed the nectar-

tube depth by an extreme amount, the fly may be better described as a nectar thief rather than a pollinator (Inouye 1980; Pauw et al. 2009). Hence, these mutualisms may be maintained by a delicate coevolutionary dance between partner species.

In a convincing counterargument drawing on meta-analyses of selection gradients and macroevolutionary patterns, Frederickson (2017) challenged the notion that mutualisms are on the verge of breakdown. Instead, Frederickson (2017) argues that a host-parasite interaction is more likely to experience an evolutionary switch to mutualism than the other way around. Combining this argument with our knowledge of coevolutionary arms races in mutualisms leads us to two key questions: Can sustained arms races within mutualistic interactions continue indefinitely without disintegrating into antagonisms? And if so, which mechanisms mediating fitness as a function of phenotype favor the stability of mutualism in the face of ongoing arms races for ever-increasing trait values?

Here, we answer these questions by analyzing the trait-differences and offset-matching models of mutualistic arms races. By tracking the absolute fitness of interacting species in response to trait evolution, our results demonstrate that unchecked mutualistic arms races mediated by a weak selection approximation to trait differences are doomed to a parasitic fate. In contrast, offset matching (with weak or strong selection) can preserve mutualistic interactions in spite of ever-increasing trait values.

Methods

Our analysis focuses on two interaction mechanisms determining the fitness of individuals as functions of their phenotype and the phenotype of an encountered individual: the trait-differences mechanism (Nuismer et al. 2007) and the offset-matching mechanism (Week and Nuismer 2019). We therefore begin the description of our methods with a brief review of trait differences and offset matching. We then describe our approach to modeling coevolutionary dynamics and present population growth rates associated with each interaction mechanism. We show that these growth rates can be additively decomposed into a component representing the overall effect on fitness due to the interaction and a component representing effects due to sources outside the interaction. By tracking the component describing the overall effect of the interaction on absolute fitness, we develop a method to model the evolutionary switching of interaction types and the transition from mutualism to parasitism in particular. In the results section, we combine our models of coevolutionary dynamics with our approach to model the evolutionary switching of interaction types to determine when trait differences and offset

matching promote an evolutionary switch from mutualism to antagonism.

Trait Differences

The trait-differences mechanism assumes that the component of fitness due to a biotic interaction changes monotonically with the trait value of the focal individual. Nuismer et al. (2007) used trait differences to derive quantitative genetic models of coevolutionary trait escalation driven by antagonistic interactions. The coevolutionary behavior implied by the trait-differences mechanism exhibits two characteristic features. First, the mean traits will diverge toward ever larger values. Second, in the limit of weak selection, the rate of trait evolution for one species does not depend on the trait value of the other, leading to independent evolutionary trajectories.

The trait-differences mechanism has frequently been modeled using a logistic curve describing the probability of a successful interaction given the trait values of the interacting individuals (Toju and Sota 2006; Nuismer et al. 2007; Nuismer 2017; Nuismer and Week 2019). However, the same bigger-is-better mechanism can also be modeled with an exponential curve. Although not as biologically realistic, the latter option simplifies analysis and yields the same dynamics as the logistic curve in the limit of weak selection (Nuismer et al. 2007; Week and Nuismer 2019). We

therefore make use of the exponential curve instead of the logistic curve in our derivations.

Denoting x the trait value of an individual of species X , denoting y the trait value of an individual of species Y , and assuming each individual engages in a single interaction, individual fitness can be captured by

$$W_X(x, y) \propto \exp(B_X(x - y)), \quad (1a)$$

$$W_Y(y, x) \propto \exp(B_Y(y - x)), \quad (1b)$$

where $B_X, B_Y > 0$ determine the sensitivity of fitness to trait values and hence mediate the strengths of selection on the respective species. We therefore refer to these parameters as the strengths of biotic selection. This interaction mechanism is summarized graphically in figure 1.

Offset Matching

The offset-matching mechanism generalizes the trait-matching mechanism frequently used to model coevolutionary interactions (Kiestler et al. 1984; Week and Nuismer 2019). In particular, offset matching relaxes the key assumption that trait values from interacting pairs must be perfectly matched to optimize fitness for both species. Instead, fitness for each species is optimized when their trait value differs from the other by a particular amount, which we refer to as the optimal offset. By assuming that

