Eco-evolutionary dynamics under limited dispersal: ecological inheritance, helping within- and harming between-species

Charles Mullon * [1] and Laurent Lehmann [1]

[1] *Department of Ecology and Evolution, University of Lausanne, 1004 Lausanne, Switzerland*

Abstract

Understanding selection on ecological interactions that take place in dispersal-limited communities is an important challenge for ecology and evolutionary biology. The problem is that local demographic stochasticity generates eco-evolutionary dynamics that are 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, but incorporating the 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- and harming between-species when these behaviours influence demography. We find helping evolves more readily when competition is for material resources rather than space because then, the costs of kin competition are paid by downstream relatives. Similarly, harming between species evolves when it alleviates downstream relatives from inter-specific competition. Beyond these examples, our approximation can help better understand the influence of ecological inheritance on a variety of eco-evolutionary dynamics, from plant-pollinator to predator-prey coevolution.

Keywords: kin selection, adaptive dynamics, meta-population, altruism, mutualism.

*Electronic address: `charles.mullon@unil.ch`

1 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 inter-species interactions emerge and are maintained depends on interdependent ecological and evolutionary processes (terHorst et al., 2018), also known as eco-evolutionary dynamics (Lion, 2017 for review). One major difficulty in understanding these dynamics is due to the spatial structuring of communities, which results from interactions between individuals occurring at a finer spatial scale than the global scale of the community. This spatial structure is captured by the notion of a "meta-community", 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, Hartl and Clark, 2007, Urban et al., 2008). When dispersal among communities is limited, individual reproductive and survival variance generate local demographic stochasticity, which in turn leads to spatial heterogeneities in community structure among patches. These heterogeneities make ecological and evolutionary analyses very challenging.

In fact, local demographic stochasticity has two complicating consequences for eco-evolutionary dynamics. First, it causes genetic stochasticity, whereby allele frequencies fluctuate within patches. These fluctuations lead to the build up 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 recipient (indirect effects) (e.g., Hamilton, 1971, Hamilton and May, 1977, Taylor, 1994, Taylor and Frank, 1996, Frank, 1998, Rousset, 2004, West et al., 2007, Lion and van Baalen, 2007b, Van Cleve, 2015). Second, local demographic stochasticity results in ecological stochasticity, whereby the abundance of different species fluctuate within and between patches. As a consequence, multi-species demography in patch structured populations is significantly more complicated that 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 eco-evolutionary feedbacks, understanding selection on traits that mediate 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, Parker and Maynard Smith, 1990, Christiansen, 1991, Grafen, 1991, Abrams et al., 1993, Metz et al., 1996, Eshel, 1996, Geritz et al., 1998, Rousset, 2004). One of the main goals is to identify local attractors of **such** adaptive dynamics. These attractors are trait values towards 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). Con-

vergence 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 understand how natural selection moulds phenotypic traits of broad biological interest, from senescence, life history, sperm competition, sex-ratio, to altruism, cumulative cultural evolution and optimal foraging (e.g., Hamilton, 1966, Charnov, 1976, Schaffer, 1982, Taylor, 1988b, Parker, 1990, Taylor, 1988a, Frank, 1998, Gardner and West, 2004, Foster, 2004, Kuijper et al., 2012, Akçay and Van Cleve, 2012, Mullon et al., 2014, Wakano and Miura, 2014, Kobayashi et al., 2015).

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 and Ronce, 2004, eqs. 23-24, Rousset, 2004, chapter 11). In principle, these expressions can be extended to consider multi-species interactions. However, even under the simplest model of limited dispersal, which is the island model of dispersal (Wright, 1931), the selection gradient on traits affecting multi-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 a 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 very 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). A biologically intuitive understanding of selection on traits that influence meta-community dynamics is therefore out of immediate reach when using the exact selection gradient.

To circumvent the difficulty of computing 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 intra-specific 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, 2007a, Lion and Gandon, 2009, 2010, Débarre et al., 2012). However, more general multi-species coevolution scenarios have not been much investigated using pair approximation. This is presumably because analytical explorations remain complicated in lattice structured populations due to isolation-by-distance.

