

# Eco-Evolutionary Dynamics in Metacommunities: Ecological Inheritance, Helping within Species, and Harming between Species

Charles Mullan\* and Laurent Lehmann

Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

Submitted November 6, 2017; Accepted June 26, 2018; Electronically published October 18, 2018

Online enhancements: appendixes, supplemental material. Dryad data: <http://dx.doi.org/10.5061/dryad.kk1qj10>.

**ABSTRACT:** Understanding selection on intra- and interspecific interactions that take place in dispersal-limited communities is a challenge for ecology and evolutionary biology. The problem is that local demographic stochasticity generates eco-evolutionary dynamics that are generally too complicated to make tractable analytical investigations. Here we circumvent this problem by approximating the selection gradient on a quantitative trait that influences local community dynamics, assuming that such dynamics are deterministic with a stable fixed point. The model nonetheless captures unavoidable kin selection effects arising from demographic stochasticity. Our approximation reveals that selection depends on how an individual expressing a trait change influences (1) its own fitness and the fitness of its current relatives and (2) the fitness of its downstream relatives through modifications of local ecological conditions (i.e., through ecological inheritance). Mathematically, the effects of ecological inheritance on selection are captured by dispersal-limited versions of press perturbations of community ecology. We use our approximation to investigate the evolution of helping within species and harming between species when these behaviors influence demography. We find that altruistic helping evolves more readily when intraspecific competition is for material resources rather than for space, because in this case the costs of kin competition tend to be paid by downstream relatives. Similarly, altruistic harming between species evolves when it alleviates downstream relatives from interspecific competition. Beyond these examples, our approximation can help better understand the influence of ecological inheritance on a variety of eco-evolutionary dynamics in metacommunities, from consumer-resource and predator-prey coevolution to selection on mating systems with demographic feedbacks.

**Keywords:** kin selection, adaptive dynamics, metacommunity, altruism, mutualism, niche construction.

## Introduction

Interactions within and between species are extremely common in nature and probably connect almost all living organisms to one another. How such intra- and interspecific interactions evolve depends on interdependent ecological and evolutionary processes (Urban 2011; Wagner et al. 2017; terHorst et al. 2018), also known as eco-evolutionary dynamics (for review, see Lion 2017). One major difficulty in understanding these dynamics is due to the spatial structuring of communities, which emerges from the physical limitations of movement and interactions. This spatial structure is captured by the notion of a metacommunity, in which individuals of different species are divided among small local patches connected to each other by dispersal (e.g., Hanski and Gilpin 1997; Clobert et al. 2001; Urban et al. 2008; Leibold and Chase 2017).

When dispersal among the patches of a metacommunity is limited, individual reproductive and survival variance generate local demographic stochasticity. This has two complicating consequences for eco-evolutionary dynamics. First, it causes genetic stochasticity, whereby allele frequencies fluctuate within and between patches. These fluctuations lead to the buildup of genetic relatedness between members of the same species. Genetic relatedness then influences selection on traits, in particular social traits, which, like helping, are traits that affect the reproductive success of both their actor (direct effects) and their recipient (indirect effects; e.g., Hamilton 1971; Hamilton and May 1977; Taylor 1994; Taylor and Frank 1996; Frank 1998; Rousset 2004; Lion and van Baalen 2007b; West et al. 2007; Van Cleve 2015). Second, local demographic stochasticity results in ecological stochasticity, whereby the abundances of different species fluctuate within and between patches. As a consequence, multispecies demography in patch-structured populations is significantly more complicated than in panmictic populations (Chesson 1978, 1981; Hubbell 2001; Neuhauser 2002; Cornell and Ovaskainen 2008). As genetic and ecological stochasticity are further likely to influence one another through

\* Corresponding author; email: [charles.mullan@unil.ch](mailto:charles.mullan@unil.ch).

**ORCID:** Mullan, <http://orcid.org/0000-0002-9875-4227>.

Am. Nat. 2018. Vol. 192, pp. 664–686. © 2018 by The University of Chicago. 0003-0147/2018/19206-58064\$15.00. All rights reserved. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0), which permits non-commercial reuse of the work with attribution. For commercial use, contact [journalpermissions@press.uchicago.edu](mailto:journalpermissions@press.uchicago.edu). DOI: 10.1086/700094

eco-evolutionary feedbacks, understanding selection on traits mediating ecological interactions is a major challenge when dispersal is limited.

In complicated demographic scenarios, fundamental insights into selection can be obtained from the long-term adaptive dynamics of quantitative traits. These dynamics follow the gradual changes of traits displayed by a population under the constant but limited influx of mutations (e.g., Eshel 1983, 1996; Parker and Maynard Smith 1990; Christiansen 1991; Grafen 1991; Abrams et al. 1993; Metz et al. 1996; Geritz et al. 1998; Rousset 2004). One of the main goals of evolutionary analysis is to identify local attractors of such adaptive dynamics. These attractors are trait values toward which selection drives a population under gradual evolution (referred to as convergence-stable phenotypes; Eshel 1983; Taylor 1989; Christiansen 1991; Geritz et al. 1998; Rousset 2004; Leimar 2009). Convergence-stable phenotypes can be identified from the selection gradient on a trait, which is the marginal change in the fitness of an individual due to this individual and all its relatives changing trait value. Such analysis has helped us understand how natural selection molds phenotypic traits of broad biological interest, from senescence, dispersal, mate choice, life history, sperm competition, and sex ratio to altruism, cumulative cultural evolution, bet hedging, and optimal foraging (e.g., Hamilton 1966; Charnov 1976; Schaffer 1982; Taylor 1988*a*, 1988*b*; Parker 1990; Frank 1998; Foster 2004; Gardner and West 2004; Kisdi and Priklopil 2010; Akçay and Van Cleve 2012; Kuijper et al. 2012; Mullon et al. 2014; Wakano and Miura 2014; Kobayashi et al. 2015; Weigang and Kisdi 2015; Nurmi et al. 2018).

Gold standard expressions for the selection gradient on traits that influence the demography of a single species, where all consequences of genetic and ecological stochasticity for natural selection are taken into account, have been worked out long ago (Rousset 2004, ch. 11; Rousset and Ronce 2004, eqq. [23], [24]). In principle, these expressions can be directly extended to consider multispecies interactions. However, even under the simplest model of limited dispersal, which is the island model of dispersal (Wright 1931), the evaluation of the selection gradient on traits affecting a single species eco-evolutionary dynamics remains dispiritingly complicated (Rousset and Ronce 2004; Lehmann et al. 2006; Alizon and Taylor 2008; Wild 2011). As a result, the selection gradient is most often computed numerically as the derivative of an invasion fitness measure, without the provision of any biological interpretation of selection on the trait under focus (Metz and Gyllenberg 2001; Cadet et al. 2003; Parvinen et al. 2003; Parvinen and Metz 2008). Only specific demographic models under limited dispersal with finite patch size have been studied analytically in detail (Comins et al. 1980; Gandon and Michalakis 1999; Lehmann et al. 2006; Rodrigues and Gardner 2012; Parvinen et al. 2018). A biologically intuitive understanding of selection on traits that influence

metacommunity dynamics is therefore out of immediate reach when using the exact selection gradient.

To circumvent the difficulty of computing and interpreting the selection gradient under limited dispersal, various approximations have been proposed. The most prominent is perhaps the heuristic pair approximation, which has been used to study intraspecific social evolution and host-parasite coevolution in lattice-structured populations (e.g., Nakamaru et al. 1997; van Baalen and Rand 1998; Le Galliard et al. 2003, 2005; Nakamaru and Iwasa 2005; Lion and van Baalen 2007*a*; Lion and Gandon 2009, 2010; Débarre et al. 2012). However, more general multispecies coevolution scenarios have not been much investigated using pair approximation. This is presumably because analytical explorations remain complicated in lattice-structured populations due to the effects of isolation by distance.

In this article, we present a novel heuristic approximation for the selection gradient on traits that influence eco-evolutionary dynamics in patch-structured populations, which do not experience isolation by distance or heterogeneities in abiotic factors (i.e., the population is structured according to the homogeneous island model of dispersal of Wright 1931; for its ecological counterpart, see Chesson 1981). The crux of this approximation is that it assumes that local population size dynamics are deterministic with a stable fixed point (i.e., we ignore ecological stochasticity and periodic or chaotic population dynamics). This assumption allows us to reach key analytical insights, which can be applied to understand a wide spectrum of multispecies interactions. Importantly, our approximation provides a biologically meaningful interpretation of selection on traits that influence ecological interactions under limited dispersal. The rest of the article is organized as follows: (1) We describe a stochastic metacommunity eco-evolutionary model. (2) We motivate an approximation of our model that ignores ecological stochasticity. (3) Under this approximation, we analytically derive the selection gradient on a trait that influences eco-evolutionary dynamics through intra- and interspecific interactions. (4) We use our approximation to study two examples of social and ecological interactions: evolution of helping within species and harming between species when these behaviors influence demography. We show that for these examples our approximation performs well compared to individual-based simulations in predicting equilibrium strategies and the species abundances these strategies generate.

## Model

### *Metacommunity Structure*

We consider an infinite number of patches that are connected by uniform dispersal (Wright's [1931] infinite island model of dispersal). On each patch, a community of up to  $S$  species may coexist. The life cycle events of each species

$i \in \{1, 2, \dots, S\}$  are as follows: (1) Each adult produces a number of offspring that follows a Poisson distribution (with a mean that may depend on local interactions within and between species). (2) Each adult then either survives or dies (with a probability that may depend on local interactions within and between species). (3) Independently of one another, each offspring either remains in its natal patch or disperses to another randomly chosen one (the dispersal probability is assumed to be nonzero for all species but may differ among species). (4) Each offspring either dies or survives to adulthood (with a probability that may depend on local population numbers, e.g., if space on each patch is a limiting factor).

### *Evolving Phenotypes and the Uninvadable Species Coalition*

Each individual expresses a genetically determined evolving phenotype, or strategy, which can affect any event, such as reproduction, survival, or dispersal, in the life cycle of any species. We assume that the expression of a strategy and its effects are independent of age (i.e., no age structure). We denote by  $\Theta_i$  the set of feasible strategies for species  $i$  (this set is either the set or a subset of the real numbers,  $\Theta_i \subset \mathbb{R}$ ). When the population of each species is monomorphic, the metacommunity (i.e., the collection of subdivided populations of each species) is described by a vector of strategies  $\theta = (\theta_1, \theta_2, \dots, \theta_S)$ , where  $\theta_i$  is the strategy expressed by all individuals of species  $i$  (i.e.,  $\theta$  denotes a monomorphic resident population).

We define  $\theta$  as an uninvadable coalition if any mutation  $\tau_i \in \Theta_i$ , which arises in any species  $i$  and which results in a unilateral deviation  $\tau_i = (\theta_1, \dots, \theta_{i-1}, \tau_i, \theta_{i+1}, \dots, \theta_S)$  of the resident vector, goes extinct. The concept of an uninvadable coalition is the same as the concepts of a multispecies evolutionary stable strategy (Brown and Vincent 1987, p. 68) and an evolutionary stable coalition (Apaloo and Butler 2009, p. 640).<sup>1</sup>

### *Adaptive Dynamics*

Under the above definition of uninvadability, it is sufficient to consider the fate of a unilateral phenotypic deviation in one species at a time to determine whether a coalition is uninvadable. We can therefore focus our attention on the evolutionary dynamics of a mutant allele, which arises in species  $i$  and codes for phenotype  $\tau_i$ , when the resident community expresses  $\theta$ . In the infinite island model of dispersal, the change  $\Delta p_i$  in (appropriately weighted) average frequency  $p_i$  of such a mutant allele over one demographic time period (one life cycle iteration) can be written as

$$\Delta p_i = p_i(1 - p_i)\delta_i S_i(\theta) + O(\delta_i^2), \quad (1)$$

1. We here refrain from using the terminology “evolutionary stability,” as it presumes that such strategies are attractors of the evolutionary dynamics (Maynard Smith 1982), which is not covered by the concept of uninvadability.

where  $\delta_i = \tau_i - \theta_i$  is the phenotypic effect of the mutation (Rousset 2004, pp. 206–207; Rousset and Ronce 2004, p. 129; with notation here adapted to introduce species-specific allele frequency change). The function  $S_i(\theta)$ , which depends on quantities evaluated only in the resident community  $\theta$ , is the selection gradient on the trait in species  $i$ . When selection is weak (so that  $|\delta| \ll 1$  and terms  $O(\delta_i^2)$  can be neglected), the selection gradient  $S_i(\theta)$  and phenotypic effect  $\delta_i$  give the direction of selection on the mutant at any allele frequency: selection favors the invasion and fixation of the mutation when  $\delta_i S_i(\theta) > 0$  and conversely extinction when  $\delta_i S_i(\theta) < 0$ . The selection gradient thus captures directional selection.

The selection gradient  $S_i(\theta)$  is useful to derive necessary conditions for a coalition to be uninvadable. First, a necessary condition for a coalition  $\theta^*$  that lies in the interior of the set of feasible strategies to be uninvadable is that

$$S_i(\theta^*) = 0 \quad \forall i. \quad (2)$$

Such a coalition  $\theta^*$  is said to be singular. The allele frequency change equation (1) also informs us whether a singular coalition  $\theta^*$  will be approached by gradual evolution from its neighborhood under the constant but limited influx of mutations, that is, whether it is convergence stable. A singular coalition is convergence stable when the eigenvalues of the  $S \times S$  Jacobian matrix,  $\mathbf{J}(\theta^*)$ , with  $(i, j)$  entry,

$$(\mathbf{J}(\theta^*))_{ij} = \left. \frac{\partial S_i(\theta)}{\partial \theta_j} \right|_{\theta=\theta^*}, \quad (3)$$

all have negative real parts (e.g., Débarre et al. 2014, eq. [7a]).

When at most two alleles can ever segregate in a given species, a convergence-stable strategy is also locally uninvadable (Débarre and Otto 2016). In that case, the collection of selection gradients acting in each species is sufficient to establish whether a coalition is uninvadable. When more than two alleles can segregate at a locus, establishing local uninvadability requires looking into the second-order effects of selection (i.e., terms of  $O(\delta_i^2)$  in eq. [1]; Taylor 1989; Geritz et al. 1998; Rousset 2004; Leimar 2009). These effects, which capture disruptive selection, are difficult to analytically characterize under limited dispersal (e.g., Ajar 2003; Mullon et al. 2016; Parvinen et al. 2018). We will therefore focus on the effects of directional selection in this article.

### *Selection Gradient in Metacommunities*

The selection gradient,  $S_i(\theta)$ , in a given species  $i$  in the island model can be written as a change in the fitness of individuals experiencing different demographic states, weighted jointly by reproductive values and relatedness coefficients (for a single-species demography dynamics, see Rousset and Ronce 2004, eqq. [26], [27]; for arbitrary demographic states, see Lehmann et al. 2016, eqq. [E.27]–[E.29]). Reproductive

values reflect the fact that individuals residing in different demographic states contribute differently to the future gene pool (Rousset and Ronce 2004; Lehmann et al. 2016). Reproductive values thus capture the effects of ecological stochasticity on selection. Relatedness, meanwhile, captures the fact that individuals from the same species that reside in the same patch are more likely to carry alleles identical by descent than randomly sampled individuals in the population (for textbook treatments, see Nagylaki 1992 and Rousset 2004). The relatedness coefficients in the selection gradient thus reflect the consequences of genetic stochasticity on selection.

In spite of the insights brought by the exact selection gradient on ecological and genetic stochasticity, its usage to study community evolution under the assumptions of our model presents two caveats. The first is that owing to the large number of possible demographic states within patches (all possible configurations of the number of individuals of all species on a patch), the necessary computations are not straightforward and would be extremely expensive numerically (as shown by the computations necessary even in the absence of interspecific interactions; Rousset and Ronce 2004; Lehmann et al. 2006; Alizon and Taylor 2008; Wild et al. 2009; Wild 2011). It is possibly due to this computational hurdle that no application of the exact selection gradient to the coevolution of multiple species that experience stochastic demography and limited dispersal can be found in the literature. The second caveat is that the expression of selection in terms of reproductive values applies to any type of demographic structuring (e.g., by age, stage, sex, or environment; Frank 1998; Rousset 2004; Grafen 2006). As a consequence, without solving reproductive values explicitly in terms of model parameters, the exact selection gradient carries little biological information about how local ecological interactions influence selection.

The goal of this article is to provide a tractable and biologically informative approximation for the selection gradient in a metacommunity. We propose to achieve this by assuming that changes in local population size are deterministic with a stable fixed point when the metacommunity is monomorphic for  $\theta$ . Resident patches will therefore experience neither stochastic ecological fluctuations nor periodic/chaotic dynamics. As a consequence, it will no longer be necessary to consider all the possible demographic states that a focal patch can transit between. Before deriving this approximation, let us first study resident demographic dynamics (i.e., in a metacommunity monomorphic for the resident  $\theta$ ) to investigate how and when these dynamics can be assumed to be deterministic.