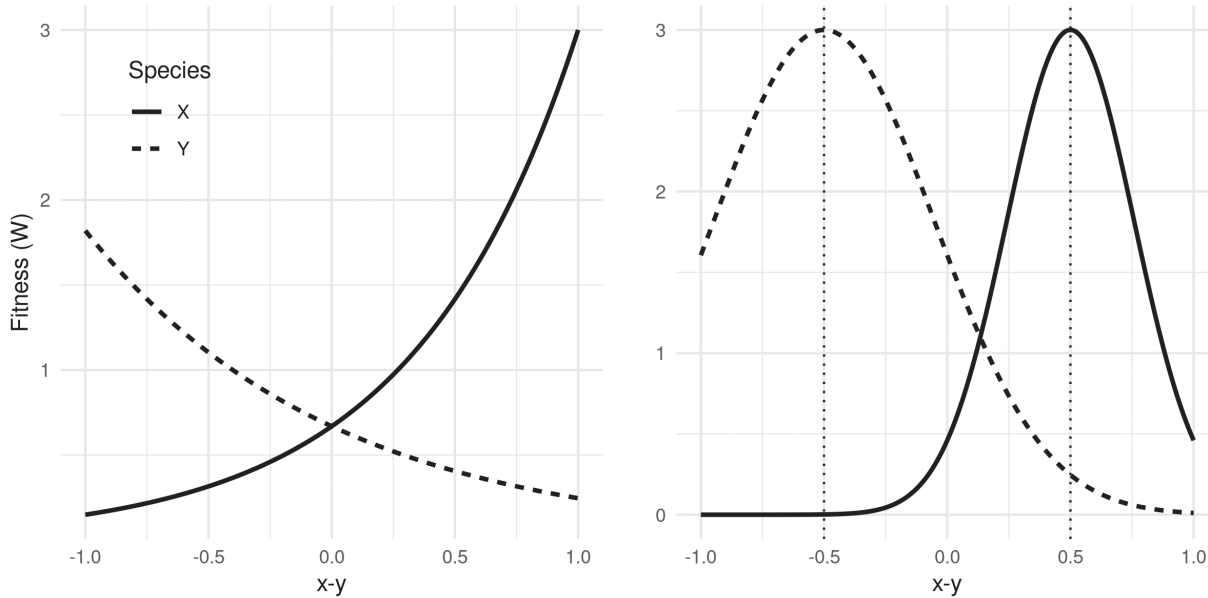


Figure 1: Fitness curves under trait differences (*left*) and offset matching (*right*) as functions of the difference in individual traits $x - y$. Solid lines represent the fitness of individuals in species X . Dashed lines represent the fitness of individuals in species Y . Vertical dotted lines mark the location of offset fitness optima. In the offset-matching case we have set $\delta = 0.5$. In both plots we have set $B_X > B_Y$ so that species X experiences stronger selection than species Y . In particular, for trait differences we set $B_X = 1.5$ and $B_Y = 1$, and for offset matching we set $B_X = 15$ and $B_Y = 5$.

fitness is maximized at a positive optimal offset for both species, we obtain a phenotypic mechanism for escalatory selection. In particular, this assumes that each species benefits from having a slightly larger trait but suffers reduced fitness beyond a certain threshold. This model of fitness was introduced by Week and Nuismer (2019) to simultaneously explain trait exaggeration and correlations of mean traits across multiple populations, enabling the use of a simple maximum likelihood approach to measure the strength of coevolution in the wild. By setting the optimal offset to zero, the trait-matching model commonly used to model mutualistic coevolution is recovered. Figure 1 compares fitness curves under trait differences and offset matching.

Following the notation introduced above, we denote by x the trait value of an individual in species X and by y a trait value in species Y . Assuming each individual participates in a single interaction, individual fitness under the offset-matching model can be described by

$$W_X(x, y) \propto \exp\left(-\frac{B_X}{2}(x + \delta - y)^2\right), \quad (2a)$$

$$W_Y(y, x) \propto \exp\left(-\frac{B_Y}{2}(y + \delta - x)^2\right), \quad (2b)$$

where δ is the optimal offset and $B_X, B_Y > 0$ again determine the strength of biotic selection on the respective species. Although it is likely that each species will experience a unique offset (say, δ_X and δ_Y), such asymmetry can be accounted for in our model by the change of variables $x' = x + \delta_X/2, y' = y + \delta_Y/2$, and $\delta = (\delta_X + \delta_Y)/2$. We therefore proceed with our assumption of identical optimal offsets without loss of generality. A formal connection with the trait-differences model can be established using a simultaneous weak selection and large optimal offset approximation. In particular, substituting $B_X, B_Y = \varepsilon^3$ and $\delta = 1/\varepsilon$, the trait-differences mechanism can be obtained using a second-order Taylor expansion around $\varepsilon \approx 0$. We illustrate this calculation in the supplemental PDF (available online).

Modeling Approach

The novelty of our approach lies in rigorous derivations of growth rates alongside trait dynamics. Beginning with the above-described models of individual fitness, we follow Week et al. (2021) and formally derive the continuous-time growth rates for interacting populations from individual-based models (detailed calculations are provided in sec. A of the supplemental PDF). The approach is essentially a generalization of the classical diffusion-limit approach introduced by Feller (1951; but see Otto and Day 2011). To illustrate the basic idea, consider a single population. First assume that the population is composed of n discrete indi-

viduals and rescale their individual contribution to population size by the amount N_0/n . Then, summing across all individuals, the initial population size is N_0 for any $n > 0$. Fitness is rescaled by $W \rightarrow W^{1/n}$ so that as individuals become more numerous and contribute less to overall population size, their expected lifetime reproductive output converges toward unity. Simultaneously, the rate at which individuals reproduce and senesce is multiplied by n . This latter step can be thought of as considering larger temporal scales over which the dynamics are occurring. Finally, the effects of mutation on offspring trait values are rescaled by the amount $1/n$ so that mutation becomes less noticeable at the individual scale with larger n . With this rescaling in place, each n corresponds to a rescaled version of the original individual-based model. While population size and mean trait value evolve in discrete jumps at birth and death events for the individual-based model, in the limit $n \rightarrow \infty$ the rescaled models converge to a limiting model with continuous evolution of mean trait and population size. This diffusion limit provides a more mathematically tractable model while preserving crucial features of the underlying biology. This approach is typically taken to study stochastic dynamics resulting from small population size. However, we make use of this method here to formally derive population-level models from interactions between individuals. To avoid eco-evolutionary feedbacks, we assume that each individual interacts with exactly one individual from the partner species. This assumption removes the dependency of selection on abundance of the partner species. To dismiss stochastic dynamics, we send the rescaled initial population sizes $N_0 \rightarrow \infty$ after we have obtained the diffusion limit from the rescaled individual-based model. Applying this approach to the models of fitness reviewed above results in a set of ordinary (deterministic) differential equations tracking the mean traits of coevolving species.

To describe our models mathematically, we denote by $m_X(x, y)$ the growth rate of individuals in species X with trait value x that interact with individuals of species Y having trait value y . Similarly, we denote by $m_Y(y, x)$ the growth rate of individuals in species Y having trait value y interacting with individuals of species X with trait value x . Averaging $m_X(x, y)$ and $m_Y(y, x)$ across interacting pairs, we obtain the continuous-time growth rates \bar{m}_X and \bar{m}_Y for species X and Y , respectively. We assume that trait values x and y are normally distributed with means \bar{x} and \bar{y} , respectively, and variances σ_x^2 and σ_y^2 , respectively. We denote by G_X and G_Y the additive genetic variances (assumed to be constant), by r_X and r_Y the intrinsic growth rates, and by e_X and e_Y the baseline effects of interactions on growth rates (i.e., the effects on growth rates due to the interaction that do not vary with trait values) for species X and Y , respectively. This notation is summarized in table 1.