In this paper, we present a novel heuristic approximation for the selection gradient on traits that influence eco-evolutionary dynamics in patch-structured populations that 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, and see Chesson, 1981, for its ecological counterpart). 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 multi-species interactions. Importantly, our approximation provides a biologically meaningful interpretation of selection on traits that influence ecological interactions under limited dispersal. The rest of the paper is organized as follows. (1) We describe a stochastic meta-community eco-evolutionary model. (2) We motivate an approximation of our model that ignores ecological stochasticity. (3) Under this approximation, we derive analytically the selection gradient on a trait that influences eco-evolutionary dynamics through intra- and inter-species interactions. (4) We use our approximation to study two examples of intra- and inter-species interactions: evolution of helping within- and harming between-species when these behaviours influence demography.

2 Model

2.1 Meta-community 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 $i \in \{1,2,...,S\}$ are as follows. (1) Each adult produces a finite number of offspring that follows a Poisson distribution. (2) Each adult then either survives or dies (with a probability that may depend on local interactions). (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 non-zero 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, for instance if space on each patch is a limiting factor).

2.2 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 Θ_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 meta-community (i.e., the collection of subdivided populations of each species) is described by a vector of strategies $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_S)$, where θ_i is the strategy expressed by all individuals of species i (i.e. $\boldsymbol{\theta}$ denotes a monomorphic resident population).

We define $\boldsymbol{\theta}$ 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 $\boldsymbol{\tau}_i = (\theta_1, ..., \theta_{i-1}, \tau_i, \theta_{i+1}, ..., \theta_S)$ of the resident vector, goes extinct. The concept of an uninvadable coalition is the same as the concepts of a multi-species evolutionary stable strategy (Brown and Vincent, 1987, p. 68) and of an evolutionary stable coalition (Apaloo and Butler, 2009, p. 640).[1]

2.3 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 in order 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 τ_i, when the resident community expresses $\boldsymbol{\theta}$. In the infinite island model of dispersal, the change Δp_i in frequency p_i of such a mutant allele over one demographic time period (one life cycle iteration) can be written as,

$$\Delta p_i = \delta_i p_i (1 - p_i) S_i(\boldsymbol{\theta}) + O(\delta_i^2), \tag{1}$$

where $\delta_i = \tau_i - \theta_i$ is the phenotypic effect of the mutation (Rousset, 2004, pp. 206–207 and Rousset and Ronce, 2004 p. 129, with notation here adapted to introduce species specific allele frequency change). The function $S_i(\boldsymbol{\theta})$, which depends on quantities evaluated only in the resident community $\boldsymbol{\theta}$, 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 $S_i(\boldsymbol{\theta})$ and phenotypic effect δ_i give the direction of selection on the mutant at any allele frequency: selection favours fixation of the mutation when $\delta_i S_i(\boldsymbol{\theta}) > 0$, and conversely, extinction when $\delta_i S_i(\boldsymbol{\theta}) < 0$. The selection gradient thus captures the influence of directional selection.

The selection gradient $S_i(\boldsymbol{\theta})$ is useful to derive necessary conditions for a coalition to be uninvadable. Specifically, the coalition $\boldsymbol{\theta}^*$ is uninvadable only if it satisfies

$$S_i(\boldsymbol{\theta}^*) = 0 \qquad \forall i. \tag{2}$$

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

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

[1] We here refrain of using the terminology "evolutionary stability" as it subsumes that such strategies are attractor of the evolutionary dynamics (Maynard Smith, 1982), which is not covered by the concept of uninvadability.

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

When at most two alleles can ever segregate at a given locus, 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 that 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 characterize analytically under limited dispersal (e.g., Ajar, 2003, Mullon et al., 2016). We will therefore focus on the effects of directional selection in this paper and investigate only the properties of the selection gradient.

2.4 The selection gradient in meta-communities

The selection gradient in a given species 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 (eqs. 26-27 of Rousset and Ronce, 2004 for a single-species demography dynamics, or eqs. E.27-29 of Lehmann et al., 2016 for arbitrary demographic states). Reproductive values reflect the fact that individuals residing in different demographic states contribute differently to the 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 (see Nagylaki, 1992 and Rousset, 2004 for textbook treatments). 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 significant problems. The first is that owing to the large number of possible demographic states within patches in our model (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 complicated computations necessary even in the absence of inter-specific 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 problem 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 carry little biological information about how local ecological interactions influence selection.

The goal of this paper is to provide a tractable and biologically informative approximation for the selection gradient in a meta-community. We propose to achieve this by assuming that changes in local population size are deterministic with a stable fixed point when the meta-community is monomorphic for $\boldsymbol{\theta}$. 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 meta-community monomorphic for the resident $\boldsymbol{\theta}$) in order to investigate how and when these dynamics can be assumed to be deterministic.