### Resident Community Dynamics

#### Deterministic Resident Community Dynamics

Our life cycle assumptions (see “Metacommunity Structure”) entail that we are considering an infinite stochastic

system, that is, we have an infinite number of interacting multidimensional Markov chains (each chain describes the community dynamics on a single patch, and chains interact with one another through dispersal; Chesson 1981, 1984). For this system, let us consider the expected number of individuals (or abundance) of each species in a focal patch at a given demographic time  $t$ , which we denote by  $\mathbf{n}_t(\theta) = (n_{1,t}(\theta), n_{2,t}(\theta), \dots, n_{S,t}(\theta))$ , where  $n_{i,t}(\theta)$  is the expected abundance of species  $i$  (note that we do not need to label patches since they are identical on average; e.g., Neuhauser 2002). This expected abundance can be written as  $n_{i,t}(\theta) = F_i(\theta, \mathbf{N}_{t-1}, \bar{\mathbf{n}}_{t-1}(\theta))$ , for  $i = 1, 2, \dots, S$ , where the population dynamic transition map  $F_i$  gives the expected number of individuals of species  $i$ , given the local community state  $\mathbf{N}_{t-1} = (N_{1,t-1}, N_{2,t-1}, \dots, N_{S,t-1})$  in the previous demographic time period (i.e.,  $N_{i,t-1}$  denotes the random variable for the number of individuals of species  $i$  in the focal patch at time  $t - 1$ ), and when the global average community state is  $\bar{\mathbf{n}}_{t-1}(\theta) = (\bar{n}_{1,t-1}(\theta), \bar{n}_{2,t-1}(\theta), \dots, \bar{n}_{S,t-1}(\theta))$  (i.e.,  $\bar{n}_{i,t-1}(\theta)$  is the average number of individuals of species  $i$  across all patches at  $t - 1$ ; note that  $\bar{\mathbf{n}}_{t-1}(\theta)$  is not a random variable because there is an infinite number of patches). We have written the transition map  $F_i(\theta, \mathbf{N}_{t-1}, \bar{\mathbf{n}}_{t-1}(\theta))$  such that it also explicitly depends on the vector of phenotypes  $\theta$  of each species in the focal patch in the previous time period, which will be useful when we introduce selection.

The basis of our approximation is to assume that the dynamics of the abundance of each species in a patch are deterministic (in an abuse of notation,  $\mathbf{N}_t \approx E[\mathbf{N}_t] = \mathbf{n}_t(\theta)$ ). This entails that the ecological dynamics on a focal patch, which are no longer stochastic, are given by

$$n_{i,t}(\theta) = F_i(\theta, \mathbf{n}_{t-1}(\theta), \bar{\mathbf{n}}_{t-1}(\theta)) \quad \text{for } i = 1, 2, \dots, S \quad (4)$$

(Chesson 1981). We further assume that these deterministic ecological dynamics are such that abundances of all species converge to a stable fixed point. From the dynamical equation (4), this ecological fixed point, which we denote as  $\hat{\mathbf{n}}(\theta) = (\hat{n}_1(\theta), \hat{n}_2(\theta), \dots, \hat{n}_S(\theta))$ , solves

$$\hat{n}_i(\theta) = F_i(\theta, \hat{\mathbf{n}}(\theta), \hat{\mathbf{n}}(\theta)) \quad \text{for } i = 1, 2, \dots, S. \quad (5)$$

Local stability of the fixed point  $\hat{\mathbf{n}}(\theta)$  entails that it is such that the local community matrix (e.g., Yodzis 1989; Case 2000)

$$\mathbf{C}(\theta) = \begin{pmatrix} \left. \frac{\partial F_1(\theta, \mathbf{n}, \hat{\mathbf{n}})}{\partial n_1} \right|_{\mathbf{n} = \hat{\mathbf{n}}} & \left. \frac{\partial F_1(\theta, \mathbf{n}, \hat{\mathbf{n}})}{\partial n_2} \right|_{\mathbf{n} = \hat{\mathbf{n}}} & \dots & \left. \frac{\partial F_1(\theta, \mathbf{n}, \hat{\mathbf{n}})}{\partial n_S} \right|_{\mathbf{n} = \hat{\mathbf{n}}} \\ \left. \frac{\partial F_2(\theta, \mathbf{n}, \hat{\mathbf{n}})}{\partial n_1} \right|_{\mathbf{n} = \hat{\mathbf{n}}} & \left. \frac{\partial F_2(\theta, \mathbf{n}, \hat{\mathbf{n}})}{\partial n_2} \right|_{\mathbf{n} = \hat{\mathbf{n}}} & \dots & \left. \frac{\partial F_2(\theta, \mathbf{n}, \hat{\mathbf{n}})}{\partial n_S} \right|_{\mathbf{n} = \hat{\mathbf{n}}} \\ \vdots & \vdots & \ddots & \vdots \\ \left. \frac{\partial F_S(\theta, \mathbf{n}, \hat{\mathbf{n}})}{\partial n_1} \right|_{\mathbf{n} = \hat{\mathbf{n}}} & \left. \frac{\partial F_S(\theta, \mathbf{n}, \hat{\mathbf{n}})}{\partial n_2} \right|_{\mathbf{n} = \hat{\mathbf{n}}} & \dots & \left. \frac{\partial F_S(\theta, \mathbf{n}, \hat{\mathbf{n}})}{\partial n_S} \right|_{\mathbf{n} = \hat{\mathbf{n}}} \end{pmatrix}, \quad (6)$$

whose  $(i, j)$  entry measures the influence of the abundance of species  $j$  on the abundance of species  $i$  over one demographic time period, has eigenvalues all with absolute value less than one.

#### Illustrating Example

To illustrate the transition function ( $F$ ; eq. [4]) and provide a basis for later comparisons with individual-based simulations, consider two semelparous species whose life cycle is as follows: (1) Each individual of species  $i \in \{1, 2\}$  in a patch with  $\mathbf{n} = (n_1, n_2)$  individuals produces a mean number  $f_i/(1 + \gamma n_i + \eta n_j)$  of offspring (with  $j \in \{1, 2\}$  and  $j \neq i$ ), where  $f_i$  is the number of offspring produced in the absence of density-dependent competition. The denominator  $1 + \gamma n_i + \eta n_j$  captures density-dependent competition within species (with intensity  $\gamma$ ) and between species (with intensity  $\eta$ ). This model for offspring production can be seen as a special case of the Leslie-Gower model of species interaction (Leslie and Gower 1958, eq. [1.1]). (2) All adults die. (3) Independently of one another, each offspring of species  $i$  disperses with probability  $m_i$ . (4) Finally, all offspring survive to adulthood.

According to this life cycle, the abundance of species 1 and 2 in the focal patch, conditional on the abundance being  $\mathbf{n}_{t-1}(\theta)$  in the previous time period in the focal patch and on the abundance in other patches being at a stable equilibrium  $\hat{\mathbf{n}}(\theta)$ , can be written as

$$\begin{aligned} n_{1,t}(\theta) &= (1 - m_1) \left( \frac{f_1}{1 + \gamma n_{1,t-1}(\theta) + \eta n_{2,t-1}(\theta)} \right) n_{1,t-1}(\theta) + m_1 \hat{n}_1(\theta), \\ n_{2,t}(\theta) &= (1 - m_2) \left( \frac{f_2}{1 + \gamma n_{2,t-1}(\theta) + \eta n_{1,t-1}(\theta)} \right) n_{2,t-1}(\theta) + m_2 \hat{n}_2(\theta). \end{aligned} \quad (7)$$

Equation (7) is an example of a so-called coupled map lattice with implicit space (e.g., Ranta et al. 2006, eq. [3.11]). The first summand in each line of equation (7) is the number of settled individuals in the focal patch that are born locally. The second summand in each line is the total number of offspring that immigrate into the focal patch from other patches. To better understand this second summand, consider that when the population is at the resident demographic equilibrium, an individual produces on average one offspring (i.e., individual fitness is equal to one), so that the total number of individuals remains constant. As a consequence,  $\hat{n}_1(\theta)$  and  $\hat{n}_2(\theta)$  gives the total number of offspring of species 1 and 2, respectively, produced in any patch other than the focal one. Therefore,  $m_1 \hat{n}_1(\theta)$  and  $m_2 \hat{n}_2(\theta)$  in equation (7) give the average number of offspring immigrating into the focal patch of species 1 and 2, respectively.

The equilibrium abundance of both species is found by substituting equation (7) in equation (5) (i.e., putting  $n_{1,t}(\theta) =$

$n_{1,t-1}(\theta) = \hat{n}_1(\theta)$  and  $n_{2,t}(\theta) = n_{2,t-1}(\theta) = \hat{n}_2(\theta)$ ) and solving for  $\hat{n}_1(\theta)$  and  $\hat{n}_2(\theta)$  simultaneously. Doing so, we find that the unique positive equilibrium (i.e.,  $\hat{n}_1(\theta) > 0$  and  $\hat{n}_2(\theta) > 0$ ) reads as

$$\begin{aligned} \hat{n}_1(\theta) &= \frac{(f_1 - 1)\gamma - (f_2 - 1)\eta}{\gamma^2 - \eta^2}, \\ \hat{n}_2(\theta) &= \frac{(f_2 - 1)\gamma - (f_1 - 1)\eta}{\gamma^2 - \eta^2}. \end{aligned} \quad (8)$$

This reveals that for the two species to coexist, it is necessary for intraspecific competition to be stronger than interspecific competition ( $\gamma > \eta$ ), which is a classical result (Case 2000).

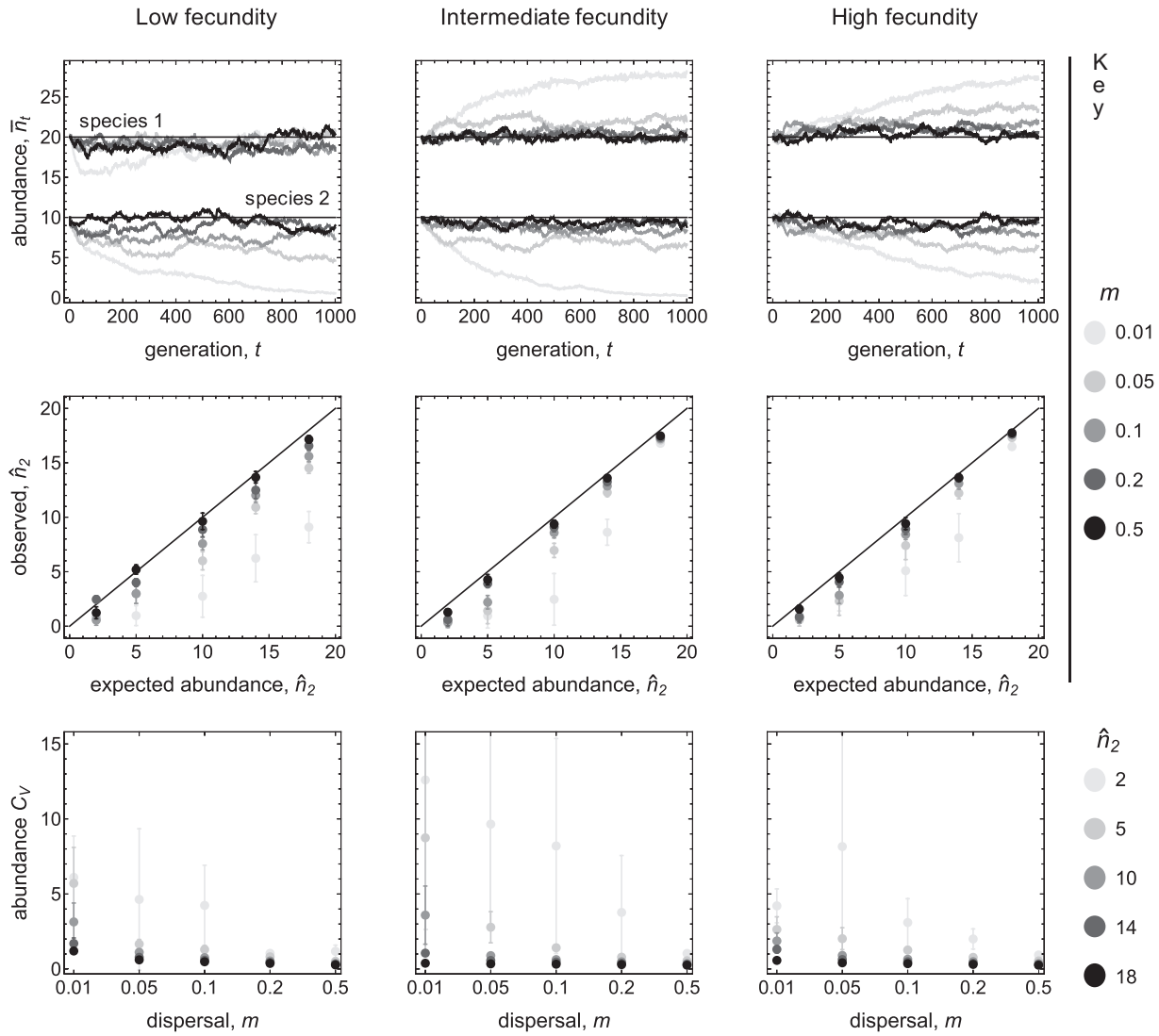
#### Comparing Deterministic and Stochastic Dynamics

To assess when ecological stochasticity can be ignored (i.e., when eq. [4] accurately reflects the true stochastic dynamics), we compared the deterministic community dynamics of the Leslie-Gower model (eq. [7]) with individual-based simulations of the full stochastic model (see app. A for a description of the simulation procedure; apps. A–C are available online).

*Ecological Stochasticity.* We find that there is a good qualitative match between the deterministic equilibrium abundance given by equation (8) and the average number of individuals of each species in a group observed in stochastic individual-based simulations (fig. 1, top row). As predicted by theory (Chesson 1981; Neuhauser 2002), the deterministic equilibrium systematically deviates from the observed average (fig. 1, top row). However, these deviations are small provided dispersal is not too weak (roughly no less than 0.1), and competition is such that deterministic abundance on a patch is greater or equal to 10 individuals (fig. 1, middle row). This suggests that under such demographic situations ecological stochasticity can be ignored.

When fecundity is low and competition is weak, one might expect that ecological stochasticity is too important to be ignored, as forces stabilizing the ecological equilibrium are weaker. Strikingly, even when fecundity is extremely low (barely above one, which is the threshold for the population to be maintained), it is still true that as long as dispersal is greater or equal to 0.1 and competition is such that deterministic abundance on a patch is greater or equal to 10 individuals, deterministic ecological dynamics are a very good approximation of the stochastic process (fig. 1, left column). The mitigating effects of dispersal on ecological stochasticity are further illustrated by the observation that in stochastic simulations variation in abundance among patches rapidly becomes vanishingly small as dispersal increases (fig. 1, bottom row).

Why dispersal mitigates the effects of ecological stochasticity can be understood as follows. Local population dynam-



**Figure 1:** Stochastic dynamics of local abundance and their deterministic approximation (eq. [7]) for the Leslie-Gower model with two species. Different columns correspond to different levels of baseline fecundity (i.e., fecundity in the absence of density of competition). The first column has extremely low baseline fecundity ( $f_1 = 1.1, f_2 = 1.095$ ), the second has intermediate fecundity ( $f_1 = 2, f_2 = 1.955$ ), and the third has high fecundity ( $f_1 = 5, f_2 = 4.955$ ). Top row: deterministic dynamics (straight line, determined from eq. [7], with  $\hat{n} = n_{t-1}$ ) and stochastic dynamics (dots, with 1,000 patches [for details on simulations, see app. sec. A1, available online], with  $m_1 = m_2 = 0.01, 0.05, 0.1, 0.2, 0.5$ —see key on right-hand side—and competition parameters  $\gamma$  and  $\eta$  chosen so that the deterministic equilibrium given by eq. [8] is  $\hat{n}_1 = 20$  and  $\hat{n}_2 = 10$ —see table 1 for values). Middle row: comparisons between the deterministic (X-axis, from eq. [8]) and stochastic (Y-axis) number of individuals of species 2, averaged over 1,000 generations starting at deterministic equilibrium in each patch (error bars give the standard deviation of the stochastic value, with  $m_1 = m_2 = 0.001, 0.01, 0.1, 0.5$ —see key on right-hand side—and competition parameters ( $\gamma, \eta$ ) chosen so that the deterministic equilibrium given by eq. [8] is  $\hat{n}_1 = 20$  and  $\hat{n}_2 = 2, 5, 10, 14, 18$ —in general, competition parameters decrease as fecundity decreases and patch size increases; see table 1 for the parameter values we used, which span two orders of magnitude, from 0.001 to 0.1). Departures from the diagonal indicate deviations between the exact process and the deterministic approximation. Bottom row: from the same simulations described in the middle row, these graphs show the coefficient of variation  $C_v$  of abundance of species 2 across patches according to dispersal level (i.e.,  $C_v$  is the ratio of the average number of individuals of species 2 per patch to its standard deviation, averaged over 1,000 generations, shown here for different values of  $\hat{n}_2$ —see key on right-hand side—error bars give the standard deviation of  $C_v$ , over 1,000 generations).

ics depend on the balance between two processes: (1) a local process at the patch level (i.e., dependence on  $N_{t-1}$ ), which has a strong stochastic component when patches have few individuals, and (2) a global process at the metacommunity

level (i.e., dependence on  $\bar{n}_{t-1}(\theta)$ ), which has a weak stochastic component when the number of patches is large (in fact, as the number of patches grows infinite, patches affect each other deterministically; Chesson 1981). As dispersal increases,

**Table 1:** Competition parameters used in simulations to generate figure 1

Equilibrium abundance of species 2— $\hat{n}_2$	Low fecundity ( $f_1, f_2$ ) = (1.1, 1.095)	Intermediate fecundity ( $f_1, f_2$ ) = (2, 1.955)	High fecundity ( $f_1, f_2$ ) = (5, 4.955)
2:			
$\gamma$	.00457071	.0456818	.182045
$\eta$	.00429293	.0431818	.179545
5:			
$\gamma$	.00406667	.0406	.1606
$\eta$	.00373333	.0376	.1576
10:			
$\gamma$	.0035	.0348333	.134833
$\eta$	.003	.0303333	.130333
14:			
$\gamma$	.00328431	.0325	.120735
$\eta$	.00245098	.025	.113235
18:			
$\gamma$	.00381579	.0369737	.115921
$\eta$	.00131579	.0144737	.0934211

Note: Values are found by solving eq. (8) for  $\gamma$  and  $\eta$  with  $\hat{n}_1 = 20$ ,  $\hat{n}_2$  given in the left column, and fecundities  $f_1$  and  $f_2$  given in the top row.

local population dynamics increasingly depend on the global process and less on the local one. As a result, local population dynamics become increasingly deterministic.