Table 1: Summary of notation

Parameter	Description
x, y	Individual trait values
\bar{x}, \bar{y}	Population mean trait values
$\sigma_x^2, \sigma_y^2 \geq 0$	Population trait variances
$G_x, G_y \geq 0$	Additive genetic variances: the heritable components of σ_x^2 and σ_y^2
$m_x(x, y), m_y(y, x)$	Growth rates as functions of focal individual trait and trait of interacting partner
\bar{m}_x, \bar{m}_y	Population growth rates: averages of m_x and m_y across interacting partners
r_x, r_y	Intrinsic growth rates: population growth rates in the absence of interaction
I_x, I_y	Overall interaction effects: effects of the interaction on population growth rates
e_x, e_y	Baseline interaction effects: effects of the interaction in the absence of biotic selection
$B_x, B_y \geq 0$	Strengths of biotic selection: coevolution requires $B_x, B_y > 0$
$\delta \geq 0$	Optimal offset: determines fitness maxima in the offset-matching model

Note: Except for those indicated, all model parameters can be positive or negative.

Using \mathcal{D} to denote results under trait differences and \mathcal{O} for offset matching, we find the following pairs of population growth rates:

$$\mathcal{D} \begin{cases} \bar{m}_x = r_x + e_x + B_x(\bar{x} - \bar{y}), \\ \bar{m}_y = r_y + e_y + B_y(\bar{y} - \bar{x}); \end{cases} \quad (3a)$$

$$\mathcal{O} \begin{cases} \bar{m}_x = r_x + e_x - \frac{B_x}{2}(\bar{y} + \delta - \bar{x})^2 - \frac{B_x}{2}(\sigma_x^2 + \sigma_y^2), \\ \bar{m}_y = r_y + e_y - \frac{B_y}{2}(\bar{x} + \delta - \bar{y})^2 - \frac{B_y}{2}(\sigma_x^2 + \sigma_y^2). \end{cases} \quad (3b)$$

In section C of the supplemental PDF, we show that the trait-specific growth rates $m_x(x, y)$ and $m_y(y, x)$ do not exhibit frequency-dependent selection. Hence, following our approach to derive deterministic dynamics, we calculate the evolution of mean traits via

$$\frac{d\bar{x}}{dt} = G_x \frac{\partial \bar{m}_x}{\partial \bar{x}}, \quad (4a)$$

$$\frac{d\bar{y}}{dt} = G_y \frac{\partial \bar{m}_y}{\partial \bar{y}}. \quad (4b)$$

This yields the following two sets of mean trait dynamics:

$$\mathcal{D} \begin{cases} \frac{d\bar{x}}{dt} = G_x B_x, \\ \frac{d\bar{y}}{dt} = G_y B_y; \end{cases} \quad (5a)$$

$$\mathcal{O} \begin{cases} \frac{d\bar{x}}{dt} = G_x B_x (\bar{y} + \delta - \bar{x}), \\ \frac{d\bar{y}}{dt} = G_y B_y (\bar{x} + \delta - \bar{y}). \end{cases} \quad (5b)$$

Under our assumption of fixed phenotypic variances, the population growth rates presented in equations (3) change only through mean trait evolution. In particular,

we find growth rates for each species can be additively partitioned as $\bar{m}_x = I_x + r_x$ and $\bar{m}_y = I_y + r_y$, where I_x and I_y represent the components due to the interspecific interaction and the intrinsic growth rates r_x and r_y represent the components due to everything else. Following equations (3), trait differences and offset matching respectively yield

$$\mathcal{D} \begin{cases} I_x = e_x + B_x(\bar{x} - \bar{y}), \\ I_y = e_y + B_y(\bar{x} - \bar{y}); \end{cases} \quad (6a)$$

$$\mathcal{O} \begin{cases} I_x = e_x - \frac{B_x}{2}[(\bar{y} + \delta - \bar{x})^2 + \sigma_x^2 + \sigma_y^2], \\ I_y = e_y - \frac{B_y}{2}[(\bar{x} + \delta - \bar{y})^2 + \sigma_x^2 + \sigma_y^2]. \end{cases} \quad (6b)$$

Hence, I_x and I_y determine the type of interaction between species X and Y . Specifically, if $I_x, I_y > 0$, the interaction is a mutualism, and if either $I_x < 0$ or $I_y < 0$, the interaction is an antagonism. Thus, by tracking the signs of I_x and I_y as mean traits evolve, we can track the evolutionary switching of ecological interaction types. In particular, if an interaction mechanism promotes the evolutionary switching from a mutualism to a parasitism, then I_x and I_y begin with positive values and eventually one of either I_x or I_y becomes negative because of the evolution of \bar{x} and \bar{y} . In the following section, we combine the coevolutionary models presented above with this criterion to determine when an interaction mechanism promotes evolutionary switching from mutualism to parasitism.

Results

Trait Differences Promotes Switching to Parasitism

We begin by exploring the conditions under which sustained trait exaggeration can occur under the model of trait differences without the interaction dissolving into antagonism. Specifically, because the trait-differences model

yields sustained trait exaggeration under positive selection strengths (i.e., $B_X, B_Y > 0$), we study how the ecological nature of the interaction evolves under this regime. Combining equations (5a) and (6a), we find that the overall effects of the interaction on species growth rates evolve via

$$\frac{dI_X}{dt} = B_X(G_X B_X - G_Y B_Y), \quad (7a)$$

$$\frac{dI_Y}{dt} = B_Y(G_Y B_Y - G_X B_X). \quad (7b)$$

Hence, unless the products of additive genetic variance and strength of biotic selection is perfectly balanced between the two species so that $G_X B_X = G_Y B_Y$, the effects on species growth rates will indefinitely evolve in opposite directions. This symmetry condition is unlikely to hold in nature, implying that one of the species will evolve to become a parasite of the other. Thus, no matter how large the baseline effects e_X and e_Y are, mutualisms mediated by a trait-differences mechanism tend toward parasitism. The left panel of figure 2 displays this steady disintegration of mutualism by tracking the evolution of overall in-

teraction effects I_X and I_Y . For the particular parameters chosen, species Y becomes the host, and species X becomes the parasite. The right panel of figure 2 displays the rate of transition from mutualism to parasitism as a function of $G_X B_X - G_Y B_Y$. When $G_X B_X < G_Y B_Y$, species Y becomes the parasite, and when $G_X B_X > G_Y B_Y$, species X becomes the parasite.