3 Resident community dynamics

3.1 Deterministic resident community dynamics

Our life-cycle assumptions (section 2.1) entail that we are considering an infinite stochastic system, i.e., 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 generation t, which we denote by $\boldsymbol{n}_t(\boldsymbol{\theta}) = (n_{1,t}(\boldsymbol{\theta}), n_{2,t}(\boldsymbol{\theta}), \ldots, n_{S,t}(\boldsymbol{\theta}))$ where $n_{i,t}(\boldsymbol{\theta})$ 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 $n_{i,t}(\boldsymbol{\theta}) = F_i\big(\boldsymbol{\theta}, \boldsymbol{N}_{t-1}, \overline{\boldsymbol{n}}_{t-1}(\boldsymbol{\theta})\big)$, for $i = 1, 2, ..., S$, where the population dynamic transition map F_i gives the expected number of individuals of species i, given the local community state $\boldsymbol{N}_{t-1} = (N_{1,t-1}, N_{2,t-1}, ..., N_{S,t-1})$ in the previous generation (i.e., $N_{i,t-1}$ denotes the random variable for the number of individuals of species i in the focal patch at generation $t-1$), and when the global average community state is $\overline{\boldsymbol{n}}_{t-1}(\boldsymbol{\theta}) = (\overline{n}_{1,t-1}(\boldsymbol{\theta}), \overline{n}_{2,t-1}(\boldsymbol{\theta}), ..., \overline{n}_{S,t-1}(\boldsymbol{\theta}))$ (i.e., $\overline{n}_{i,t-1}(\boldsymbol{\theta})$ is the average number of individuals of species i across all patches at generation $t-1$, note that $\overline{\boldsymbol{n}}_{t-1}(\boldsymbol{\theta})$ is not a random variable because there is an infinite number of patches). We have written the transition map $F_i\big(\boldsymbol{\theta}, \boldsymbol{N}_{t-1}, \overline{\boldsymbol{n}}_{t-1}(\boldsymbol{\theta})\big)$ such that it also depends explicitly on the vector of phenotypes $\boldsymbol{\theta}$ of each species in the focal patch in the previous generation, 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, $\boldsymbol{N}_t \approx \mathrm{E}[\boldsymbol{N}_t] = \boldsymbol{n}_t(\boldsymbol{\theta})$), so that the ecological dynamics on a focal patch,

which are no longer stochastic, are given by

$$n_{i,t}(\boldsymbol{\theta}) = F_i\big(\boldsymbol{\theta}, \boldsymbol{n}_{t-1}(\boldsymbol{\theta}), \overline{\boldsymbol{n}}_{t-1}(\boldsymbol{\theta})\big) \qquad \text{for } i = 1,2,...,S \tag{4}$$

(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 eq. (4), this ecological fixed point, which we denote as $\hat{\boldsymbol{n}}(\boldsymbol{\theta}) = (\hat{n}_1(\boldsymbol{\theta}), \hat{n}_2(\boldsymbol{\theta}), \ldots, \hat{n}_S(\boldsymbol{\theta}))$, solves

$$\hat{n}_i(\boldsymbol{\theta}) = F_i(\boldsymbol{\theta}, \hat{\boldsymbol{n}}(\boldsymbol{\theta}), \hat{\boldsymbol{n}}(\boldsymbol{\theta})) \qquad \text{for } i = 1,2,...,S. \tag{5}$$

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

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

whose $(i\,j)$-entry measures the influence of the abundance of species j on the abundance of species i over one generation, has eigenvalues with absolute value less than one.

3.2 Illustrating example

In order to illustrate the transition function (F_i, 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 $\boldsymbol{n} = (n_1, n_2)$ individuals produces a mean number $f_i / \big(1 + \gamma n_i + \eta n_j\big)$ 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 γ), and between species (with intensity η). 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 $m_i(\boldsymbol{\theta})$. (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 $\boldsymbol{n}_{t-1}(\boldsymbol{\theta})$ in the previous time period on the focal patch and on the abundance in other patches being at a

stable equilibrium $\hat{\boldsymbol{n}}(\boldsymbol{\theta})$, can be written as