*Genetic Stochasticity.* The above analysis suggests that ecological stochasticity can be ignored when dispersal values are roughly above 0.1 and demographic parameters are such that equilibrium abundance within species is greater or equal to 10 (fig. 1). This raises the question of whether genetic stochasticity can also be ignored for such parameter values. The consequence of genetic stochasticity for selection can be ignored when relatedness coefficients are very small. The standard relatedness coefficient in the island model is the probability  $r_i(\theta)$  that two individuals from the same species  $i$ , which are randomly sampled in the same patch, carry an allele that is identical by descent when the population is monomorphic for  $\theta$  (also referred to as pairwise relatedness; Frank 1998; Rousset 2004). Let us consider this probability when the community has reached its (deterministic) demographic equilibrium  $\hat{n}(\theta)$  (given by eq. [8]). Owing to our assumption that fecundity is Poisson distributed, pairwise relatedness satisfies the relationship

$$r_i(\theta) = (1 - m_i)^2 \left( \frac{1}{\hat{n}_i(\theta)} + \left( 1 - \frac{1}{\hat{n}_i(\theta)} \right) r_i(\theta) \right), \quad (9)$$

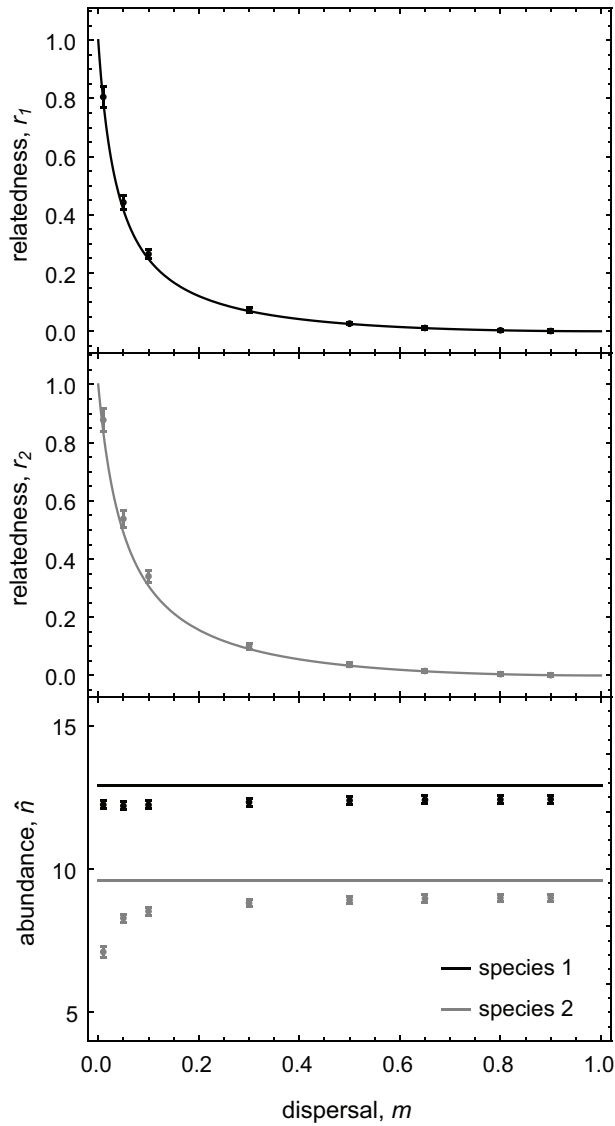
which can be understood as follows. With probability  $(1 - m_i)^2$ , two randomly sampled individuals of species  $i$  are both of philopatric origin (i.e., they were born in the focal patch). Then, with probability  $1/\hat{n}_i(\theta)$ , these individuals descend from the same parent, so their relatedness is one. With complementary probability  $1 - 1/\hat{n}_i(\theta)$ , they descend

from different parents, so their relatedness is  $r_i(\theta)$ . The solution to equation (9) is

$$r_i(\theta) = \frac{(1 - m_i)^2}{1 + [1 - (1 - m_i)^2](\hat{n}_i(\theta) - 1)}, \quad (10)$$

which is equivalent to the standard  $F_{ST}$  quantity (when individuals are sampled without replacement; e.g., Rousset 2004; Hartl and Clark 2007). Note, however, that in contrast to most mathematical treatments of  $F_{ST}$ , the number of individuals  $\hat{n}(\theta)$  here is endogenously determined by an explicit demographic process (eqq. [7], [8]).

Inspection of equation (10) reveals that relatedness can build up to significant values, even when dispersal is within a range under which we can legitimately approximate mean abundance by the deterministic model (e.g., local population size must be greater than 422 for relatedness to be less than 0.01 when dispersal is equal to 0.1). This shows that there exists a demographic regime under which ecological stochasticity can be neglected but genetic stochasticity cannot (which is in line with the fact that genetic stochasticity can lead to significant levels of relatedness even when patch size is constant and there is no ecological stochasticity; Rousset 2004; Hartl and Clark 2007). We will therefore take into account the effects of genetic stochasticity when deriving our approximation for the selection gradient. It is noteworthy that we find an excellent match between pairwise relatedness observed in individual-based simulations and pairwise relatedness calculated from the deterministic ecological approximation (i.e., eq. [10] with eq. [8]; fig. 2). This lends further support to the usefulness of the deterministic ecological approximation to study populations at ecological equilibrium (eq. [5]).



**Figure 2:** Relatedness and average local abundance under stochastic individual-based simulations and their deterministic approximation. The top two graphs show the relatedness for species 1 and 2, respectively, under the Leslie-Gower demographic model (see “Illustrating Example”). Relatedness values obtained from the deterministic approximation are shown in solid lines (i.e., obtained from eq. [10] with the ecological equilibrium derived from the deterministic approximation, eq. [8], parameter values  $f_1 = 2, f_2 = 1.8, \gamma = 0.07,$  and  $\eta = 0.01$ ). Relatedness computed from individual-based simulations are shown as points, with standard deviation shown by error bars (time average of population mean over 5,000 generations after burn-in of 5,000 generations with 1,000 patches, with  $m = 0.01, 0.05, 0.1, 0.3, 0.5, 0.65, 0.8, 0.9$ ; for details on calculations of relatedness, see app. sec. A1, available online). The bottom graph displays the average local abundance of species 1 (in black) and 2 (in gray). Full lines are for the deterministic approximation (eq. [8]). Points are results obtained from stochastic individual-based simulations (time average of population mean over 5,000 generations after burn-in of 5,000 generations, with  $m = 0.01, 0.05, 0.1, 0.3, 0.5, 0.8, 0.9$ ). Error bars show standard deviation.

### Evolutionary Analysis

We now specify the (approximate) selection gradient on a trait expressed in species  $i$  when the effects of ecological stochasticity are neglected. First, we characterize ecological dynamics when they can be influenced by the presence of genetic mutants.

### Mutant Community Dynamics

We now assume that two alleles segregate in the focal species  $i$ : a rare mutant that codes for phenotype  $\tau_i$  and a resident for  $\theta_i$ . We focus on a focal patch in which both alleles are present, while other patches are considered to be monomorphic for the resident  $\theta$ , and at their ecological equilibrium,  $\hat{n}(\theta)$  (eq. [5]). In this focal patch, we assume that the number of individuals,  $n_{j,t}(\tau_i)$ , of species  $j$  at time  $t$  is given by

$$n_{j,t}(\tau_i) = F_j(\bar{\tau}_{i,t-1}, \mathbf{n}_{t-1}(\tau_i), \hat{n}(\theta)) \quad \text{for } j = 1, 2, \dots, S, \quad (11)$$

where  $F_j$  is the map introduced in “Deterministic Resident Community Dynamics” (eq. [4]) but the first argument of this map is now  $\bar{\tau}_{i,t-1} = (\theta_1, \dots, \theta_{i-1}, \bar{\tau}_{i,t-1}, \theta_{i+1}, \dots, \theta_S)$ , which is a vector collecting the average phenotypes expressed in each species in the focal patch at demographic time  $t - 1$  (in species  $j \neq i$  other than the focal, this average is simply the resident  $\theta_j$ ; in the focal species  $i$ , this average is denoted by  $\bar{\tau}_{i,t}$ ). Since the average phenotype in the focal species,  $\bar{\tau}_{i,t}$ , depends on the number of genetic mutants, the first argument of  $F_j$  in equation (11) captures the effect of the genetic state on local abundance. The dependence of  $F_j$  on the average phenotype approximates possibly more complicated relationships between genetic state and abundance to the first order, which is sufficient to evaluate the selection gradient (Rousset 2004, p. 95). The map  $F_j$  also depends on  $\mathbf{n}_{t-1}(\tau_i) = (n_{1,t-1}(\tau_i), \dots, n_{S,t-1}(\tau_i))$ , which is the ecological state of the focal patch at time  $t - 1$ , and on the equilibrium  $\hat{n}(\theta)$ , which is the ecological state of other patches.

Because we take genetic stochasticity into account, the number of mutants, and hence the average phenotype in the focal species,  $\bar{\tau}_{i,t}$ , fluctuate randomly and should be considered as random variables. As a result, the abundance  $n_{j,t}(\tau_i)$  at time  $t$  given by equation (11) is also a random variable. Importantly, this stochasticity in abundance is due only to genetic stochasticity in our approximation. When the average phenotype  $\bar{\tau}_{i,t}$  is fixed (e.g., for the resident,  $\bar{\tau}_{i,t} = \theta_i$ ), ecological dynamics are fully deterministic and given by the recurrence equation (11). In other words, our approximation ignores the influence of ecological stochasticity on ecological dynamics (for further details, see app. sec. B3).

### Inclusive Fitness Effect for the Interactive Community

To derive our approximation for the selection gradient, we use the basic reproductive number as an invasion fitness proxy



(e.g., Stearns 1992; Charlesworth 1994; Case 2000; Metz and Gyllenberg 2001; Lehmann et al. 2016). This allows us to drastically simplify our calculations and equivalently characterize directional selection (i.e., the first-order effects of selection on allele frequency change). These points are further detailed in appendix B, where we show that the selection gradient on an evolving trait in species  $i$  in a  $\theta$  community can be approximated as

$$S_i(\theta) \approx \epsilon(\theta)s_i(\theta), \tag{12}$$

where  $\epsilon(\theta) > 0$  is a factor of proportionality that depends on  $\theta$  only (see eqq. [B3]–[B19]), and

$$s_i(\theta) = s_{w,i}(\theta) + s_{e,i}(\theta) \tag{13}$$

is the approximate selection gradient. Since  $\epsilon(\theta) > 0$ ,  $s_i(\theta)$  is sufficient to ascertain singular trait values and their convergence stability (by replacing  $S_i(\theta)$  with  $s_i(\theta)$  in eqq. [2] and [3]). The approximate selection gradient,  $s_i(\theta)$  (eq. [13]), consists of the sum of two terms. The first term,  $s_{w,i}(\theta)$ , captures selection owing to the trait’s intratemporal effects (effects within a demographic period). The second term,  $s_{e,i}(\theta)$ , captures selection owing to the trait’s intertemporal effects (effects between demographic periods), which emerge as a result of ecological inheritance (i.e., modified environmental conditions passed down to descendants; Odling-Smee et al. 2003; Bonduriansky 2012). We detail the two components of selection,  $s_{w,i}(\theta)$  and  $s_{e,i}(\theta)$ , in the next two sections.

*Selection on Intratemporal Effects.* The first term of equation (13) can be expressed as

$$s_{w,i}(\theta) = \left. \frac{\partial w_i(\tau_{\cdot,i}, \tilde{\tau}_i, \mathbf{n}, \hat{\mathbf{n}})}{\partial \tau_{\cdot,i}} \right|_{\substack{\tau_i = \theta_i \\ \mathbf{n} = \hat{\mathbf{n}}}} + \left. \frac{\partial w_i(\tau_{\cdot,i}, \tilde{\tau}_i, \mathbf{n}, \hat{\mathbf{n}})}{\partial \tilde{\tau}_i} \right|_{\substack{\tau_i = \theta_i \\ \mathbf{n} = \hat{\mathbf{n}}}} \cdot r_i(\theta) \tag{14}$$

(for derivation, see app. sec. B3.4), where  $w_i$  is the individual fitness of a focal individual of species  $i$  (i.e., the expected number of successful offspring produced over one life cycle iteration by the focal, including itself if it survives). Individual fitness,  $w_i(\tau_{\cdot,i}, \tilde{\tau}_i, \mathbf{n}, \hat{\mathbf{n}})$ , is written as a function of four variables: (1) the phenotype  $\tau_{\cdot,i}$  of the focal individual, (2) the vector  $\tilde{\tau}_i = (\theta_1, \dots, \theta_{i-1}, \tilde{\tau}_i, \theta_{i+1}, \dots, \theta_s)$  of average phenotypes of neighbors in the focal patch (where  $\tilde{\tau}_i$  is the average phenotype among the neighbors of species  $i$  of the focal individual), (3) the vector of abundances in the focal community  $\mathbf{n}$ , and (4) the vector of average abundance across the metacommunity, which is at its equilibrium  $\hat{\mathbf{n}}(\theta)$  (explicit examples of such a fitness function are given later when we apply our method; see eqq. [23], [36], and [C1]). Note that individual fitness may also depend on the phenotype expressed in patches other than the focal, which is the resident  $\theta$ , but we have chosen not to write this dependency explicitly.

The two derivatives in equation (14), which are evaluated in the resident population (i.e., with resident phenotype  $\tau_i = \theta_i$  and resident ecological equilibrium  $\mathbf{n} = \hat{\mathbf{n}}(\theta)$ ), capture different fitness effects of the trait. The first derivative represents the change in the fitness of a focal individual of species  $i$  resulting from this individual switching from the resident to the mutant phenotype (i.e., the direct effect of the trait). The second derivative can be interpreted as the change in the fitness of the whole set of same-species patch neighbors resulting from the focal individual switching from the resident to the mutant phenotype (i.e., the indirect effect of the trait). This second derivative is weighted by the neutral relatedness coefficient,  $r_i(\theta)$ , which gives the probability that any same-species neighbor also carries the mutation in the monomorphic resident.

*Selection on Intertemporal Feedback Effects due to Ecological Inheritance.* We now detail the second term of the selection gradient,  $s_{e,i}(\theta)$ , which captures selection on a trait’s intertemporal effects.

*Feedback between local ecology and evolution.* We find that the second term of equation (13) can be written as

$$s_{e,i}(\theta) = \sum_{j=1}^s \left. \frac{\partial \bar{n}_j}{\partial \tau_i} \right|_{\substack{\tau_i = \theta_i \\ \mathbf{n} = \hat{\mathbf{n}}}} \cdot \left. \frac{\partial w_i(\tau_{\cdot,i}, \tilde{\tau}_i, \mathbf{n}, \hat{\mathbf{n}})}{\partial n_j} \right|_{\substack{\tau_i = \theta_i \\ \mathbf{n} = \hat{\mathbf{n}}}} \tag{15}$$

(for details, see app. sec. B3.5, eq. [B28]), where  $\partial \bar{n}_j / \partial \tau_i$  is the effect of the mutation on the local abundance of species  $j$  experienced by a mutant of species  $i$  that is randomly sampled from its local lineage (i.e., the lineage of carriers of the mutant trait  $\tau_i$  that reside in the focal patch in which the mutation first appeared). The second derivative in equation (15) is the effect that this abundance change of species  $j$  has on the fitness of a focal individual of species  $i$ . By multiplying these two effects and summing them over all species  $j$  of the community, equation (15) therefore captures how selection depends on the feedback between local community ecology and evolution.