There are three important caveats to this result: (1) the lack of stabilizing selection, (2) the assumed independence of ecological and evolutionary dynamics, and (3) the assumption of weak selection. Each may have important consequences for our predictions, which we work through in detail in the discussion section.

Offset Matching Stabilizes Mutualistic Interactions

Here, we assume that fitness for an individual is maximized when its trait value exceeds that of its partners' by some specific amount $\delta > 0$, which we refer to as the optimal offset (see fig. 1). Under these assumptions, equation (5b) yields ever-increasing trait values. Combining equations (5b) and (6b), we find that the effects on species growth rates evolve via

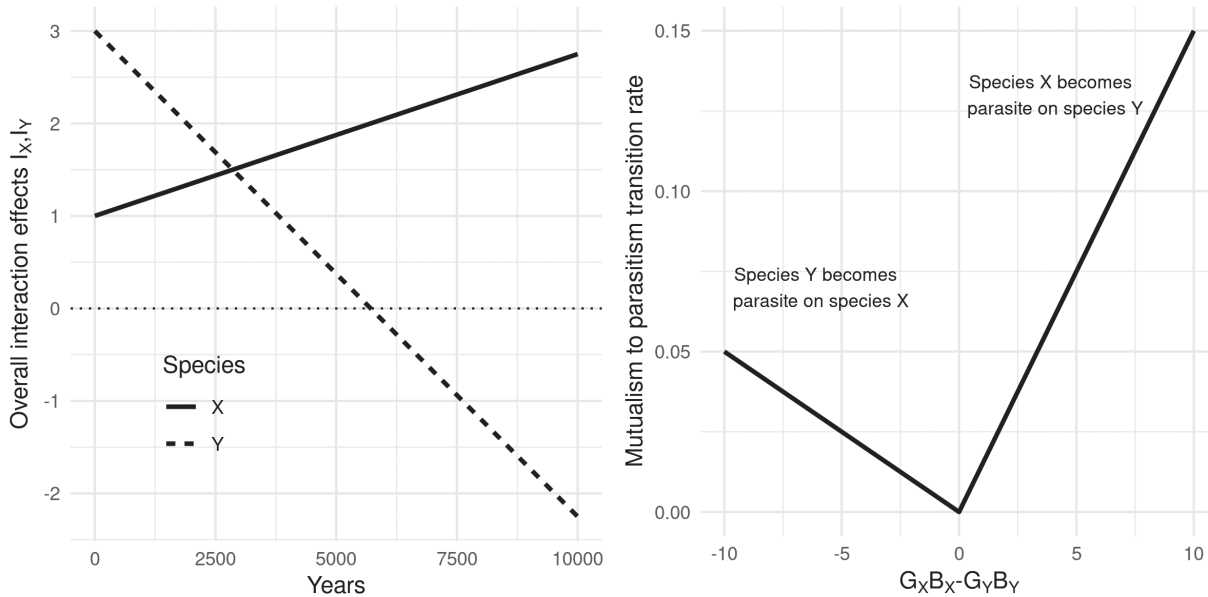


Figure 2: *Left*, time-series plot illustrating the trajectories of overall interaction effects I_X and I_Y . The solid line represents the overall interaction effect on species X , and the dashed line represents the overall interaction effect on species Y . The horizontal dotted line marks the threshold determining the interaction type. When I_X or I_Y is below this line, the interaction is an antagonism. Hence, for the particular parameters chosen, this graph shows that species X evolves to become a parasite of species Y . *Right*, rate of transition from parasitism to mutualism as a function of $G_X B_X - G_Y B_Y$ under the assumption $B_X < B_Y$. On the right side of the plane, species X becomes a parasite of species Y . On the left side of the plane, this relationship is reversed. For both plots we set $B_X = 5e - 3$ and $B_Y = 1.5e - 2$. For the left plot, we used initial values $I_X = 1$ and $I_Y = 3$ and set $G_X = 10$ and $G_Y = 1$.

$$\frac{dI_X}{dt} = -B_X(\delta - \Delta)\frac{d\Delta}{dt}, \quad (8a)$$

$$\frac{dI_Y}{dt} = -B_Y(\delta + \Delta)\frac{d\Delta}{dt}, \quad (8b)$$

where $\Delta = \bar{x} - \bar{y}$. In turn, the difference in mean traits Δ evolves via

$$\frac{d\Delta}{dt} = (G_X B_X + G_Y B_Y) \left(\frac{G_X B_X - G_Y B_Y}{G_X B_X + G_Y B_Y} \delta - \Delta \right). \quad (9)$$

Since the offset-matching model requires positive selection strengths $B_X, B_Y > 0$, equation (9) implies that the difference in mean trait values Δ will always evolve to a stable equilibrium, even though the mean traits \bar{x} and \bar{y} evolve indefinitely greater values. In particular, we find that the difference in mean trait values at equilibrium can be expressed as

$$\hat{\Delta} = \frac{G_X B_X - G_Y B_Y}{G_X B_X + G_Y B_Y} \delta. \quad (10)$$

In spite of ever-escalating mean trait values \bar{x} and \bar{y} , when Δ is at its equilibrium, the components of species growth rates I_X and I_Y have also reached a stable equilibrium, given by

$$\hat{I}_X = e_X - \frac{B_X}{2} \left[\left(\frac{2G_Y B_Y \delta}{G_X B_X + G_Y B_Y} \right)^2 + (\sigma_X^2 + \sigma_Y^2) \right], \quad (11a)$$

$$\hat{I}_Y = e_Y - \frac{B_Y}{2} \left[\left(\frac{2G_X B_X \delta}{G_X B_X + G_Y B_Y} \right)^2 + (\sigma_X^2 + \sigma_Y^2) \right]. \quad (11b)$$

Hence, trait escalation under the offset-matching model promotes the stabilization of interaction effects in the absence of all other evolutionary forces. This means that if the baseline effects e_X and e_Y are large enough, the interaction will remain a mutualism even as trait values escalate indefinitely toward larger values. Furthermore, this result implies novel interactions that begin as a parasitism may evolve toward mutualism when mediated by an offset-matching mechanism. Figure 3 illustrates these results by tracking the evolution of overall interaction effects I_X and I_Y for two scenarios.