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

Eq. (7) is an example of a co-called coupled map lattice (e.g., eq. 5 of Jansen and Lloyd, 2000). The first summand in each line of eq. (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. In order to understand better this second summand, consider that when the population is at the resident demographic equilibrium, an individual on average produces one offspring (so that the total number of individuals remains constant). As a consequence, $\hat{n}_1(\boldsymbol{\theta})$ and $\hat{n}_2(\boldsymbol{\theta})$ give the total number of offspring produced of species 1 and 2 respectively, in any patch other than the focal one. Therefore, $m_1(\boldsymbol{\theta})\hat{n}_1(\boldsymbol{\theta})$ and $m_2(\boldsymbol{\theta})\hat{n}_2(\boldsymbol{\theta})$ in eq. (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 eq. (7) in eq. (5) (i.e., putting $n_{1,t}(\boldsymbol{\theta}) = n_{1,t-1}(\boldsymbol{\theta}) = \hat{n}_1(\boldsymbol{\theta})$ and $n_{2,t}(\boldsymbol{\theta}) = n_{2,t-1}(\boldsymbol{\theta}) = \hat{n}_2(\boldsymbol{\theta})$) and solving for $\hat{n}_1(\boldsymbol{\theta})$ and $\hat{n}_2(\boldsymbol{\theta})$ simultaneously. Doing so, we find that the unique positive equilibrium (i.e., $\hat{n}_1(\boldsymbol{\theta}) > 0$ and $\hat{n}_2(\boldsymbol{\theta}) > 0$) reads as

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

This reveals that for the two species to co-exist, it is necessary for intra-specific competition to be stronger than inter-specific competition ($\gamma > \eta$), which is a classical result (Case, 2000).

3.3 Comparing deterministic and stochastic dynamics

In order 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 detailed in the above section 3.2 with individual based simulations of the full stochastic model (see Appendix A for a description of the simulation procedure).

3.3.1 Ecological stochasticity

We find that there is a good qualitative match between the deterministic dynamics given by eq. (7), and the average number of individuals of each species in a group observed in individual based simulations (Fig. 1, top

panel). As predicted by theory (Chesson, 1981, Neuhauser, 2002), the deterministic dynamics deviate systematically from the observed average (Fig. 1, top panel, grey lines). However, these deviations are small provided dispersal is not too weak (roughly when dispersal is greater than 0.1, Fig. 1, lower panel).

Why dispersal mitigates the effects of ecological stochasticity can be understood as follows. Local population dynamics depend on the balance between two processes: 1) a local process at the patch level (i.e., dependence on $\boldsymbol{N}_{t-1}$), which has a strong stochastic component when patches have few individuals; and 2) a global process at the meta-community level (i.e., dependence on $\overline{\boldsymbol{n}}_{t-1}(\boldsymbol{\theta})$), which has a weak stochastic component when the number of patches is large (in fact, as the number of patches grow infinite, patches affect each other deterministically, Chesson, 1981). As dispersal increases, local population dynamics depend increasingly on the global process and less on the local one. As a result, local population dynamics become increasingly deterministic.

3.3.2 Genetic stochasticity

The above analysis suggests that ecological stochasticity can be ignored when dispersal values are roughly above 0.1 (Fig. 1, lower panel). 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 $r_i(\boldsymbol{\theta})$ that two individuals from the same species $i \in \{1,2\}$, which are randomly sampled in the same patch, carry an allele that is identical-by-descent when the population is monomorphic for $\boldsymbol{\theta}$ (also referred to as pairwise relatedness, Frank, 1998, Rousset, 2004). Let us consider this probability when the community has reached its (deterministic) demographic equilibrium $\hat{\boldsymbol{n}}(\boldsymbol{\theta})$ (given by eq. 8). Owing to our assumption that fecundity is Poisson distributed, pairwise relatedness satisfies the relationship

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

which can be understood as follows. With probability $(1 - m_i(\boldsymbol{\theta}))^2$, two randomly sampled individuals of species i are both of philopatric origin (i.e., they were born in the focal patch). Then, with probability $1/\hat{n}_i(\boldsymbol{\theta})$, these individuals descend from the same parent so their relatedness is one. With complementary probability $1 - 1/\hat{n}_i(\boldsymbol{\theta})$, they descend from different parent so their relatedness is $r_i(\boldsymbol{\theta})$. The solution to eq. (9) is

$$r_i(\boldsymbol{\theta}) = \frac{(1 - m_i(\boldsymbol{\theta}))^2}{1 - \left[1 - (1 - m_i(\boldsymbol{\theta}))^2\right](\hat{n}_i(\boldsymbol{\theta}) - 1)}, \tag{10}$$

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 $\hat{\boldsymbol{n}}(\boldsymbol{\theta})$ here is endogenously determined by an explicit demographic process (eqs. 7-8).