*Lineage-centered perspective on the ecological influence of a trait.* The feedback effect captured by equation (15) reveals that a phenotypic change will be selected when such a change results in local ecological conditions that are favorable for the lineage of those that express the change (i.e., when  $\partial \bar{n}_j / \partial \tau_i \cdot \partial w_i(\tau_{\cdot,i}, \tilde{\tau}_i, \mathbf{n}, \hat{\mathbf{n}}) / \partial n_j > 0$ ). This brings us to the question of what is the nature of the influence of a local lineage on its own ecology, which is captured by the derivative  $\partial \bar{n}_j / \partial \tau_i$ . We find that this derivative can be expressed as

$$\left. \frac{\partial \bar{n}_j}{\partial \tau_i} \right|_{\substack{\tau_i = \theta_i \\ \mathbf{n} = \hat{\mathbf{n}}}} = \sum_{k=1}^s \Psi_{k,i}(\theta) \sum_{h=1}^{\infty} K_{j,k,h}(\theta) [\hat{n}_i(\theta) r_{i,h}(\theta)] \tag{16}$$

(for details, see app. sec. B3.5, eq. [B36]). To understand equation (16), consider a focal individual from species  $i$  ex-

pressing the mutant allele who lives at a demographic time that we arbitrarily label as time zero ( $t = 0$ ; see star in fig. 3). The first term of equation (16),

$$\Psi_{k,i}(\theta) = \frac{1}{\hat{n}_i(\theta)} \left. \frac{\partial F_k(\bar{\tau}_i, \mathbf{n}, \hat{\mathbf{n}})}{\partial \bar{\tau}_i} \right|_{\substack{\tau_i = \theta_i \\ \mathbf{n} = \hat{\mathbf{n}}}} \quad (17)$$

is the effect that a trait change in the focal individual has on the abundance of species  $k$  in the focal patch at the next demographic time, that is, at  $t = 1$  (solid gray arrows in fig. 3).

The second term in equation (16),  $K_{jk,h}(\theta)$ , is given by the  $(j, k)$  element of the matrix

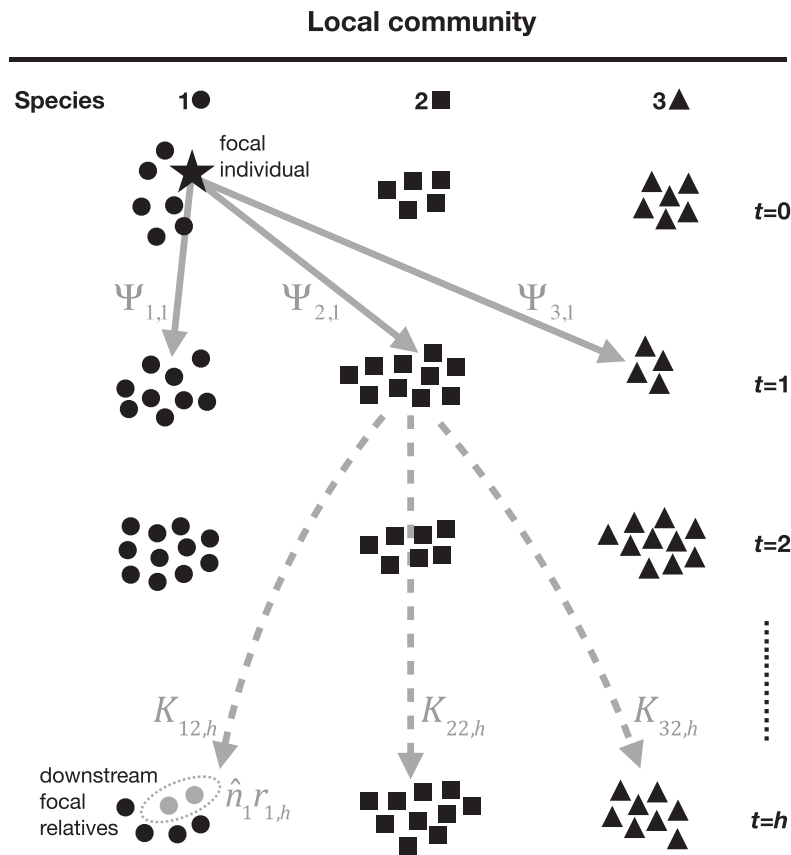
$$\mathbf{K}_h(\theta) = \mathbf{C}(\theta)^{h-1}, \quad (18)$$

where  $\mathbf{C}(\theta)$  is the community matrix given in equation (6). Equation (18) reveals that  $K_{jk,h}(\theta)$  in equation (16) is the effect that a change in the abundance of species  $k$  at time

$t = 1$  has on the abundance of species  $j$  at time  $t = h$ . Importantly, this effect takes into account the influence that species have on one another's abundance, cumulated over  $h - 1$  demographic time periods (as indicated by the exponent  $h - 1$  in eq. [18]; e.g., dashed gray arrows in fig. 3).

Finally, the term in square brackets in equation (16) can be interpreted as the expected number of carriers of identical-by-descent copies of the mutant allele (number of "relatives" of the focal individual) that live in the focal patch at time  $t = h \geq 1$  in the future (and that therefore experience the mutant-modified ecological conditions at time  $h$ ; e.g., gray disks in fig. 3). Indeed, this term consists of the product of the equilibrium abundance,  $\hat{n}_i(\theta)$ , with

$$r_{i,h}(\theta) = (1 - m_i(\theta))^h \underbrace{\left( \frac{1}{\hat{n}_i(\theta)} + \frac{\hat{n}_i(\theta) - 1}{\hat{n}_i(\theta)} r_i(\theta) \right)}_{\bar{r}_i(\theta)}, \quad (19)$$



**Figure 3:** Intertemporal effects of a focal individual. As an example, consider a local community of three species, labeled 1 (circle), 2 (square), and 3 (triangle). We consider the effect of a mutation in species 1. A focal carrier of a mutation  $\tau_1$  in species 1 living at time  $t = 0$  (denoted by a star) first directly influences the population dynamics of species 1, 2, and 3 at time  $t = 1$  according to  $\Psi_{1,1}(\theta)$ ,  $\Psi_{2,1}(\theta)$ , and  $\Psi_{3,1}(\theta)$  (solid gray arrows; eq. [17]). This change in abundance at time  $t = 1$  affects the abundance of the other species through time due to ecological interactions. For example, a change in abundance of species 2 at time  $t = 1$  influences the abundance of species 1, 2, and 3 at time  $t = h$  according to  $K_{12,h}(\theta)$ ,  $K_{22,h}(\theta)$ , and  $K_{32,h}(\theta)$ , respectively (dashed gray arrows; eq. [18]). These changes are experienced by  $\hat{n}_1(\theta)r_{1,h}(\theta)$  relatives of the focal individual at time  $t = h$  (gray circles; eq. [19]).

which is the relatedness between two individuals of species  $i$  that are sampled  $h$  demographic time periods apart in the focal patch in the resident population (i.e., the probability that these two individuals share a common ancestor that resided in the focal patch; for a formal definition, see also eq. [B35]). Note that in general  $m_i(\theta)$  is the backward probability of dispersal, which is defined as the probability that a randomly sampled individual of species  $i$  in the resident population is an immigrant. This will of course be influenced by dispersal behavior but also by other organismal aspects depending on the life cycle (e.g., on adult survival from one time period to the next).

The above considerations (eqq. [17], [18]) show that the influence of a local lineage on the abundance of its own or another species  $j$  (eq. [16]) can be intuitively understood as the effect of a trait change in a focal individual on the abundance of species  $j$ , which is experienced by all its downstream relatives residing in the focal patch (for a diagram, see fig. 3; for mathematical details, see app. sec. B3.5).

*Evolutionary press perturbations.* To evaluate equation (16) explicitly, we can use the fact that under our assumptions that patches are not totally isolated from one another (i.e.,  $m_i(\theta) > 0$ ) and that the resident community is at a stable fixed point (i.e.,  $\mathbf{C}(\theta)$  has eigenvalues with absolute value less than one), the infinite sum in equation (16) converges. This leads to the following expression (for details, see app. sec. B3.5, eq. [B37]):

$$\left. \frac{\partial \bar{n}_j}{\partial \tau_i} \right|_{\substack{\tau_i = \theta_i \\ \mathbf{n} = \hat{\mathbf{n}}}} = \sum_{k=1}^s \Psi_{k,i}(\theta) \hat{n}_i(\theta) \Lambda_{j,k,i}(\theta) (1 - m_i(\theta)) \bar{r}_i(\theta), \quad (20)$$

where  $\bar{r}_i(\theta)$  is the relatedness between two individuals sampled with replacement in the same patch (for definition, see eq. [19]) and  $\Lambda_{j,k,i}(\theta)$  is given by the  $(j, k)$  entry of the matrix

$$\Lambda_i(\theta) = (\mathbf{I} - [1 - m_i(\theta)]\mathbf{C}(\theta))^{-1}. \quad (21)$$

The term  $\Lambda_{j,k,i}(\theta)$  in equation (20) captures the effect of a change in the abundance of species  $k$  on the abundance of species  $j$ , experienced by all individuals of species  $i$  descending from a single ancestor in the focal patch. Interestingly, as dispersal goes to zero in the focal species ( $m_i(\theta) \rightarrow 0$ ), the matrix  $\Lambda_i(\theta)$  (eq. [21]) tends to the matrix of press perturbations of community ecology (i.e.,  $\Lambda_i(\theta) \rightarrow (\mathbf{I} - \mathbf{C}(\theta))^{-1}$ ). The entries of this matrix measure how a constant and persistent change in the abundance of one species influences the equilibrium abundance of another through multispecies interactions (e.g., Yodzis 1989; Case 2000). The correspondence between equation (21) and press perturbation matrices reflects that as  $m_i(\theta) \rightarrow 0$ , the mutant lineage may persist locally forever and thus experience persistent changes in the abundance of other species. But as dispersal  $m_i(\theta)$  increases, the mutant lineage will spend fewer time periods locally,

which means that its experience of changes in local species abundance will last fewer time periods (and so  $\Lambda_i(\theta)$  approaches the identity matrix as dispersal becomes complete, i.e.,  $\Lambda_i(\theta) \rightarrow \mathbf{I}$  as  $m_i(\theta) \rightarrow 1$ ).

*Connections with Previous Results on Selection Gradients and Ecological Feedback.* The selection gradient we have derived is closely connected to existing gradients in the literature. To see these connections, consider first the case when dispersal is complete ( $m_i(\theta) = 1$  so that  $r_i(\theta) = 0$ ). In this case, the selection gradient reduces to

$$s_i(\theta) = \left. \frac{\partial w_i(\tau_{\cdot,i}, \bar{r}_i, \mathbf{n}, \hat{\mathbf{n}})}{\partial \tau_{\cdot,i}} \right|_{\substack{\tau_i = \theta_i \\ \mathbf{n} = \hat{\mathbf{n}}}} \quad (22)$$

which embodies the classical ecological feedback considered in evolutionary analyses (e.g., Michod 1979; Charlesworth 1994; and see in particular eq. [29] of Lion 2017): the invasion of a rare mutant depends on resident-set ecological conditions only (i.e., on  $\theta$  and  $\hat{\mathbf{n}}(\theta)$  only), and if the mutant invades it becomes the resident and thereby modifies these conditions. The simplicity of equation (22) reflects the fact that when dispersal is complete a globally rare mutant is also always locally rare. As a consequence, the selection gradient depends only on the effect that a mutant carrier has on its own individual fitness.

When dispersal is limited ( $m_i(\theta) < 1$ ), however, a globally rare mutant may become locally common and remain so over multiple demographic time periods. This has two implications that are important for the way selection targets this mutant. First, mutants living in the same time period interact directly with one another. This effect is captured by the relatedness-weighted fitness effect of neighbors in  $s_{w,i}(\theta)$  (i.e., the second summand of eq. [14]). In fact,  $s_{w,i}(\theta)$  (eq. [14]) is equivalent to the standard selection gradient in the island model with constant demography (Taylor and Frank 1996; Frank 1998; Rousset 2004). But in contrast to the selection gradient under constant demography, abundance in  $s_{w,i}(\theta)$  (eq. [14]) is endogenously determined and evaluated at the resident ecological equilibrium,  $\hat{\mathbf{n}}(\theta)$ . As such,  $s_{w,i}(\theta)$  approximates the exact selection gradient in a demographically structured population when the trait has no demographic effect (denoted as  $S_i$ ; Rousset and Ronce 2004, eq. [26]; Lehmann et al. 2016, eq. [E-28]). The main difference between the approximation and the exact expression is that the latter depends on reproductive values while the approximation  $s_{w,i}(\theta)$  (eq. [14]) does not. This is because we ignore stochastic demographic fluctuations here and thus do not need to consider fitness effects in all possible demographic states.

The second implication of limited dispersal for the way selection targets a mutant is that a rare mutant can modify the demographic/ecological conditions experienced by its

own lineage. Put differently, mutants living at different time periods interact indirectly through heritable ecological modifications. Selection due to these indirect interactions is captured by the second term of the selection gradient,  $s_{e,i}(\theta)$  (eq. [13]). With constant demography in the focal evolving species, the ecological inheritance term  $s_{e,i}(\theta)$  is consistent with the selection gradient on intertemporal within-species altruism (Lehmann 2007, eq. [9]; Sozou 2009, eq. [4.11]) and niche construction traits affecting the abundance of a local abiotic resource (Lehmann 2008, eq. [A39] in the absence of isolation by distance). With fluctuating demography,  $s_{e,i}(\theta)$  (eq. [13]) approximates the part of the exact selection gradient that captures selection on a trait due to its demographic effect on the focal species only (sometimes denoted as  $S_p$ ; Rousset and Ronce 2004, eq. [27]; Lehmann et al. 2016, eq. [E-29]).

### Summary

In summary, we have shown that the selection gradient on a trait,  $s_i(\theta)$ , depends on how a trait change in a focal individual affects (1) its own fitness and the fitness of its current relatives through intratemporal interactions ( $s_{w,i}(\theta)$ ; eq. [15]) and (2) the fitness of its downstream relatives living in the focal patch through heritable modifications of the ecological environment ( $s_{e,i}(\theta)$ ; eqq. [16], [17]). This reveals that under limited dispersal, selection on intra- and interspecific interactions can generally be interpreted in terms of intertemporal inclusive fitness effects, that is, in terms of the effect that a trait change in a focal individual has on the fitness of this focal and of all its relatives (current and downstream). Such a perspective allows for an intuitive understanding of selection on ecological interactions that take place in dispersal-limited communities. In particular, our approximation highlights the nature of intertemporal effects and their roles in the molding of functional traits. We illustrate more concretely the potential importance of intertemporal effects when we apply our approximation to specific models in the next section.

### Applications

Here we use our approximation to study the evolution of two traits that underlie intra- and interspecific interactions under limited dispersal. The first is the evolution of helping within species, which has received considerable attention. This will allow us to contextualize our approach to study intraspecific interactions when such interactions influence demography. The second example is the evolution of harming between species, which has so far not been investigated under limited dispersal. Analytical calculation checks of these examples, as well as the codes for the associated individual-based simulations, are available in a zip

file containing a Mathematica Notebook (available online).<sup>2</sup>

### Helping within Species

**Biological Scenario.** We focus on a single species and study the evolution of a social trait or behavior that increases the fitness of patch neighbors but comes at a fitness cost to self. We consider the following life cycle: (1) Adults reproduce. A focal individual has mean fecundity  $f(\tau, \bar{\tau})/(1 + \gamma n)$ , where  $f(\tau, \bar{\tau})$  is its fecundity in the absence of density-dependent competition. The latter has intensity  $\gamma$ . Maximal fecundity  $f(\tau, \bar{\tau})$  decreases with the level  $\tau$ , of helping of the focal individual but increases with the average level  $\bar{\tau}$  of helping among its neighbors in the focal patch ( $\partial f(\tau, \bar{\tau})/\partial \tau < 0$  and  $\partial f(\tau, \bar{\tau})/\partial \bar{\tau} > 0$ ). (2) All the adults die. (3) Each offspring independently disperses with a probability  $m$ . (4) All offspring survive to adulthood (i.e., no competition for space among offspring).

Our assumptions for the life cycle can be biologically interpreted as individuals competing locally to acquire material resources, and the transformation of these resources into offspring depends on the level of helping within the patch (e.g., because individuals share resources).