Inspection of equations (11) reveals that the minimal baseline benefits e_X and e_Y needed to preserve mutualism increase logarithmically with the selection strength of the partner species. For example, the minimal value of e_X required to maintain benefits for species X does not increase indefinitely with B_Y . In contrast, we see that these minimal quantities do increase indefinitely with the selection strengths of their focal species, the optimal offset δ , and phenotypic

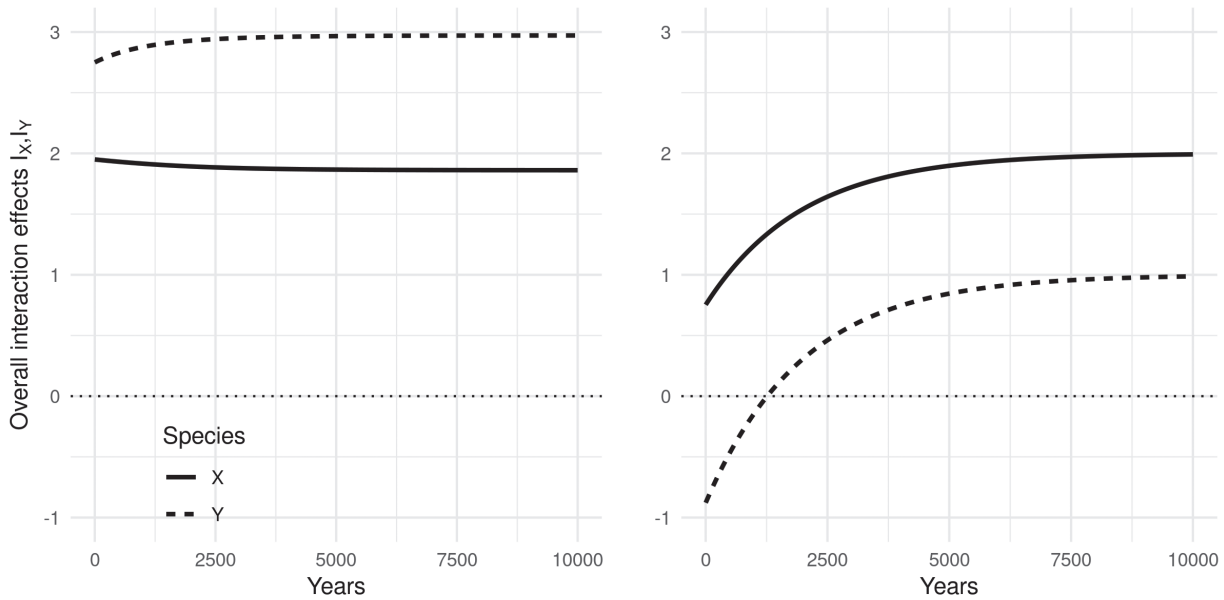


Figure 3: Time-series plots of overall interaction effects I_X and I_Y under two scenarios. Solid lines represent the overall interaction effect on species X , and dashed lines represent the overall interaction effect on species Y . The horizontal dotted lines mark the threshold determining the interaction type. When I_X or I_Y is below this line, the interaction is an antagonism. The left plot illustrates a mutualism that is preserved in spite of indefinitely escalating trait values. The right plot illustrates a novel parasitism that experiences an evolutionary switch to mutualism. Parameter values for the left plot are $\Delta_0 = 0$, $\delta = 100$, $B_X = 1e - 5$, $B_Y = 5e - 5$, $e_X = 2$, $e_Y = 3$, and $G_X = G_Y = \sigma_X^2 = \sigma_Y^2 = 10$. Parameter values for the right plot are $\Delta_0 = 500$, $\delta = 0.5$, $B_X = 1e - 5$, $B_Y = 1.5e - 5$, $e_X = 2$, $e_Y = 1$, and $G_X = G_Y = \sigma_X^2 = \sigma_Y^2 = 10$.

variances σ_x^2 and σ_y^2 . To summarize, mutualisms mediated by an offset-matching mechanism are more likely to dissolve into parasitisms when fitness is maximized for large offsets, when either species maintains a sufficiently large trait variance, or when either species experiences strong selection.

Discussion

We have shown that mutualistic interactions can be preserved in the face of indefinite coevolutionary trait escalation. However, this result depends on the mechanism mediating interspecific interactions. In particular, we found that interactions mediated by a weak selection approximation of the trait-differences mechanism, which has been traditionally employed to model trait escalation in coevolving antagonisms (Toju and Sota 2006; Nuismer et al. 2007), encourages the dissolution of mutualism into parasitism. In contrast, employing the recently introduced offset-matching mechanism (Week and Nuismer 2019), which generalizes the classical trait-matching mechanism proposed by Kiester et al. (1984), we found that mutualistic interactions are preserved when the baseline benefits of the interaction are sufficiently large (eqq. [11]).

Our analyses are particularly relevant for the study of pairwise plant-pollinator interactions. Indeed, the primary motivation for this work is the hypothesis that exaggerated floral tubes and pollinator proboscises observed in the wild are explained by pairwise coevolutionary races. Plant-pollinator interactions provide classical examples of mutualisms in which the plant provides a nutritional or metabolic resource to the pollinator in trade for movement of pollen. Under the trait-differences mechanism, our results imply that the difference between average proboscis length and average floral tube depth will increase indefinitely, leading to one of two outcomes: (1) floral tube depth will eventually exceed pollinator proboscis length, such that the pollinator is essentially tricked into transferring pollen without access to any reward, or (2) pollinator proboscis length will eventually exceed floral tube depth, such that the pollinator drains the flower of nectar without transferring pollen. In both of these cases, one species is eventually exploited while the other continues to profit from the interaction. Quantitatively, this implies that the overall interaction effect on absolute fitness becomes negative for one of the species and hence results in a host-parasite relationship.