Inspection of eq. (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. We will therefore take into account the effects 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, Figure 2). This lends further support to the usefulness of the deterministic ecological approximation to study populations at ecological equilibrium (eq. 5).

4 Evolutionary analysis

We now specify the (approximate) selection gradient on a trait expressed in species i, assuming that ecological dynamics are deterministic. First, we characterise such dynamics when they can be influenced by the presence of genetic mutants.

4.1 Mutant community dynamics

In contrast to the above section 3, two alleles now segregate in the focal species i: a mutant that codes for phenotype τ_i and a resident for θ_i. We focus on a focal patch in which both alleles are present, while other patches are considered to be monomorphic for the resident $\boldsymbol{\theta}$, and at their ecological equilibrium, $\hat{\boldsymbol{n}}(\boldsymbol{\theta})$ (eq. 5). For this focal patch, the dynamics of species abundance are now given by

$$n_{j,t}(\boldsymbol{\tau}_i) = F_j\big(\bar{\boldsymbol{\tau}}_{i,t-1}, \boldsymbol{n}_{t-1}(\boldsymbol{\tau}_i), \hat{\boldsymbol{n}}(\boldsymbol{\theta})\big) \qquad \text{for } j = 1, 2, ..., S, \tag{11}$$

where F_j is the one-generational map introduced in section 3.1 (see section 2.2 for a definition of $\boldsymbol{\tau}_i$). This map now depends on $\bar{\boldsymbol{\tau}}_{i,t} = (\theta_1, ..., \theta_{i-1}, \bar{\tau}_{i,t}, \theta_{i+1}, ..., \theta_S)$, which is a vector collecting the average phenotypes expressed in each species in the focal patch at generation t. In species j other than the focal ($j \neq i$), this average is simply the resident θ_j. In the focal species i, the average phenotype, which is denoted by $\bar{\tau}_{i,t}$, is given by

$$\bar{\tau}_{i,t} = \frac{M_{i,t}}{n_{i,t}}\tau_i + \left(1 - \frac{M_{i,t}}{n_{i,t}}\right)\theta_i, \tag{12}$$

where $M_{i,t} \in \{0, 1, 2, ..., n_{i,t}\}$ is the number of mutant individuals of species i at demographic time t in the focal patch (dependence of F_j on the average phenotype is a first-order approximation of possibly more complicated relationships between the distribution of mutants in the patch and local abundance, which is sufficient to evaluate the selection gradient, Rousset, 2004, p. 95). Since we take genetic stochasticity into account, $M_{i,t}$ should be viewed as a random variable. The ecological dynamics in a patch that contains mutants and residents are therefore stochastic (as the one-generational change in abundance depends on $M_{i,t}$, eq. 11). Once a focal patch is monomorphic (i.e., $M_{i,t} = 0$ or $M_{i,t} = 1$), however, ecological dynamics become fully deterministic again.

4.2 The inclusive fitness effect for the interactive community

In order 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 to drastically simplify our calculations and to equivalently characterise 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 $\boldsymbol{\theta}$ community can be approximated as

$$S_i(\boldsymbol{\theta}) \approx \epsilon(\boldsymbol{\theta}) s_i(\boldsymbol{\theta}), \tag{13}$$

where $\epsilon(\boldsymbol{\theta}) > 0$ is a factor of proportionality that depends on $\boldsymbol{\theta}$ only (see eqs. B.3 and B.6), and

$$s_i(\boldsymbol{\theta}) = s_{\mathrm{w},i}(\boldsymbol{\theta}) + s_{\mathrm{e},i}(\boldsymbol{\theta}) \tag{14}$$

is the approximate selection gradient. Since $\epsilon(\boldsymbol{\theta}) > 0$, $s_i(\boldsymbol{\theta})$ is sufficient to ascertain singular trait values and their convergence stability (by replacing $S_i(\boldsymbol{\theta})$ with $s_i(\boldsymbol{\theta})$ in eqs. 2 and 3). The approximate selection gradient, $s_i(\boldsymbol{\theta})$, consists of two terms: $s_{\mathrm{w},i}(\boldsymbol{\theta})$, which captures selection owing to the trait's intra-generational effects; and $s_{\mathrm{e},i}(\boldsymbol{\theta})$, which captures selection owing to the trait's inter-generational effects that 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 these two components of selection in the next two sections.