**Necessary Components.** We first specify the components necessary to compute the selection gradient (i.e., the terms that appear in eqq. [13]–[20]). According to the life cycle assumptions for the model of helping, the fitness of a focal individual that expresses a level of helping  $\tau$ , in a patch of size  $n$  when its average neighbor expresses level  $\bar{\tau}$  is

$$w(\tau, \bar{\tau}, n, \hat{n}(\theta)) = \frac{f(\tau, \bar{\tau})}{1 + \gamma n}. \quad (23)$$

Note that here fitness does not depend on species abundance in patches other than the focal (i.e., it does not depend on  $\hat{n}(\theta)$ ). This is because we have assumed that competition occurs locally for material resources (for an example of a fitness function that depends on  $\hat{n}(\theta)$ , see eq. [C1]). Following the same argument used to derive equation (7), we find that the population dynamic (i.e., the abundance in the focal patch after one iteration of the life cycle, given that the average level of helping in the patch is  $\bar{\tau}$ , that abundance at the previous time period was  $n$ , and that other patches are at equilibrium  $\hat{n}(\theta)$ ) can be written as

$$F(\bar{\tau}, n, \hat{n}(\theta)) = (1 - m) \left( \frac{f(\bar{\tau}, \bar{\tau})}{1 + \gamma n} \right) n + m \hat{n}(\theta). \quad (24)$$

2. Code that appears in *The American Naturalist* is provided as a convenience to the readers. It has not necessarily been tested as part of the peer review.

The equilibrium population size  $\hat{n}(\theta)$  in the resident population is found by solving  $\hat{n} = F(\theta, \hat{n}, \hat{n})$  for  $\hat{n}$ , which yields

$$\hat{n}(\theta) = \frac{f(\theta, \theta) - 1}{\gamma}. \tag{25}$$

This equilibrium population size further allows us to obtain the pairwise relatedness  $r(\theta)$ , which is given by substituting equation (25) into equation (10), that is,

$$r(\theta) = \frac{(1 - m)^2}{1 + [1 - (1 - m)^2][(f(\theta, \theta) - 1)/\gamma - 1]}. \tag{26}$$

This shows that pairwise relatedness increases as intraspecific competition  $\gamma$  increases because this leads to smaller patch size (eq. [25]). As expected, relatedness increases as dispersal becomes limited ( $m \rightarrow 0$ ). From here it is straightforward to obtain the other necessary relatedness coefficient,  $\bar{r}(\theta)$  (for definition, see eq. [19]).

*Selection on Helping.* We now proceed to calculate the selection gradient on helping under our scenario. Note that the selection gradient on a single trait in a single species can be written as  $s(\theta) = s_w(\theta) + s_c(\theta)$ , where  $s_w(\theta)$  captures the intratemporal effects and  $s_c(\theta)$  the intertemporal effects.

*Intratemporal effects of helping.* Let us first study selection on helping according to its intratemporal effects (i.e., by looking at  $s_w(\theta)$ ; eq. [14]). These effects can be expressed as

$$\begin{aligned} s_w(\theta) &= \frac{\partial w(\tau, \bar{\tau}, n, \hat{n})}{\partial \tau} + \frac{\partial w(\tau, \bar{\tau}, n, \hat{n})}{\partial \bar{\tau}} r(\theta) \\ &= \frac{1}{f(\theta, \theta)} \left( \frac{\partial f(\tau, \bar{\tau})}{\partial \tau} + \frac{\partial f(\tau, \bar{\tau})}{\partial \bar{\tau}} r(\theta) \right), \end{aligned} \tag{27}$$

where we used equations (23) and (25). Note that since helping is individually costly but increases the fecundity of neighbors, the direct and indirect fitness effects of helping are negative and positive, respectively (i.e.,  $\partial w/\partial \tau < 0$  and  $\partial w/\partial \bar{\tau} > 0$ ). Hence, the helping trait in our model is altruistic sensu evolutionary biology (e.g., Hamilton 1964; Rousset 2004; West and Gardner 2010). Equation (27) shows that altruistic helping is favored by high relatedness. From the relatedness equation (26), we therefore expect limited dispersal and intraspecific competition to favor the evolution of helping, owing to its intratemporal effects. However, selection on helping also depends on its intertemporal effects, which we investigate in the next paragraph.

*Intertemporal effects of helping.* When a single species is under scrutiny, selection on intertemporal effects (i.e.,  $s_c(\theta)$ ; eq. [15]) can be expressed as

$$s_c(\theta) = \frac{\partial \bar{n}}{\partial \tau} \cdot \frac{\partial w(\tau, \bar{\tau}, n, \hat{n})}{\partial n}. \tag{28}$$

Using equation (20), the effect of helping on the lineage-experienced equilibrium abundance can be written as

$$\begin{aligned} \frac{\partial \bar{n}}{\partial \tau} &= \frac{\partial F(\bar{\tau}, n, \hat{n})}{\partial \bar{\tau}} \left[ 1 - (1 - m) \frac{\partial F(\bar{\tau}, n, \hat{n})}{\partial n} \right]^{-1} (1 - m) \bar{r}(\theta) \\ &= r(\theta) \frac{[f(\theta, \theta) - 1]/\gamma}{f(\theta, \theta) - (1 - m)^2} \left( \frac{\partial f(\tau, \bar{\tau})}{\partial \tau} + \frac{\partial f(\tau, \bar{\tau})}{\partial \bar{\tau}} \right). \end{aligned} \tag{29}$$

We will assume that  $\partial f(\tau, \bar{\tau})/\partial \tau + \partial f(\tau, \bar{\tau})/\partial \bar{\tau} > 0$ , so that helping increases equilibrium abundance (i.e.,  $\partial \bar{n}/\partial \tau > 0$ ). In turn, this increase in abundance feeds back negatively on the fitness of downstream individuals according to

$$\frac{\partial w(\tau, \bar{\tau}, n, \hat{n})}{\partial n} = -\frac{\gamma}{f(\theta, \theta)} < 0 \tag{30}$$

(using eq. [23]). This is because greater abundance leads to stronger intraspecific competition (according to  $\gamma$ ). As a result, the selective intertemporal fitness effects of helping,

$$s_c(\theta) = -\frac{r(\theta)}{f(\theta, \theta)} \underbrace{\left( \frac{f(\theta, \theta) - 1}{f(\theta, \theta) - (1 - m)^2} \right)}_{=\beta(\theta) > 0} \left( \frac{\partial f(\tau, \bar{\tau})}{\partial \tau} + \frac{\partial f(\tau, \bar{\tau})}{\partial \bar{\tau}} \right) < 0, \tag{31}$$

are negative (found by substituting eqq. [29] and [30] into [28]).

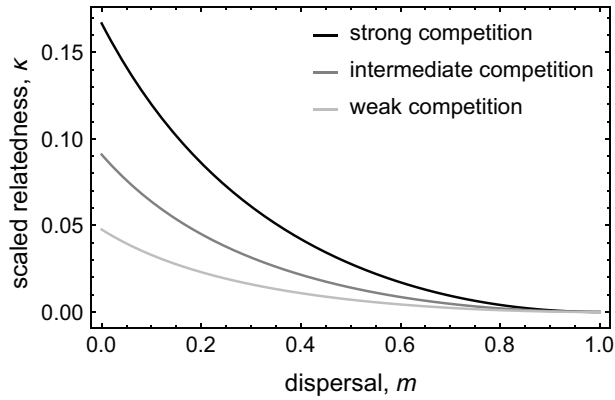
*Balance between intra- and intertemporal effects.* Summing equations (27) and (31), we find that the selection gradient is proportional to

$$s(\theta) \propto \frac{\partial f(\tau, \bar{\tau})}{\partial \tau} + \kappa(\theta) \frac{\partial f(\tau, \bar{\tau})}{\partial \bar{\tau}}, \tag{32}$$

where

$$\kappa(\theta) = \frac{(1 - m)^2}{\hat{n}(\theta)f(\theta, \theta) - (1 - m)^2(\hat{n}(\theta) - 1)} = r(\theta) \frac{1 - \beta(\theta)}{1 - r(\theta)\beta(\theta)} \tag{33}$$

is a scaled relatedness coefficient, which decreases with dispersal ( $m$ ; for the definition of  $\beta(\theta) > 0$ , see fig. 4 and eq. [31]). This scaled relatedness can be understood by looking at the right-hand side of equation (33). There the relatedness coefficient  $r(\theta)$  in the numerator reflects selection on helping due to its positive intratemporal indirect effects (eq. [27]). This positive effect, however, is discounted by a factor  $[1 - \beta(\theta)]/[1 - r(\theta)\beta(\theta)] < 1$ , due to the negative intertemporal indirect effects of helping (eq. [31]). Scaled relatedness coefficient  $\kappa(\theta)$  thus reflects how selection on helping depends on the balance between the positive intratemporal indirect effects of helping and its negative intertemporal



**Figure 4:** Scaled relatedness for the evolution of social interactions within species. Equation (33) with equation (25) plotted against dispersal with strong ( $\gamma = 0.2$ , black), intermediate ( $\gamma = 0.1$ , dark gray), and weak ( $\gamma = 0.05$ , light gray) levels of competition. Other parameters:  $f(\theta, \theta) = 2$ . Scaled relatedness therefore decreases with dispersal and patch size (since smaller values of  $\gamma$  lead to larger equilibrium patch size).

indirect effects owing to increased competition for relatives living in the future.

To better understand the balance between intra- and intertemporal effects, it is noteworthy that relatedness among individuals decreases with the number of generations that separate them (eq. [19]). As a result, selection on fitness effects becomes increasingly weak over generations. This is reflected in the fact that the scaled relatedness coefficient is nonnegative (i.e.,  $\kappa(\theta) \geq 0$ ; eq. [33]; fig. 4). In fact, provided  $r(\theta) > 0$  (so that  $\kappa(\theta) > 0$ ), altruistic helping can evolve in our model. This can be seen more explicitly if we further assume that maximal fecundity is given by

$$f(\tau, \tilde{\tau}) = f_b \cdot (1 - C\tau^2 + B\tilde{\tau}), \quad (34)$$

where  $f_b$  is a baseline fecundity;  $C$  is the cost of helping, which increases quadratically with the investment of the focal into helping; and  $B$  is the benefit of helping, which results from one unit invested into helping. Substituting equation (34) into equations (32) and (33) and solving  $s(\theta^*) = 0$  allows us to find the singular strategy  $\theta^*$ . When both  $C$  and  $B$  are small (of the order of a parameter  $\epsilon \ll 1$ ), the singular strategy  $\theta^*$  can be found by solving a first-order Taylor expansion of the selection gradient about  $\epsilon = 0$ . Doing so, we obtain a simple expression for the convergence-stable strategy,

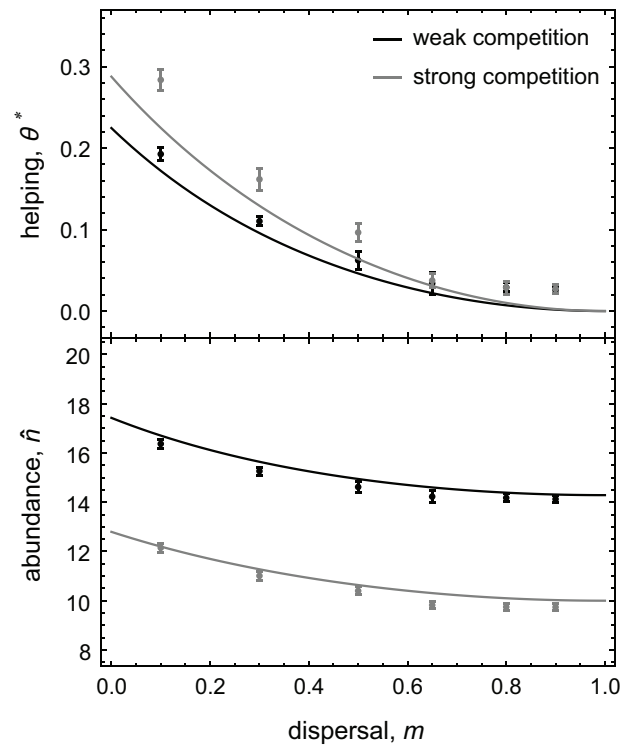
$$\theta^* = \frac{B}{2C} \kappa(0) = \frac{B}{2C} \cdot \frac{1}{1 + \hat{n}(0) \left( \frac{f_b}{(1-m)^2} - 1 \right)} \geq 0. \quad (35)$$

Equation (35) makes it straightforward to see that helping can evolve in spite of its negative intertemporal indirect ef-

fects. It further shows how the equilibrium level of helping,  $\theta^*$ , decreases with dispersal  $m$  and local abundance in the absence of helping,  $\hat{n}(0) = (f_b - 1)/\gamma$ .

More generally, by solving equations (32) and (33) with equation (34) numerically, we find that predictions generated from our approximation fit qualitatively and quantitatively well with observations from individual-based simulations, as much for the value of the convergence-stable level of helping  $\theta^*$  (fig. 5, top panel) as for the concomitant equilibrium group size  $\hat{n}(\theta^*)$  this generates (fig. 5, bottom panel).

*Connections to Previous Results on Altruism Evolution.* Our finding that altruism decreases as dispersal and local abundance increase is a standard result of evolutionary bi-



**Figure 5:** Convergence-stable level of within-species helping and the concomitant local abundance it generates. Solid lines are the convergence-stable strategies (top graph) and concomitant local abundance (bottom graph) obtained from the selection gradient (obtained by finding the  $\theta^*$  value solving eq. [32]) under weak (black,  $\gamma = 0.07$ ) and strong (gray,  $\gamma = 0.1$ ) intraspecific competition (other parameters:  $B = 0.5$ ,  $C = 0.05$ ,  $f = 2$ ). Points are the results obtained from individual-based simulations (time average over 40,000 generations after 10,000 generations of evolution; error bars show standard deviation). Parameters for simulations: 1,000 patches;  $m = 0.1, 0.3, 0.5, 0.65, 0.8, 0.9$ ; probability of a mutation = 0.01; standard deviation of the quantitative effect of a mutation = 0.005 (for details on simulations, see app. sec. A1, available online).

ology. However, our model of altruism and its results depart in two ways from the literature on this topic (e.g., Taylor 1992; van Baalen and Rand 1998; Taylor and Irwin 2000; Gardner and West 2006; Lehmann et al. 2006; Lehmann 2007; El Mouden and Gardner 2008; Johnstone and Cant 2008; Lion and Gandon 2009, 2010; Sozou 2009; Wild 2011; Bao and Wild 2012; Johnstone et al. 2012; Kuijper and Johnstone 2012; Ohtsuki 2012; Rodrigues and Gardner 2012). First, the vast majority of previous analyses assume that density-dependent competition occurs for space after dispersal (i.e., space or “open sites” is the limiting factor; e.g., Tilman 1982, ch. 8). In this case, intratemporal kin competition effects strongly inhibit the benefits of interacting among relatives. Here we have assumed that competition occurs for resources before dispersal. In this situation, we found that intratemporal kin competition effects do not abate the selective advantage of interacting with relatives (this can be seen from eq. [27], which depends only on pairwise relatedness). Rather, by increasing abundance, altruism increases kin competition for future generations (eq. [31]). This also hinders the evolution of altruism but only moderately so, because relatedness between individuals of different generations is on average lower than individuals of the same generation.

A second important difference between our model and previous models of social evolution with endogenous patch dynamics is that the latter had to exclusively rely on numerical approaches to compute the selection gradient (in the island model of dispersal; e.g., for models of altruism evolution: Lehmann et al. 2006; Alizon and Taylor 2008; Wild et al. 2009; Wild 2011; for models of dispersal evolution: Metz and Gyllenberg 2001; Cadet et al. 2003; Parvinen et al. 2003; Rousset and Ronce 2004). This reliance on numerical analysis makes it more difficult to understand how scaled relatedness  $\kappa$  and selection on altruism vary with demographic parameters (e.g., Lehmann et al. 2006, eq. [12]). Here our approximation yields a simple and intuitive expression for the selection gradient (eq. [32]), which nonetheless fits well with simulation results (fig. 5).

It is noteworthy that the selection gradient we have derived for this example (eq. [32]) applies to any type of social interactions within the life cycle given in “Biological Scenario.” In fact, the selection gradient equation (32) can be adjusted to study other social behaviors simply by changing the fecundity function (e.g., eq. [34]). Such selection gradients written in terms of marginal fecundity effects of behavior have also been derived for lattice-structured populations using the pair approximation (Lion and Gandon 2009, eq. [14]; Lion and Gandon 2010, eq. [19]). Comparing these expressions with ours would be interesting, in particular to investigate the effects of isolation by distance (which are ignored here).