Alternatively, results based on the offset-matching mechanism, which assumes that fitness is maximized at a particular offset instead of indefinitely increasing or decreasing with the difference in mean trait values, imply that this parasitic fate is not universal among mutualistic pairs engaged in a coevolutionary arms race. Returning to the plant-

pollinator example given above, the offset-matching mechanism implies that the difference in average proboscis length and average floral tube depth will converge to a stable equilibrium proportional to the optimal offset, even though these mean traits will themselves continue to escalate indefinitely. In this case, there are three conditions required to maintain the mutualism: (1) the optimal offset must be sufficiently small, (2) biotic selection cannot be too strong (as determined by e_x and e_y ; see eqq. [11]), and (3) local phenotypic variances cannot be too large. If all three of these conditions are satisfied, then the phenotypic distribution of each species remains in a range that benefits the other species, even though these ranges are dynamic.

One hypothesized mechanism for the evolutionary breakdown of mutualism is the evolution of cheating individuals that exploit the benefits of the interaction without reciprocation (Bronstein 2001; Ferrière et al. 2007). To understand whether this mechanism explains the evolutionary switch from mutualism to parasitism predicted by the trait-differences model, we can apply the definition of cheating provided by Jones et al. (2015). According to this definition, cheaters must satisfy two conditions. First, a cheater must enjoy greater fitness than the mean fitness of their focal population. Second, a cheater must decrease the fitness of their partner below the average fitness of the partner population. Since we assume that individuals encounter each other at random, individuals are expected to cheat under the trait-differences model whenever their trait value exceeds the average trait value of their focal population. Furthermore, since we assume that traits are symmetrically distributed, half of the individuals in each species are expected to cheat at any stage of the coevolutionary arms race. In particular, this implies that the mechanism of mutualistic breakdown studied here is not explained by the evolution of cheaters when using the definition provided by Jones et al. (2015). However, if we define cheating using absolute fitness instead of relative fitness, the frequency of cheating in the eventual parasitic species would evolve toward unity. Hence, under an alternative definition of cheating, this result is equivalent to an explanation of mutualistic breakdown by cheating. Thus, whether the mechanism of mutualistic breakdown studied here is distinct from the evolution of cheating depends on the definition of cheating employed.

Although our mathematical analyses shed light on the outcomes of pairwise interactions, plant-pollinator interactions based on a trade between pollen movement and food tend to be generalized (Waser et al. 1996; Vázquez and Aizen 2004). These interactions may therefore vary as a result of their ecological context. To understand the dynamics of evolutionary interaction switching in the presence of multispecific interactions, the analytical methods

developed here can be extended to models of coevolving mutualistic networks (e.g., Medeiros et al. 2018). In the case that a subset of the community engages in a multi-specific coevolutionary arms race, the species left behind are at risk of becoming parasitized. For example, in the fly-flower pollination system investigated by Pauw et al. (2009), the long-proboscid fly *Moegistorhynchus longirostris* is known to visit at least 20 long-tubed flower species (Manning and Goldblatt 1997). This plant guild includes *Lapeirousia anceps*, which is the most abundant and widespread member (Pauw et al. 2009). The exaggerated nectar tube of *L. anceps* exhibits strong spatial correlations with the proboscis length of *M. longirostris*. By interfacing these patterns of trait exaggeration and spatial correlation with coevolutionary theory, Week and Nuismer (2019) were able to provide quantitative evidence for a coevolutionary arms race between the two species. However, in the course of this arms race it is likely that *M. longirostris* imposed and received selection pressures from other members of the long-tubed plant guild, leading to a compartment of species engaged in a coevolutionary arms race (Pauw et al. 2009). As a consequence, mutualistic interactions between *M. longirostris* and flower species outside this compartment are at risk of disintegrating into parasitisms.

A likely antagonism that may have been produced indirectly as a by-product of coevolutionary escalation as described above occurs between *M. longirostris* and the iris *Babiana thunbergii*. *Moegistorhynchus longirostris* has been observed to frequently visit flowers of *B. thunbergii*, which are usually pollinated by malachite sunbirds and have a much shorter and wider floral morphology. Pauw et al. (2009) noted that *M. longirostris* rarely contacts the reproductive organs of *B. thunbergii* while draining its nectar and hence acts as a parasitic nectar thief. In the case that *B. thunbergii* originally profited from visits of *M. longirostris*, before the evolution of extreme proboscis lengths, external sources of selection, such as those due to sunbirds, may have countered selection for floral elongation. In this case, the disintegration of mutualism follows a different path from what we have described above for pairwise interactions. Instead of direct pairwise interactions leading to an intimate race for exaggerated trait values, indirect coevolutionary effects due to interactions with other community members impose inconsistent patterns of selection for each member of the focal pair. A recent theoretical result supports this view by suggesting that such indirect effects are as important for shaping coevolutionary dynamics in ecological networks as direct effects (Guimarães et al. 2017).

For the sake of analytical tractability and clarity of results, we have kept our analysis as simple as possible while recognizing that our assumptions are unlikely to be satisfied literally in real biological systems. For instance, ig-

norning stabilizing selection implies the absence of well-established physiological and anatomical constraints that prevent indefinite runaway coevolution. Similarly, we have ignored the possible depletion of heritable variation and the influence of abundance dynamics, and in the case of trait differences we considered only weak coevolution. Although these assumptions come at the cost of biological realism, they have allowed us to understand and illuminate the tendency of trait differences and offset matching to motivate or resist evolutionary transitions from mutualism to parasitism. Hence, our results should be interpreted in much the same way as results obtained from experiments conducted in controlled but biologically unrealistic environments. It would therefore be premature, for example, to combine our results with the argument made by Frederickson (2017)—namely, that parasitisms are more likely to experience a switch to mutualism than vice versa—to conclude that mutualistic arms races observed in the wild tend to be mediated by offset mechanisms. Indeed, mutualism in the face of an arms race may instead be explained by combining the trait-differences mechanism with abiotic stabilizing selection. Thus, to predict and understand when wild populations of coevolving mutualists are doomed to an inevitable parasitic fate, we must confront these complications directly and with more complex and biologically realistic models.