4.2.1 Selection on intra-generational effects

The first term of eq. (14) can be expressed as

$$s_{\mathrm{w},i}(\boldsymbol{\theta}) = \left.\frac{\partial w_i(\tau_{\bullet,i}, \tilde{\boldsymbol{\tau}}_i, \boldsymbol{n}, \hat{\boldsymbol{n}})}{\partial \tau_{\bullet,i}}\right|_{\substack{\tau_i=\theta_i \\ \boldsymbol{n}=\hat{\boldsymbol{n}}}} + \left.\frac{\partial w_i(\tau_{\bullet,i}, \tilde{\boldsymbol{\tau}}_i, \boldsymbol{n}, \hat{\boldsymbol{n}})}{\partial \tilde{\tau}_i}\right|_{\substack{\tau_i=\theta_i \\ \boldsymbol{n}=\hat{\boldsymbol{n}}}} \times r_i(\boldsymbol{\theta}) \tag{15}$$

(see Appendix B.3.1, eq. B.18 for derivation), 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 is written as a function of four variables: (1) the phenotype $\tau_{\bullet,i}$ of the focal individual; (2) the vector $\tilde{\boldsymbol{\tau}}_i = (\theta_1, ..., \theta_{i-1}, \tilde{\tau}_i, \theta_{i+1}, ..., \theta_S)$ of average phenotypes of neighbours 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 $\boldsymbol{n}$; and (4) the vector of average abundance across the meta-community, which is at its equilibrium $\hat{\boldsymbol{n}}(\boldsymbol{\theta})$ (explicit examples of such a fitness function are given later when we apply our method, see eqs. 24, 37 and C.1 in Appendix C). Note that individual fitness may also depend on the phenotype expressed in patches other than the focal, which is the resident $\boldsymbol{\theta}$, but we have chosen not to write this dependency explicitly.

The two derivatives in eq. (15), which are evaluated in the resident population (i.e., with resident phenotype $\tau_i = \theta_i$ and resident ecological equilibrium $\boldsymbol{n} = \hat{\boldsymbol{n}}(\boldsymbol{\theta})$), 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 neighbours resulting from the focal individual switching from the resident to mutant phenotype (i.e., the indirect effect of the trait). This second derivative is weighted by the relatedness coefficient, $r_i(\boldsymbol{\theta})$, which gives the probability that any neighbours also carries the mutation in the monomorphic resident.

4.2.2 Selection on inter-generational feedback effects due to ecological inheritance

The feedback between local ecology and evolution. We find that the second term of the selection gradient eq. (14) can be written as

$$s_{\mathrm{e},i}(\boldsymbol{\theta}) = \sum_{j=1}^{S} \left.\frac{\partial \overline{n}_j}{\partial \tau_i}\right|_{\substack{\tau_i=\theta_i \\ \boldsymbol{n}=\hat{\boldsymbol{n}}}} \times \left.\frac{\partial w_i(\tau_{\bullet,i}, \tilde{\boldsymbol{\tau}}_i, \boldsymbol{n}, \hat{\boldsymbol{n}})}{\partial n_j}\right|_{\substack{\tau_i=\theta_i \\ \boldsymbol{n}=\hat{\boldsymbol{n}}}} \tag{16}$$

(see Appendix B.3.2, eq. B.20 for details), where $\overline{n}_j$ is defined as the 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 τ_i that reside in the focal patch in which the mutation first appeared, see eq. B.21 in the appendix for a formal definition). The first derivative in eq. (16) is thus the effect that a trait increase in the focal species i has on the local abundance of species j, experienced by a local lineage of species i. The second derivative in eq. (16) is the effect that this abundance change of species j has on the fitness of an individual from the focal species i. By multiplying these two effects and summing them over all species j of the community, eq. (16) therefore captures how selection depends on the feedback between local community ecology and evolution.