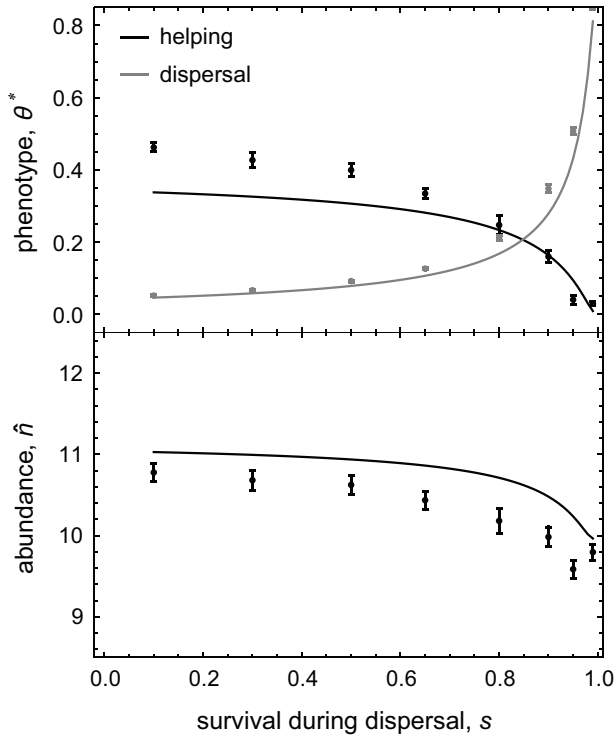
*Coevolution of Helping and Dispersal.* Our social evolution model assumes that dispersal is fixed. Dispersal, however, is likely to be an evolving trait. Because dispersal determines whether individuals interact and compete with relatives, dispersal evolution is important for selection on social behavior (e.g., Le Galliard et al. 2005; Purcell et al. 2012; Mullon et al. 2018). Dispersal evolution can also influence demography, in particular when individuals compete for space (e.g., when offspring survival after dispersal depends on local abundance; Metz and Gyllenberg 2001; Cadet et al. 2003; Parvinen et al. 2003; Rousset and Ronce 2004).

To test whether our approximation could capture the interplay between social behavior, dispersal, and demography, we used it to study a model of the coevolution between altruistic helping and dispersal when offspring compete for space following dispersal. We assumed that dispersal is costly, with offspring surviving dispersal with a probability  $s < 1$ . Details on this model and its analysis are given in appendix C.

We find that dispersal increases as survival during dispersal,  $s$ , increases (fig. 6, top panel, gray curve). This in turn selects for lower levels of helping (fig. 6, top panel, black curve), in line with previous models of helping-dispersal coevolution that assume that demography is constant (e.g., Mullon et al. 2018). Here we further find that as survival during dispersal  $s$  increases, the resulting collapse in helping and increase in dispersal leads to fewer individuals populating each patch (fig. 6, bottom panel). The predictions derived from our approximation agree well with observations we made from individual-based simulations for the equilibria of the two traits and the concomitant abundance these equilibria generate (fig. 6). This supports the idea that the approximate selection gradient (eqq. [13]–[16]) can be used to model dispersal evolution, in particular when local demography, genetic structure, and social traits feed back on one another. Our approximation, however, cannot be used to investigate disruptive selection, which can emerge when helping and dispersal coevolve (e.g., Purcell et al. 2012; Mullon et al. 2018). Such an investigation would require studying the second-order effects of selection, which is beyond the scope of this article.

### *Harming between Species*

*Biological Scenario.* To illustrate how the approximate selection gradient can be applied to study ecological interactions among species, we now model the evolution of antagonistic interactions among two species, species 1 and species 2. Specifically, we model the evolution of a trait in species 1 that is costly to express and that harms individuals of species 2. Our two species go through the following life cycle: (1) Individuals reproduce. A focal individual of species 1 has mean fecundity  $f_1(\tau, s_1)/(1 + \gamma n_1 + \eta n_2)$ , which



**Figure 6:** Coevolutionary convergence-stable level of helping and dispersal and the concomitant local abundance it generates. Solid lines are the convergence-stable strategies (top graph) for helping (black) and dispersal (gray) and concomitant local abundance (bottom graph) obtained from the selection gradient (eq. [C7], with  $B = 0.5$ ,  $C = 0.05$ ,  $f = 2$ ,  $\gamma = 0.05$ ). Points are the results obtained from individual-based simulations (time average over 40,000 generations after 10,000 generations of evolution; error bars show standard deviation). Parameters for simulations: 1,000 patches; survival during dispersal  $s = 0.1, 0.3, 0.5, 0.65, 0.8, 0.9, 0.95, 0.99$ ; probability of a mutation = 0.01; standard deviation of the quantitative effect of a mutation on each trait = 0.005; and no covariance (for details on simulations, see app. sec. A1, available online).

decreases with intra- and interspecific competition (respectively measured by parameters  $\gamma$  and  $\eta$ ). The maximal fecundity of a focal individual of species 1,  $f_1(\tau_{\cdot,1})$ , decreases with its investment  $\tau_{\cdot,1}$  into harming (i.e.,  $f_1'(\tau_{\cdot,1}) < 0$ ). A focal individual of species 2 has mean fecundity  $f_2/(1 + \gamma n_2)$ , where  $f_2$  is the maximal fecundity of species 2 and  $\gamma$  is the level of intraspecific competition. Note that only species 1 experiences interspecific competition. This would occur, for instance, because species 2 is a generalist consumer while species 1 is a specialist. (2) Adult individuals of species 1 kill offspring of species 2 in amount  $D(\bar{\tau}_1)$  per capita, which increases with the average  $\bar{\tau}_1$  level of harming in the focal patch (i.e.,  $D'(\bar{\tau}_1) > 0$ ). Costly interspecific harming could for instance occur through the release of chemicals into the environment that suppress the growth and establishment

of offspring (i.e., through allelopathy; Lankau 2008). (3) All adults of both species die. (4) Surviving offspring of species 1 and 2 disperse with probability  $m_1$  and  $m_2$ , respectively. (5) All offspring survive to adulthood.

*Necessary Components.* We first specify the components necessary for deriving the selection gradient on interspecific harming. According to the above, a focal individual from species 1 that invests  $\tau_{\cdot,1}$  into harming in a patch with  $\mathbf{n} = (n_1, n_2)$  individuals of species 1 and 2, respectively, has fitness

$$w_1(\tau_{\cdot,1}, \mathbf{n}, \hat{\mathbf{n}}(\boldsymbol{\theta})) = \frac{f_1(\tau_{\cdot,1})}{1 + \gamma n_1 + \eta n_2}. \quad (36)$$

The abundances of both species in the focal patch after one iteration of the life cycle, given that (1) the average level of harming in the patch is  $\bar{\tau}_1$ , (2) the abundance at the previous generation was  $\mathbf{n} = (n_1, n_2)$ , and (3) other patches are at equilibrium  $\hat{\mathbf{n}}(\boldsymbol{\theta}) = (\hat{n}_1(\theta_1), \hat{n}_2(\theta_1))$ , are given by

$$F_1(\bar{\tau}_1, \mathbf{n}, \hat{\mathbf{n}}(\boldsymbol{\theta})) = (1 - m_1) \left( \frac{f_1(\bar{\tau}_1)}{1 + \gamma n_1 + \eta n_2} \right) n_1 + m_1 \hat{n}_1(\theta_1),$$

$$F_2(\bar{\tau}_1, \mathbf{n}, \hat{\mathbf{n}}(\boldsymbol{\theta})) = (1 - m_2) \left( \frac{f_2}{1 + \gamma n_2} - D(\bar{\tau}_1) \frac{n_1}{n_2} \right) n_2 + m_2 \hat{n}_2(\theta_1). \quad (37)$$

The resident ecological equilibrium, which is found by solving  $\hat{\mathbf{n}}(\boldsymbol{\theta}) = (F_1(\theta_1, \hat{\mathbf{n}}(\boldsymbol{\theta}), \hat{\mathbf{n}}(\boldsymbol{\theta})), F_2(\theta_1, \hat{\mathbf{n}}(\boldsymbol{\theta}), \hat{\mathbf{n}}(\boldsymbol{\theta})))$  simultaneously, is too complicated to be presented here for the general case. Note however that when the resident level  $\theta_1$  of harming is small, a first-order Taylor expansion of the resident ecological equilibrium around  $\theta_1 = 0$  gives

$$\begin{aligned} \hat{n}_1(\theta_1) &= \hat{n}_1(0) + \frac{\theta_1}{\gamma} \left( \frac{f_2}{f_2 - 1} \hat{n}_1(0) \eta D'(0) - \underbrace{c_h(0)}_{-f_1(0)} \right) + O(\theta_1^2), \\ \hat{n}_2(\theta_1) &= \hat{n}_2(0) - \theta_1 \left( \frac{f_2}{f_2 - 1} \hat{n}_1(0) D'(0) \right) + O(\theta_1^2), \end{aligned} \quad (38)$$

where

$$\begin{aligned} \hat{n}_1(0) &= \frac{f_1(0) - 1}{\gamma} - \frac{\eta}{\gamma} \hat{n}_2(0), \\ \hat{n}_2(0) &= \frac{f_2 - 1}{\gamma} \end{aligned} \quad (39)$$

are the abundances in the absence of harming (assuming that  $D(0) = 0$ ). Equation (38) reveals that harming of species 2 reduces its abundance (i.e.,  $\hat{n}_2(\theta_1) \leq \hat{n}_2(0)$ ). The abun-



dance of species 1,  $\hat{n}_1(\theta_1)$ , depends on the balance between two opposite effects of harming. On one hand, by reducing the abundance of species 2, harming increases the abundance of species 1 due to interspecific competition (this is captured by the first summand within brackets on the first line of eq. [38]). On the other hand, abundance decreases with the cost of harming (which is captured by  $c_h(0) = -f'_1(0) > 0$  in eq. [38]). The final component necessary to derive the selection gradient is the relatedness coefficient for species 1. It is given by equation (10), with the resident ecological equilibrium for species 1,  $\hat{n}_1(\theta_1)$ , which is defined by equation (37).

*Selection on Harming.* We now detail the selection gradient on harming, according to the intra- and intertemporal effects of this behavior.

*Intratemporal effects.* Substituting equation (36) into equation (14), we obtain that selection on the intratemporal effects of harming,

$$s_{w,1}(\theta_1) = -\frac{c_h(\theta_1)}{f_1(\theta_1)} < 0, \tag{40}$$

where  $c_h(\theta_1) = -f'_1(\theta_1) > 0$ , is always negative. This is because harming is intratemporally costly to express at the individual level and does not provide any intratemporal indirect fitness benefits. Hence, the only way for harming to evolve in this model is if this cost is compensated by future benefits received by downstream relatives, which we investigate below.

*Intertemporal effects.* Selection on harming due to its effects on the fitness of downstream relatives is captured by the intertemporal part of the selection gradient equation (15). Substituting equations (36)–(38) into equations (15) and (20), we find that selection on harming due to its intertemporal effects is given by

$$s_{e,1}(\theta_1) = \frac{1 - m_1}{f_1(\theta_1)} \left( (1 - m_1) \frac{c_h(\theta_1)}{f_1(\theta_1)} \gamma + (1 - m_2) D'(\theta_1) \eta \right) + O(\epsilon^4), \tag{41}$$

where  $\epsilon$  is such that  $O(1 - m_1) \sim O(1 - m_2) \sim O(\epsilon)$ . From equation (41), we see that overall, intertemporal effects favor the evolution of interspecific harming (i.e.,  $s_{e,1}(\theta_1) > 0$  since  $c_h(\theta_1) = -f'_1(\theta_1) > 0$  and  $D'(\theta_1) > 0$ ). This is due to two intertemporal effects of harming. First, as captured by the first summand within brackets of equation (41), harming benefits downstream relatives that remain philopatric because by paying the cost of harming, a focal individual produces less offspring and thus diminishes local intraspecific competition. Accordingly, this first effect scales with the tendency to remain philopatric in the harming species,  $1 - m_1$ , the cost of harming,  $c_h(\theta_1)$ , and the strength of intraspe-

cific competition,  $\gamma$ . The second way that harming benefits downstream relatives is by reducing the local abundance of the harmed species (species 2, as captured by the second summand within brackets of eq. [41]), which reduces interspecific competition. In line with this, harming is favored when the effect of harming,  $D'(\theta_1)$ , and the intensity of interspecific competition,  $\eta$ , are large and when dispersal is limited in both species (in species 1 to ensure that relatives benefit from the reduction of interspecific competition and in species 2 since otherwise the local abundance of species 2 in downstream generations depends only on the process of immigration and not on local harming; eq. [37]).

*Convergence-stable equilibrium of harming.* To explicitly test the effect of limited dispersal on the evolution of harming, we assumed that the fecundity of an individual of species 1 that expresses a level  $\tau_{\cdot,1}$  of harming is

$$f_1(\tau_{\cdot,1}) = f_{1,b} \cdot (1 - C\tau_{\cdot,1}^2), \tag{42}$$

where  $f_{1,b}$  is a baseline fecundity in species 1 and  $C$  is the individual cost of harming. We further assumed that an individual of species 2 that is in a patch in which the average harming level is  $\bar{\tau}_1$  suffers a fecundity cost given by

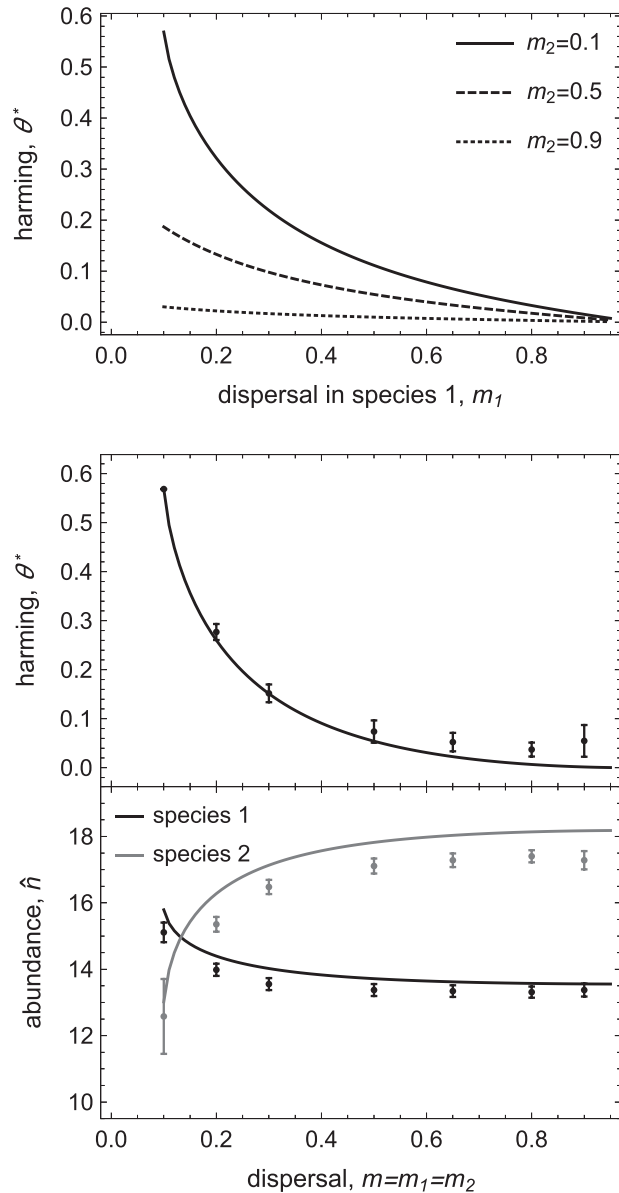
$$D(\bar{\tau}_1) = \alpha \bar{\tau}_1, \tag{43}$$

where  $\alpha$  is a parameter tuning the deleteriousness of harming. The convergence-stable level of harming, which is found by solving  $s_1(\theta_1^*) = s_{w,1}(\theta_1^*) + s_{e,1}(\theta_1^*) = 0$  for  $\theta_1^*$ , is shown in figure 7 as a function of dispersal.

In line with equation (41), we find that individually costly harming does not evolve when dispersal is complete (fig. 7, top and middle panels). This is because in that case, downstream relatives can never benefit from a decrease of interspecific competition owing to harming. As dispersal becomes limited in both species, this intertemporal benefit increasingly goes to relatives so that harming evolves (fig. 7, top and middle panels). This evolution in turn causes a significant reduction in the abundance of species 2 and an increase of species 1 (fig. 7, bottom panel). These results were confirmed using individual-based simulations, further supporting the goodness of fit of our approximation (fig. 7, middle and bottom panels).

### Discussion

Due to the physical limitations of movement, a community of species is typically structured in space to form a metacommunity (e.g., Tilman 1982; Clobert et al. 2001; Urban et al. 2008; Leibold and Chase 2017). Understanding selection in such a metacommunity is challenging due to the feedback between local ecology and trait composition that emerges when dispersal is limited and local demography is stochastic.