Of the large set of processes we have controlled for, perhaps the most important are eco-evolutionary feedbacks. Future models using our framework can incorporate the effects of eco-evolutionary feedbacks by replacing the assumption that each individual interacts with a single individual of their partner species with the assumption that individual interaction frequencies depend on abundances. As an example, assuming that each individual of species X interacts once with each individual of species Y (and vice versa) yields baseline benefits e_x and e_y and strengths of biotic selection B_x and B_y that depend linearly on the abundance of the partner species. Hence, this assumption leads to generalizations of classical mass action models of population dynamics, such as the Lotka-Volterra predator-prey model (Lotka 1925; Volterra 1926), that allow for the evolutionary switching of interaction type in tandem with the fluctuations of abundance. In this case, the overall interaction effects can be expressed as $I_x = \alpha_x N_y$ and $I_y = \alpha_y N_x$, where α_x and α_y represent the interaction coefficients corresponding to species X and Y , respectively. In principle, the same approach developed here to analyze evolutionary interaction switching can be directly extended to this more complicated scenario.

Conclusion

We have introduced a novel approach to model the evolutionary switching of interaction types by tracking the

evolution of absolute fitness in addition to trait dynamics. Applying our approach to two models of coevolutionary trait escalation inspired by plant-pollinator interactions revealed that mutualisms mediated by a bigger-is-better (i.e., trait-differences) mechanism inevitably dissolve into parasitism. In contrast, our results show that mutualisms mediated by an offset-matching mechanism, a generalization of classical trait matching, are preserved when the baseline benefits of the interaction (i.e., the components that do not vary with trait values) are large enough. Our results are based on minimal models of mean trait coevolution determined by different interaction mechanisms and hence ignore the effects of external evolutionary processes and eco-evolutionary feedbacks. Predicting when wild populations of mutualists are prone to disintegration via coevolutionary arms races will likely require more complex models that account for a variety of processes, such as abiotic stabilizing selection, gene flow, random genetic drift, and multispecific interactions. Future work applying the analytical approach developed here to models that account for realistic sets of evolutionary processes can enrich our understanding of the dynamical nature of ecological relationships observed in the wild and produce novel statistical tools to forecast the evolutionary stability of ecological relationships.

Acknowledgments

We thank Editor Jennifer A. Lau, Associate Editor Adam M. Siepielski, and two anonymous reviewers for their insightful and encouraging comments.

Statement of Authorship

B.W. conceived the project, formulated the model, performed the analysis, produced the figures, and wrote the original draft and revision. S.L.N. provided supervision, reviewed and edited the manuscript, and secured funding.

Data and Code Availability

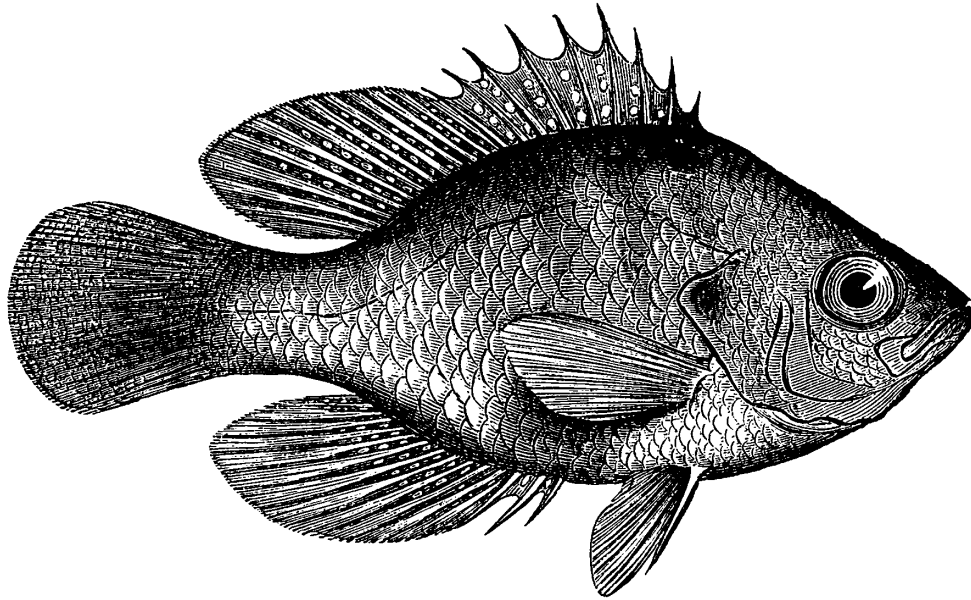
There is no code or data for this article.