A lineage-centred perspective on the ecological influence of a trait. The feedback effect captured by eq. (16) reveals that a phenotypic change will be selected when such a change results in local ecological conditions that are favourable for the lineage of those that express the change (i.e., when $\partial\overline{n}_j/\partial\tau_i \times \partial w_i(\tau_{\bullet,i}, \tilde{\boldsymbol{\tau}}_i, \boldsymbol{n}, \hat{\boldsymbol{n}})/\partial n_j > 0$). 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 $\partial\overline{n}_j/\partial\tau_i$. We find that this derivative can be expressed as

$$\left.\frac{\partial\overline{n}_j}{\partial\tau_i}\right|_{\substack{\tau_i=\theta_i\\ \boldsymbol{n}=\hat{\boldsymbol{n}}}} = \sum_{k=1}^{S}\Psi_{k,i}(\boldsymbol{\theta})\sum_{h=1}^{\infty}\mathrm{K}_{jk,h}(\boldsymbol{\theta})\left[\hat{n}_i(\boldsymbol{\theta})r_{i,h}(\boldsymbol{\theta})\right] \tag{17}$$

(see Appendix B.3.2, eq. B.28 for details). In order to understand eq. (17), consider a focal individual from species i that lives at a generation that we arbitrarily label as generation zero ($t = 0$, see blue star Fig. 3). The first term of eq. (17),

$$\Psi_{k,i}(\boldsymbol{\theta}) = \frac{1}{\hat{n}_i(\boldsymbol{\theta})}\left.\frac{\partial F_k(\bar{\boldsymbol{\tau}}_i, \boldsymbol{n}, \hat{\boldsymbol{n}})}{\partial\bar{\tau}_i}\right|_{\substack{\tau_i=\theta_i\\ \boldsymbol{n}=\hat{\boldsymbol{n}}}} \tag{18}$$

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 generation, i.e., at $t = 1$ (grey arrows, Fig. 3).

The second term in eq. (17), $\mathrm{K}_{jk,h}(\boldsymbol{\theta})$, is given by the (jk) element of the matrix

$$\mathbf{K}_h(\boldsymbol{\theta}) = \mathbf{C}(\boldsymbol{\theta})^{h-1}, \tag{19}$$

where $\mathbf{C}(\boldsymbol{\theta})$ is the community matrix given in eq. (6). Eq. (19) reveals that $\mathrm{K}_{jk,h}(\boldsymbol{\theta})$ in eq. (17) is the effect that a change in the abundance of species k at generation $t = 1$ has on the abundance of species j at generation $t = h$. Importantly, this effect takes into account the influence that species have on one anothers' abundance, cumulated over $h-1$ generations (as indicated by the exponent $h-1$ in eq. 19, see green arrows in Fig. 3 for e.g.).

Finally, the term in square brackets in eq. (17) can be interpreted as the expected number of relatives of the focal individual that live in the focal patch at generation $t = h \geq 1$ (and which therefore experience the mutant-modified ecological conditions at generation h, light blue disks, Fig. 3). Indeed, this term consists of the product of the equilibrium abundance, $\hat{n}_i(\boldsymbol{\theta})$, with

$$r_{i,h}(\boldsymbol{\theta}) = \left(1 - m_i(\boldsymbol{\theta})\right)^h \underbrace{\left(\frac{1}{\hat{n}_i(\boldsymbol{\theta})} + \frac{\hat{n}_i(\boldsymbol{\theta}) - 1}{\hat{n}_i(\boldsymbol{\theta})}r_i(\boldsymbol{\theta})\right)}_{\bar{r}_i(\boldsymbol{\theta})}, \tag{20}$$

which is the relatedness between two individuals of species i that are sampled h generations 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, see also eq. B.27 for a formal definition). Note that in general, $m_i(\boldsymbol{\theta})$ 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 behaviour, but also on other organismal aspects depending on the life-cycle (e.g., on adult survival from one generation to the next).

The above considerations (eqs. 18-19) show that the influence of a local lineage on the abundance of its own or another species j (eq. 17) 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 its downstream relatives residing in the focal patch (see Figure 3 for a diagram and Appendix B.3.2 for more mathematical details).