**Figure 7:** Convergence-stable level of between-species harming and the concomitant local abundance it generates. Solid lines are the convergence-stable strategies (top and middle graphs) and concomitant local abundance (bottom graph) in species 1 (black) and 2 (gray) obtained from the selection gradient (eqq. [13]–[20] along with eqq. [36]–[43], with  $C_1 = 0.0085$ ,  $f_1 = 2.2$ ,  $f_2 = 2$ ,  $\gamma = 0.055$ ,  $\eta = 0.025$ ,  $\alpha = 0.24$ ; top: dispersal in species 2,  $m_2 = 0.1$  (solid),  $0.5$  (dashed),  $0.9$  (smaller dashed); middle and bottom: dispersal in species  $m_2 = m_1 = m$ . Points are the results obtained from individual-based simulations (time average over 40,000 generations after 10,000 generations of evolution; error bars show standard deviation). Parameters for simulations: 1,000 patches;  $m = 0.1, 0.2, 0.3, 0.5, 0.65, 0.8, 0.9$ ; probability of a mutation = 0.01; standard deviation of the quantitative effect of a mutation = 0.005.

To better understand these eco-evolutionary dynamics, we derived in this article an approximation for the selection gradient on a quantitative trait that influences local ecology in the island model of dispersal.

The basis of our approximation is to neglect ecological stochasticity and to assume that the resulting deterministic ecological dynamics have a single fixed point (i.e., we do not consider periodic or chaotic dynamics). We nonetheless take into account the consequences of genetic stochasticity for selection. We found that this approximation works well qualitatively for all models and conditions that we studied. We further found that it is quantitatively accurate in predicting ecological and evolutionary dynamics as long as dispersal is not excessively weak. As a rule of thumb, effective dispersal rate should be no less than 0.1 when patches are small (with fewer than 10 individuals; figs. 1, 5–7). Such demographic regime leads to an  $F_{ST}$  well within the range of  $F_{ST}$  values that have been estimated across a wide spectrum of taxa (when  $m_i(\theta) = 0.1$  and  $\hat{n}_i(\theta) = 10$ , eq. [10] gives  $F_{ST} = 0.30$  for haploids; for diploids, eq. [6.23] of Hartl and Clark 2007 gives  $F_{ST} = 0.20$ ; equivalently, this regime entails one migrant per demographic time period; i.e.,  $\hat{n}_i(\theta)m_i(\theta) = 1$ ; for empirical estimates, see Barton 2001, p. 334; Hartl and Clark 2007, p. 302). This suggests that our approximation takes into account dispersal levels that are relevant to many species (Bohonak 1999).

The simplicity of our approximate selection gradient allows us to investigate convergence-stable species coalitions and to intuitively understand community evolution under limited dispersal. In particular, our selection gradient reveals that selection can be decomposed into intra-temporal (eq. [14]) and intertemporal (eqq. [15], [16]) effects. Intertemporal effects reflect the interaction between kin selection and local eco-evolutionary dynamics. This interaction can be understood by considering that when a focal individual perturbs species abundance locally, this perturbation leads to changes in community composition in the future due to ecological interactions (eq. [16]; fig. 3). These changes then feed back on the fitness of individuals living in the future, who potentially carry genes that are identical by descent to the focal (i.e., who are relatives). In other words, intertemporal effects emerge because individuals inherit not only their genes but also an ecological environment that has been transformed by their ancestors (Odling-Smee et al. 2003; Bonduriansky 2012). By considering the effects of such ecological inheritance on multispecies interactions, our model generalizes previous models of local ecological interactions in the presence of relatives that ignored trait-induced changes in abundance (either altogether [Frank 1994; Foster and Wenseleers 2006; Wyatt et al. 2013; Akçay 2017] or in the evolving species [Lehmann 2008]).

Interestingly, the eco-evolutionary, intertemporal feedbacks that emerge in our model are captured mathemati-

cally by analogues of press perturbations (eqq. [20], [21]). These are central notions of classical community ecology. Press perturbations traditionally measure how a persistent change in the abundance of a given species alters the equilibrium abundance of another due to ecological interactions (Yodzis 1989; Case 2000). In our model, the change in abundance is initiated by a phenotypic change in a focal individual, and its persistence is measured over the time it is evolutionarily significant from the perspective of the focal, which is the time the focal's lineage of relatives experiences it (fig. 3). Because this time increases as dispersal becomes more limited, intertemporal effects are more important for selection when individuals remain in philopatry (i.e., when dispersal is limited). This can also be understood from a gene-centered perspective: because limited dispersal ties in the fate of a trait-changing mutation with its intertemporal ecological effects, these effects become more important for how selection targets this mutation.

We applied our model to the evolution of two behavioral traits with demographic and ecological consequences. First, we studied the evolution of altruistic helping within species. Our model follows a rich literature on this topic (see “Connections to Previous Results on Altruism Evolution”), which traditionally assumes that local patch size remains constant and that competition occurs for space after dispersal. In this case, the benefits of helping relatives are offset or partially offset by the intratemporal cost of kin competition (e.g., Taylor 1992; Gardner and West 2006; Lion and Gandon 2009). By contrast, our model assumes that competition occurs for material resources before dispersal and that helping increases baseline fecundity, thereby influencing patch size dynamics. As a result, the costs of kin competition are delayed and paid by downstream relatives. Because these intertemporal costs are weaker than intratemporal ones, we found that selection favors intraspecific altruism for a large range of parameters, in particular when dispersal is limited (fig. 5). Our results therefore suggest that altruistic traits are more likely to be found in species in which competition occurs for material resources rather than for space.

Second, we studied the evolution of individually costly harming between species. We found that harming evolution strongly depends on intertemporal effects. Specifically, we found that harming evolves when it alleviates interspecific competition for downstream relatives by reducing the other species' abundance, which requires that dispersal is limited in both interacting species. Our analysis thus makes the empirical prediction that antagonism is more likely when dispersal is limited (fig. 7). Previous theory has focused on understanding how “altruism” or mutualism between species can evolve in the presence of relatives (Frank 1994; Foster and Wenseleers 2006; Wyatt et al. 2013; Akçay 2017). These studies have highlighted

the fact that the evolution of mutualism is facilitated by among-species genetic correlations, which emerge when dispersal is limited. Here our model reveals that antagonism between species can also evolve in this case, which raises the interesting question of whether mutualistic or antagonistic interactions are more likely to evolve under limited dispersal. Presumably, this would depend on the degree of interspecific competition.

Beyond the examples presented here, our analysis helps identify the conditions under which selection on a trait in one species depends on intertemporal effects that arise from the effects of this trait on the demographic dynamics of other species. First, when expressed in a focal individual, a change in this trait should have a large influence on the local community over one demographic time period (i.e.,  $\Psi_{k,i}(\theta)$  should be large; eq. [17]). We expect this to be the case for traits that are directly involved in interspecific interactions—such as defenses against predators, traits that attract mutualists, or resource-extraction strategies—especially when expressed by keystone or dominant species (i.e., species with large effects on communities). Our analysis further reveals that there are more opportunities for selection on intertemporal effects when the local abundances of the different species that are part of the community are interdependent (so that [1] fitness in the focal species depends on community composition,  $\partial w_i / \partial n_j \neq 0$ ; [2] the community matrix  $\mathbf{C}(\theta)$  is nonsparse; and [3] evolutionary press perturbations  $\Lambda_i(\theta)$  are large; eq. [21]). This ensures that the ecological perturbation initiated by a focal individual has multiple downstream effects through indirect ecological interactions (e.g., focal trait in species 1 increases the abundance of species 2, who is a competitor of species 3, who itself is a competitor of species 1; e.g., terHorst et al. 2018). Multiple downstream effects of a perturbation by a focal individual then increase the likelihood that this perturbation feeds back on the fitness of downstream relatives of the focal. This leads to the broad prediction that communities that are tightly interconnected are more likely to show traits whose intertemporal ecological effects are under selection.

One crucial condition for intertemporal ecological effects to be under selection is that dispersal is limited. This needs to be the case in the focal species to ensure that relatives experience trait-driven ecological changes, but also in other species of the community so that the effect of local interactions on abundance is not swamped by immigration. In our example of harming between species, for instance, we found that harming did not evolve if either the harming or the recipient species showed full dispersal (fig. 7, top panel). Plant communities would be ideal to test the notion that traits within spatially limited communities are more likely to have intertemporal effects that have been shaped by selection. For instance, many plants are engaged in interspecific chemical warfare, with lasting

effects on soil composition (for review, see Inderjit et al. 2011). In light of our results, it would be interesting to study how these intertemporal effects of allelopathy vary with the degree of dispersal (or gene flow, which can be estimated from  $F_{ST}$  values). In particular, we expect allelopathy to be most adapted among competitors that show interspecific genetic correlations.

Of course, the approximate selection gradient derived here cannot be applied to all evolutionary scenarios. It should generally be supplemented with simulation checks, in particular when dispersal is severely limited and patches are very small (e.g., when populations are close to extinction). In fact, it would be useful to analyze our model with greater mathematical rigor to obtain a sharper understanding of the conditions under which ecological stochasticity can be neglected (e.g., by generalizing the results of Chesson 1981). One major limitation to our approach is that it relies on the assumption that ecological dynamics converge to a fixed point. This assumption, which allowed us to improve the understanding of selection on traits affecting metacommunity stochastic demography, precludes the consideration of limit cycles or spatiotemporal fluctuations in abundance, which are thought to be prevalent in many ecological systems (e.g., Yodzis 1989; Case 2000). It would therefore be very relevant to extend our approach to derive the selection gradient under more complicated ecological dynamics. Another assumption we have made is that reproduction occurs as a discrete time process. It would thus be relevant to derive the selection gradient under continuous time, but this is unlikely to change our main qualitative results (as this essentially requires replacing sums by integrals and individual fitness by individual growth rates and by calculating intertemporal relatedness coefficients in continuous time; e.g., Sozou 2009).

To conclude, our heuristic approximation is a step further toward the integration of multispecies ecological theory and kin selection theory. Owing to its simplicity and intuitive interpretation, the approximate selection gradient we have derived can provide a useful guide to answer questions that lie at the intersection of ecology and evolution. In particular, it can be straightforwardly applied to study plant-pollinator, host-parasite, or predator-prey coevolution under limited dispersal or the eco-evolutionary dynamics of sex-specific dispersal. These and other applications should help explain how selection molds intra- and interspecific interactions when dispersal is limited.

### Acknowledgments

We thank James Rodger for a useful discussion on dispersal evolution. We also thank Scott Nuismer, Minus Van Baalen, and an anonymous reviewer for helpful comments on the manuscript. C.M. is funded by Swiss National Science Foundation grant PP00P3-123344 to L.L.

### Literature Cited

- Abrams, P. A., H. Matsuda, and Y. Harada. 1993. Evolutionarily unstable fitness maxima and stable fitness minima of continuous traits. *Evolutionary Ecology* 7:465–487.
- Ajar, E. 2003. Analysis of disruptive selection in subdivided populations. *BMC Evolutionary Biology* 3:22.
- Akçay, E. 2017. Population structure reduces benefits from partner choice in mutualistic symbiosis. *Proceedings of the Royal Society B* 284:1–10.
- Akçay, E., and J. Van Cleve. 2012. Behavioral responses in structured populations pave the way to group optimality. *American Naturalist* 179:257–269.
- Alizon, S., and P. D. Taylor. 2008. Empty sites can promote altruistic behavior. *Evolution* 62:1335–1344.
- Apaloo, J., and S. Butler. 2009. Evolutionary stabilities in multidimensional-traits and several-species models. *Evolutionary Ecology Research* 11:637–650.
- Bao, M., and G. Wild. 2012. Reproductive skew can provide a net advantage in both conditional and unconditional social interactions. *Theoretical Population Biology* 82:200–208.
- Barton, N. H. 2001. The evolutionary consequences of gene flow and local adaptation: future approaches. Pages 329–340 in J. Clobert, E. Danchin, A. A. Dhondt, and J. D. Nichols, eds. *Dispersal*. Oxford University Press, New York.
- Bohonak, A. J. 1999. Dispersal, gene flow, and population structure. *Quarterly Review of Biology* 74:21–45.
- Bonduriansky, R. 2012. Rethinking heredity, again. *Trends in Ecology and Evolution* 27:330–336.
- Brown, J. S., and T. L. Vincent. 1987. Coevolution as an evolutionary game. *Evolution* 41:66–79.
- Cadet, C., R. Ferrière, J. A. J. Metz, and M. van Baalen. 2003. The evolution of dispersal under demographic stochasticity. *American Naturalist* 162:427–441.
- Case, T. J. 2000. *An illustrated guide to theoretical ecology*. Oxford University Press, Oxford.
- Charlesworth, B. 1994. *Evolution in age-structured populations*. 2nd ed. Cambridge University Press, Cambridge.
- Charnov, E. L. 1976. Optimal foraging, the marginal value theorem. *Theoretical Population Biology* 9:129–136.
- Chesson, P. L. 1978. Predator-prey theory and variability. *Annual Review of Ecology and Systematics* 9:323–347.
- . 1981. Models for spatially distributed populations: the effect of within-patch variability. *Theoretical Population Biology* 19:288–325.
- . 1984. Persistence of a Markovian population in a patchy environment. *Zeitschrift für Wahrscheinlichkeitstheorie* 66:97–107.
- Christiansen, F. B. 1991. On conditions for evolutionary stability for a continuously varying character. *American Naturalist* 138:37–50.
- Clobert, J., E. Danchin, A. A. Dhondt, and J. D. Nichols, eds. 2001. *Dispersal*. Oxford University Press, New York.
- Comins, H. N., W. D. Hamilton, and R. M. May. 1980. Evolutionarily stable dispersal strategies. *Journal of Theoretical Biology* 82:205–230.
- Cornell, S. J., and O. Ovaskainen. 2008. Exact asymptotic analysis for metapopulation dynamics on correlated dynamic landscapes. *Theoretical Population Biology* 74:209–225.
- Débarre, F., S. Lion, M. van Baalen, and S. Gandon. 2012. Evolution of host life-history traits in a spatially structured host-parasite system. *American Naturalist* 179:52–63.