Literature Cited

- Anderson, B., and S. D. Johnson. 2007. The geographical mosaic of coevolution in a plant-pollinator mutualism. *Evolution* 62:220–225.
- . 2009. Geographical covariation and local convergence of flower depth in a guild of fly-pollinated plants. *New Phytologist* 182:533–540.
- Anderson, B., J. S. Terblanche, and A. G. Ellis. 2010. Predictable patterns of trait mismatches between interacting plants and insects. *BMC Evolutionary Biology* 10:204.
- Benkman, C. W., T. L. Parchman, A. Favis, and A. M. Siepielski. 2003. Reciprocal selection causes a coevolutionary arms race between crossbills and lodgepole pine. *American Naturalist* 162:182–194.
- Brodie, E. D., III, C. R. Feldman, C. T. Hanifin, J. E. Motychak, D. G. Mulcahy, B. L. Williams, and E. D. Brodie Jr. 2005. Parallel arms races between garter snakes and newts involving tetrodotoxin as the phenotypic interface of coevolution. *Journal of Chemical Ecology* 31:343–356.
- Bronstein, J. L. 1994. Conditional outcomes in mutualistic interactions. *Trends in Ecology and Evolution* 9:214–217.
- . 2001. The exploitation of mutualisms. *Ecology Letters* 4:277–287.
- Darwin, C. 1862. On the various contrivances by which British and foreign orchids are fertilised by insects: and on the good effects of intercrossing. J. Murray, London.
- de Andreazzi, C. S., J. Astegiano, and P. R. Guimarães. 2019. Coevolution by different functional mechanisms modulates the structure and dynamics of antagonistic and mutualistic networks. *Oikos* 129:224–237.
- Feller, W. 1951. Diffusion processes in genetics. Pages 227–246 in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press, Berkeley.
- Ferrière, R., M. Gauduchon, and J. L. Bronstein. 2007. Evolution and persistence of obligate mutualists and exploiters: competition for partners and evolutionary immunization. *Ecology Letters* 10:115–126.
- Frederickson, M. E. 2017. Mutualisms are not on the verge of breakdown. *Trends in Ecology and Evolution* 32:727–734.
- Gavrilets, S. 1997. Coevolutionary chase in exploiter-victim systems with polygenic characters. *Journal of Theoretical Biology* 186:527–534.
- Gavrilets, S., and A. Hastings. 1998. Coevolutionary chase in two-species systems with applications to mimicry. *Journal of Theoretical Biology* 191:415–427.
- Gomulkiewicz, R., S. L. Nuismer, and J. N. Thompson. 2003. Coevolution in variable mutualisms. *American Naturalist* 162:S80–S93.
- Guimarães, P. R., M. M. Pires, P. Jordano, J. Bascompte, and J. N. Thompson. 2017. Indirect effects drive coevolution in mutualistic networks. *Nature* 550:511–514.
- Inouye, D. W. 1980. The terminology of floral larceny. *Ecology* 61:1251–1253.
- Jones, E. I., M. E. Afkhami, E. Akçay, J. L. Bronstein, R. Bshary, M. E. Frederickson, K. D. Heath, et al. 2015. Cheaters must prosper: reconciling theoretical and empirical perspectives on cheating in mutualism. *Ecology Letters* 18:1270–1284.
- Kiester, A. R., R. Lande, and D. W. Schemske. 1984. Models of coevolution and speciation in plants and their pollinators. *American Naturalist* 124:220–243.
- Lotka, A. 1925. *Elements of physical biology*. Williams & Wilkins, Baltimore.
- Manning, J. C., and P. Goldblatt. 1997. The *Moegistorhynchus longirostris* (Diptera: Nemestrinidae) pollination guild: long-tubed flowers and a specialized long-proboscid fly pollination system in southern Africa. *Plant Systematics and Evolution* 206:51–69.
- Medeiros, L. P., G. Garcia, J. N. Thompson, and P. R. Guimarães. 2018. The geographic mosaic of coevolution in mutualistic networks. *Proceedings of the National Academy of Sciences of the USA* 115:12017–12022.

- Muchhala, N., and J. D. Thomson. 2010. Fur versus feathers: pollen delivery by bats and hummingbirds and consequences for pollen production. *American Naturalist* 175:717–726.
- Nuismer, S. L. 2017. Introduction to coevolutionary theory. W. H. Freeman, San Francisco.
- Nuismer, S. L., R. Gomulkiewicz, and M. T. Morgan. 2003. Coevolution in temporally variable environments. *American Naturalist* 162:195–204.
- Nuismer, S. L., B. J. Ridenhour, and B. P. Oswald. 2007. Antagonistic coevolution mediated by phenotypic differences between quantitative traits. *Evolution* 61:1823–1834.
- Nuismer, S. L., J. N. Thompson, and R. Gomulkiewicz. 1999. Gene flow and geographically structured coevolution. *Proceedings of the Royal Society B* 266:605–609.
- Nuismer, S. L., and B. Week. 2019. Approximate Bayesian estimation of coevolutionary arms races. *PLoS Computational Biology* 15:e1006988.
- Otto, S. P., and T. Day. 2011. *A biologist's guide to mathematical modeling in ecology and evolution*. Princeton University Press, Princeton, NJ.
- Pauw, A., J. Stofberg, and R. J. Waterman. 2009. Flies and flowers in Darwin's race. *Evolution* 63:268–279.
- Thompson, J. N. 2014. *Interaction and coevolution*. University of Chicago Press, Chicago.
- Thompson, J. N., and O. Pellmyr. 1992. Mutualism with pollinating seed parasites amid co-pollinators: constraints on specialization. *Ecology* 73:1780–1791.
- Toju, H., and T. Sota. 2006. Adaptive divergence of scaling relationships mediates the arms race between a weevil and its host plant. *Biology Letters* 2:539–542.
- Vázquez, D. P., and M. A. Aizen. 2004. Asymmetric specialization: a pervasive feature of plant-pollinator interactions. *Ecology* 85:1251–1257.
- Volterra, V. 1926. Fluctuations in the abundance of a species considered mathematically. *Nature* 118:558–560.
- Wallace, A. R. 1867. Creation by law. *Quarterly Journal of Science* 4:470–488.
- Waser, N. M., L. Chittka, M. V. Price, N. M. Williams, and J. Ollerton. 1996. Generalization in pollination systems, and why it matters. *Ecology* 77:1043–1060.
- Week, B., and S. L. Nuismer. 2019. The measurement of coevolution in the wild. *Ecology Letters* 22:717–725.
- Week, B., S. L. Nuismer, L. J. Harmon, and S. M. Krone. 2021. A white noise approach to evolutionary ecology. *Journal of Theoretical Biology* 521:110660.

Associate Editor: Adam M. Siepielski
Editor: Jennifer A. Lau



“The habits of the species are those of the Centrarchidae generally, modified in so far as being merely more of a mud-loving species. So purely a mud-dwelling fish are they that we have frequently found them in water so shallow, that they marked the mud with their pectoral fins in swimming; preferring such shallow water, with the mud, to that which was deeper, to which they had access, because it was over a stony bed.” Figured: “Spotted Sun-fish (*Enneacanthus guttatus*).” From “Mud-Loving Fishes” by Charles C. Abbott (*The American Naturalist*, 1870, 4:385–391).