- Débarre, F., S. L. Nuismer, and M. Doebeli. 2014. Multidimensional (co)evolutionary stability. *American Naturalist* 184:158–171.
- Débarre, F., and S. P. Otto. 2016. Evolutionary dynamics of a quantitative trait in a finite asexual population. *Theoretical Population Biology* 108:75–88.
- El Mouden, C., and A. Gardner. 2008. Nice natives and mean migrants: the evolution of dispersal-dependent social behaviour in viscous populations. *Journal of Evolutionary Biology* 21:1480–1491.
- Eshel, I. 1983. Evolutionary and continuous stability. *Journal of Theoretical Biology* 103:99–111.
- . 1996. On the changing concept of evolutionary population stability as a reflection of a changing point of view in the quantitative theory of evolution. *Journal of Mathematical Biology* 34:485–510.
- Foster, K. R. 2004. Diminishing returns in social evolution: the not-so-tragic commons. *Journal of Evolutionary Biology* 17:1058–1072.
- Foster, K. R., and T. Wenseleers. 2006. A general model for the evolution of mutualisms. *Journal of Evolutionary Biology* 19:1283–1293.
- Frank, S. A. 1994. Genetics of mutualism: the evolution of altruism between species. *Journal of Theoretical Biology* 170:393–400.
- . 1998. *Foundations of social evolution*. Princeton University Press, Princeton, NJ.
- Gandon, S., and Y. Michalakis. 1999. Evolutionary stable dispersal rate in a metapopulation with extinction and kin competition. *Journal of Theoretical Biology* 199:275–290.
- Gardner, A., and S. A. West. 2004. Cooperation and punishment, especially in humans. *American Naturalist* 164:753–764.
- . 2006. Demography, altruism, and the benefits of budding. *Journal of Evolutionary Biology* 19:1707–1716.
- 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* 12:35–57.
- Grafen, A. 1991. Modeling in behavioural ecology. Pages 5–31 in J. R. Krebs and N. Davies, eds. *Behavioural ecology*. Blackwell Scientific, Oxford.
- . 2006. A theory of Fisher's reproductive value. *Journal of Mathematical Biology* 53:15–60.
- Hamilton, W. D. 1964. The genetical evolution of social behaviour. I. *Journal of Theoretical Biology* 7:1–16.
- . 1966. The moulding of senescence by natural selection. *Journal of Theoretical Biology* 12:12–45.
- . 1971. Selection of selfish and altruistic behaviour in some extreme models. Pages 59–91 in J. Eisenberg and W. Dillon, eds. *Man and beast: comparative social behavior*. Smithsonian Institution, Washington, DC.
- Hamilton, W. D., and R. M. May. 1977. Dispersal in stable habitats. *Nature* 269:578–581.
- Hanski, I., and M. Gilpin. 1997. *Metapopulation biology: ecology, genetics, and evolution*. Academic Press, San Diego, CA.
- Hartl, D., and A. G. Clark. 2007. *Principles of population genetics*. 4th ed. Sinauer, Sunderland, MA.
- Hubbell, S. P. 2001. *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton, NJ.
- Inderjit, D. A. Wardle, R. Karban, and R. M. Callaway. 2011. The ecosystem and evolutionary contexts of allelopathy. *Trends in Ecology and Evolution* 26:655–662.
- Johnstone, R. A., and M. A. Cant. 2008. Sex differences in dispersal and the evolution of helping and harming. *American Naturalist* 172:318–330.
- Johnstone, R. A., M. A. Cant, and J. Field. 2012. Sex-biased dispersal, haplodiploidy and the evolution of helping in social insects. *Proceedings of the Royal Society B* 279:787–793.
- Kisdi, E., and T. Priklopil. 2010. Evolutionary branching of a magic trait. *Journal of Mathematical Biology* 63:361–397.
- Kobayashi, Y., J. Y. Wakano, and H. Ohtsuki. 2015. A paradox of cumulative culture. *Journal of Theoretical Biology* 379:79–88.
- Kuijper, B., and R. A. Johnstone. 2012. How dispersal influences parent-offspring conflict over investment. *Behavioral Ecology* 23:898–906.
- Kuijper, B., I. Pen, and F. J. Weissing. 2012. A guide to sexual selection theory. *Annual Review of Ecology, Evolution, and Systematics* 43:287–311.
- Lankau, R. 2008. A chemical trait creates a genetic trade-off between intra- and interspecific competitive ability. *Ecology* 89:1181–1187.
- Le Galliard, J., R. Ferrière, and U. Dieckmann. 2003. The adaptive dynamics of altruism in spatially heterogeneous populations. *Evolution* 57:1–17.
- . 2005. Adaptive evolution of social traits: origin, trajectories. *American Naturalist* 165:206–224.
- Lehmann, L. 2007. The evolution of trans-generational altruism: kin selection meets niche construction. *Journal of Evolutionary Biology* 20:181–189.
- . 2008. The adaptive dynamics of niche constructing traits in spatially subdivided populations: evolving posthumous extended phenotypes. *Evolution* 62:549–566.
- Lehmann, L., C. Mullon, E. Akçay, and J. Van Cleve. 2016. Invasion fitness, inclusive fitness, and reproductive numbers in heterogeneous populations. *Evolution* 70:1689–1702.
- Lehmann, L., N. Perrin, and F. Rousset. 2006. Population demography and the evolution of helping behaviors. *Evolution* 60:1137–1151.
- Leibold, M. A., and J. M. Chase. 2017. *Metacommunity ecology*. Vol. 59. Princeton University Press, Princeton, NJ.
- Leimar, O. 2009. Multidimensional convergence stability. *Evolutionary Ecology Research* 11:191–208.
- Leslie, P. H., and J. C. Gower. 1958. The properties of a stochastic model for two competing species. *Biometrika* 45:316–330.
- Lion, S. 2017. Theoretical approaches in evolutionary ecology: environmental feedback as a unifying perspective. *American Naturalist* 191:21–44.
- Lion, S., and S. Gandon. 2009. Habitat saturation and the spatial evolutionary ecology of altruism. *Journal of Evolutionary Biology* 22:1487–1502.
- . 2010. Life history, habitat saturation and the evolution of fecundity and survival altruism. *Evolution* 64:1594–1606.
- Lion, S., and M. van Baalen. 2007a. From infanticide to parental care: why spatial structure can help adults be good parents. *American Naturalist* 170:E26–E46.
- . 2007b. Self-structuring in spatial evolutionary ecology. *Ecology Letters* 11:277–295.
- Maynard Smith, J. 1982. *Evolution and the theory of games*. Cambridge University Press, Cambridge.
- Metz, J. A. J., S. A. H. Geritz, G. Meszéna, F. J. A. Jacobs, and J. S. van Heerwaarden. 1996. Adaptive dynamics: a geometrical study of the consequences of nearly faithful reproduction. Pages 183–231 in S. J. van Strien and S. M. Verduyn Lunel, eds. *Stochastic and spatial structures of dynamical systems*. North-Holland, Amsterdam.
- Metz, J. A. J., and M. Gyllenberg. 2001. How should we define fitness in structured metapopulation models? including an applica-

- tion to the calculation of evolutionary stable dispersal strategies. *Proceedings of the Royal Society B* 268:499–508.
- Michod, R. E. 1979. Evolution of life histories in response to age-specific mortality factors. *American Naturalist* 113:531–550.
- Mullon, C., L. Keller, and L. Lehmann. 2016. Evolutionary stability of jointly evolving traits in subdivided populations. *American Naturalist* 188:175–195.
- . 2018. Social polymorphism is favoured by the co-evolution of dispersal with social behaviour. *Nature Ecology and Evolution* 2:132–140.
- Mullon, C., M. Reuter, and L. Lehmann. 2014. The evolution and consequences of sex-specific reproductive variance. *Genetics* 196:235–252.
- Nagylaki, T. 1992. *Introduction to population genetics*. Springer, Heidelberg.
- Nakamaru, M., and Y. Iwasa. 2005. The evolution of altruism by costly punishment in lattice-structured populations: score-dependent viability versus score-dependent fertility. *Evolutionary Ecology Research* 7:853–870.
- Nakamaru, M., H. Matsuda, and Y. Iwasa. 1997. The evolution of cooperation in a lattice-structured population. *Journal of Theoretical Biology* 184:65–81.
- Neuhauser, C. 2002. Effects of local interactions and local migration on stability. *Theoretical Population Biology* 62:297–308.
- Nurmi, T., K. Parvinen, and V. Selonen. 2018. Joint evolution of dispersal propensity and site selection in structured metapopulation models. *Journal of Theoretical Biology* 444:50–72.
- Odling-Smee, F. J., K. L. Laland, and M. W. Feldman. 2003. *Niche construction*. Princeton University Press, Princeton, NJ.
- Ohtsuki, H. 2012. Does synergy rescue the evolution of cooperation? an analysis for homogeneous populations with non-overlapping generations. *Journal of Theoretical Biology* 307:20–28.
- Parker, G. A. 1990. Sperm competition games: raffles and roles. *Proceedings of the Royal Society B* 242:120–126.
- Parker, G. A., and J. Maynard Smith. 1990. Optimality theory in evolutionary biology. *Science* 349:27–33.
- Parvinen, K., U. Dieckmann, M. Gyllenberg, and J. A. J. Metz. 2003. Evolution of dispersal in metapopulations with local density dependence and demographic stochasticity. *Journal of Evolutionary Biology* 16:143–153.
- Parvinen, K., and J. A. J. Metz. 2008. A novel fitness proxy in structured locally finite metapopulations with diploid genetics, with an application to dispersal evolution. *Theoretical Population Biology* 73:517–528.
- Parvinen, K., H. Ohtsuki, and J. Y. Wakano. 2018. Spatial heterogeneity and evolution of fecundity-affecting traits. *Journal of Theoretical Biology* 454:190–204.
- Purcell, J., A. Brelsford, and L. Avilés. 2012. Co-evolution between sociality and dispersal: the role of synergistic cooperative benefits. *Journal of Theoretical Biology* 312:44–54.
- Ranta, E., P. Lundberg, and V. Kaitala. 2006. *Ecology of populations*. Cambridge University Press, Cambridge.
- Rodrigues, A. M. M., and A. Gardner. 2012. Evolution of helping and harming in heterogeneous populations. *Evolution* 66:2065–2079.
- Rousset, F. 2004. *Genetic structure and selection in subdivided populations*. Princeton University Press, Princeton, NJ.
- Rousset, F., and O. Ronce. 2004. Inclusive fitness for traits affecting metapopulation demography. *Theoretical Population Biology* 65:127–141.
- Schaffer, W. M. 1982. The application of optimal control theory to the general life history problem. *American Naturalist* 121:418–431.
- Sozou, P. D. 2009. Individual and social discounting in a viscous population. *Proceedings of the Royal Society B* 276:2955–2962.
- Stearns, S. 1992. *The evolution of life histories*. Oxford University Press, Oxford.
- Taylor, P. D. 1988a. An inclusive fitness model for dispersal of offspring. *Journal of Theoretical Biology* 130:363–378.
- . 1988b. Inclusive fitness models with two sexes. *Theoretical Population Biology* 34:145–168.
- . 1989. Evolutionary stability in one-parameter models under weak selection. *Theoretical Population Biology* 36:125–143.
- . 1992. Altruism in viscous populations—an inclusive fitness model. *Evolutionary Ecology* 6:352–356.
- . 1994. Sex ratio in a stepping-stone population with sex-specific dispersal. *Theoretical Population Biology* 45:203–218.
- Taylor, P. D., and S. A. Frank. 1996. How to make a kin selection model. *Journal of Theoretical Biology* 180:27–37.
- Taylor, P. D., and A. J. Irwin. 2000. Overlapping generations can promote altruistic behavior. *Evolution* 54:1135–1141.
- terHorst, C. P., P. C. Zee, K. D. Heath, T. E. Miller, A. I. Pastore, S. Patel, S. J. Schreiber, M. J. Wade, and M. R. Walsh. 2018. Evolution in a community context: trait responses to multiple species interactions. *American Naturalist* 191:368–380.
- Tilman, D. 1982. *Resource competition and community structure*. Princeton University Press, Princeton, NJ.
- Urban, M. C. 2011. The evolution of species interactions across natural landscapes. *Ecology Letters* 14:723–732.
- Urban, M. C., M. A. Leibold, P. Amarasekare, L. De Meester, R. Gomulkiewicz, M. E. Hochberg, C. A. Klausmeier, et al. 2008. The evolutionary ecology of metacommunities. *Trends in Ecology and Evolution* 23:311–317.
- van Baalen, M., and A. Rand. 1998. The unit of selection in viscous populations and the evolution of altruism. *Journal of Theoretical Biology* 193:631–648.
- Van Cleve, J. 2015. Social evolution and genetic interactions in the short and long term. *Theoretical Population Biology* 103:2–26.
- Wagner, C. E., R. J. Best, L. J. Harmon, and B. Matthews. 2017. Evolution in a community context: on integrating ecological interactions and macroevolution. *Trends in Ecology and Evolution* 32:291–304.
- Wakano, J. Y., and C. Miura. 2014. Trade-off between learning and exploitation: the Pareto-optimal versus evolutionarily stable learning schedule in cumulative cultural evolution. *Theoretical Population Biology* 91:37–43.
- Weigang, H. C., and E. Kisdi. 2015. Evolution of dispersal under a fecundity-dispersal trade-off. *Journal of Theoretical Biology* 371:145–153.
- West, S. A., and A. Gardner. 2010. Altruism, spite, and greenbeards. *Science* 327:1341–1344.
- West, S. A., A. S. Griffin, and A. Gardner. 2007. Evolutionary explanations for cooperation. *Current Biology* 17:661–672.
- Wild, G. 2011. Inclusive fitness from multitype branching processes. *Bulletin of Mathematical Biology* 73:1028–1051.
- Wild, G., A. Gardner, and S. A. West. 2009. Adaptation and the evolution of parasite virulence in a connected world. *Nature* 459:983–986.
- Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16:97–159.
- Wyatt, G. A., S. A. West, and A. Gardner. 2013. Can natural selection favour altruism between species? *Journal of Evolutionary Biology* 26:1854–1865.

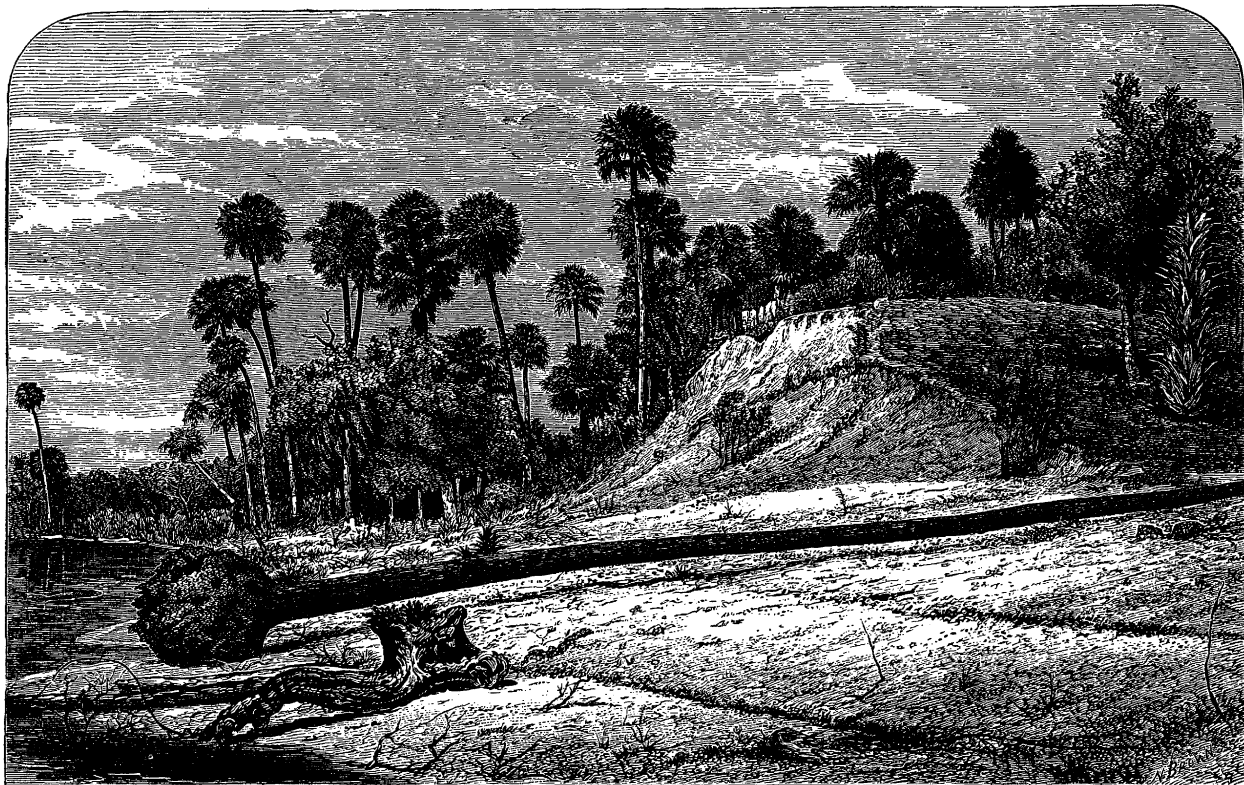
Yodzis, P. 1989. Introduction to theoretical ecology. Harper & Row, New York.

### References Cited Only in the Online Appendixes

- Caswell, H. 2000. Matrix population models. Sinauer, Sunderland, MA.
- Cohen, J. E. 1979. Long-run growth rates of discrete multiplicative processes in Markovian environments. *Journal of Mathematical Analysis and Applications* 69:243–251.
- Epperson, B. K. 1999. Gene genealogies in geographically structured populations. *Genetics* 152:797–806.
- Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, MA.
- Malécot, G. 1973. Génétique des population diploïdes naturelle dans le cas d'un seul locus. III. Parenté, mutations et migration. *Annale de Génétique et de Selection Animale* 5:333–361.
- Metz, J. A. J., and C. G. F. de Kovel. 2013. The canonical equation of adaptive dynamics for Mendelian diploids and haplo-diploids. *Interface Focus* 3:20130025.

- Metz, J. A. J., and O. Leimar. 2011. A simple fitness proxy for structured populations with continuous traits, with case studies on the evolution of haplo-diploids and genetic dimorphisms. *Journal of Biological Dynamics* 5:163–190.
- Meyn, S., and R. L. Tweedie. 2009. Markov chains and stochastic stability. 2nd ed. Cambridge University Press, Cambridge.
- Mullon, C., and L. Lehmann. 2018. Data from: Eco-evolutionary dynamics in metacommunities: ecological inheritance, helping within species, and harming between species. *American Naturalist*, Dryad Digital Repository, <http://dx.doi.org/10.5061/dryad.kk1qj10>.
- Sydsaeter, K., P. Hammond, A. Seierstad, and A. Strøm. 2008. Further mathematics for economic analysis. 2nd ed. Prentice Hall, Essex.
- Tuljapurkar, S. 1989. An uncertain life: demography in random environments. *Theoretical Population Biology* 35:227–294.
- Tuljapurkar, S., C. C. Horvitz, and J. B. Pascarella. 2003. The many growth rates and elasticities of populations in random environments. *American Naturalist* 162:489–502.
- Wolfram Research. 2016. Mathematica. Version 10.0.1.0. Wolfram Research, Champaign, IL.

Associate Editor: Scott L. Nuismer  
 Editor: Russell Bonduriansky



“These shell-mounds, built up exclusively of fresh-water species, are peculiar, in being formed mainly by accumulations of *Ampullarias* and *Paludinas*, with a small percentage of mussel shells (*Unios*), as elsewhere these heaps are entirely formed of *Unios*, the other shells being either very scantily represented or altogether absent.” From the review of Wyman’s *Fresh-Water Shell-Mounds of the St. John’s River, Florida* (*The American Naturalist*, 1876, 10:165–169).