Evolutionary press perturbations. In order to evaluate eq. (17) explicitly, we can use the fact that under our assumptions that patches are not totally isolated from one another (i.e., $m_i(\boldsymbol{\theta}) > 0$) and that the resident community is at a stable fixed point (i.e., $\mathbf{C}(\boldsymbol{\theta})$ has eigenvalues with absolute value less than one), the infinite sum in eq. (17) converges. This leads to the following expression (see Appendix B.3.2, eq. B.29 for details),

$$\left.\frac{\partial \overline{n}_j}{\partial \tau_i}\right|_{\substack{\tau_i=\theta_i \\ \boldsymbol{n}=\hat{\boldsymbol{n}}}} = \sum_{k=1}^{S} \Psi_{k,i}(\boldsymbol{\theta})\,(1-m_i(\boldsymbol{\theta}))\,\hat{n}_i(\boldsymbol{\theta})\bar{r}_i(\boldsymbol{\theta})\Lambda_{jk,i}(\boldsymbol{\theta}), \tag{21}$$

where $\bar{r}_i(\boldsymbol{\theta})$ is the relatedness between two individuals sampled with replacement in the same patch (see eq. 20 for definition), and $\Lambda_{jk,i}(\boldsymbol{\theta})$ is given by the (jk)-entry of the matrix

$$\Lambda_i(\boldsymbol{\theta}) = \left(\mathbf{I} - [1-m_i(\boldsymbol{\theta})]\mathbf{C}(\boldsymbol{\theta})\right)^{-1}. \tag{22}$$

The term $\Lambda_{jk,i}(\boldsymbol{\theta})$ in eq. (21) 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 ($m_i(\boldsymbol{\theta}) \to 0$), the matrix $\Lambda_i(\boldsymbol{\theta})$ (eq. 22) tends to the matrix of press perturbations of community ecology [i.e., $\Lambda_i(\boldsymbol{\theta}) \to (\mathbf{I}-\mathbf{C}(\boldsymbol{\theta}))^{-1}$]. 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 multi-species interactions (e.g., Yodzis, 1989, Case, 2000). The correspondence between eq. (22) and press perturbation matrices reflects that as $m_i(\boldsymbol{\theta}) \to 0$, the mutant lineage may persist locally forever and thus experience persistent changes in the abundance of other species. But as dispersal $m_i(\boldsymbol{\theta})$ increases, the mutant lineage will spend fewer generations locally, which means that its experience of changes in local species abundance will last fewer generations (and so $\Lambda_i(\boldsymbol{\theta})$ approaches the identity matrix as dispersal becomes complete, i.e., $\Lambda_i(\boldsymbol{\theta}) \to \mathbf{I}$ as $m_i(\boldsymbol{\theta}) \to 1$).

4.2.3 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 ($m_i(\boldsymbol{\theta}) = 1$ so that $r_i(\boldsymbol{\theta}) = 0$). In this case, the selection gradient reduces to

$$s_i(\boldsymbol{\theta}) = \left.\frac{\partial w_i(\tau_{\bullet,i}, \tilde{\boldsymbol{\tau}}_i, \boldsymbol{n}, \hat{\boldsymbol{n}})}{\partial \tau_{\bullet,i}}\right|_{\substack{\tau_i=\theta_i \\ \boldsymbol{n}=\hat{\boldsymbol{n}}}}, \tag{23}$$

which embodies the classical ecological feedback considered by the field of adaptive dynamics (e.g., eq. 29 of Lion, 2017): the invasion of a rare mutant depends on resident-set ecological conditions only (i.e., on $\boldsymbol{\theta}$ and $\hat{\boldsymbol{n}}(\boldsymbol{\theta})$ only), and if the mutant invades, it becomes the resident and thereby modifies these conditions. The simplicity of eq. (23) reflects 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 has on its own fitness.

When dispersal is limited ($m_i(\boldsymbol{\theta}) < 1$), however, a globally rare mutant may become locally common and remain so over multiple generations. This has two implications that are important for the way selection targets this mutant. First, mutants living in the same generation interact directly with one another. This effect is captured by the relatedness-weighted fitness effect of neighbours in $s_{\mathrm{w},i}(\boldsymbol{\theta})$ (i.e., the second summand of eq. 15). In fact, $s_{\mathrm{w},i}(\boldsymbol{\theta})$ (eq. 15) 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, demography in $s_{\mathrm{w},i}(\boldsymbol{\theta})$ (eq. 15) is endogenously determined and evaluated at the resident ecological equilibrium, $\hat{\boldsymbol{n}}(\boldsymbol{\theta})$. As such, $s_{\mathrm{w},i}(\boldsymbol{\theta})$ approximates the exact selection gradient in a demographically-structured population when the trait has no demographic effect (denoted as S_{f}, eq. 26 of Rousset and Ronce, 2004, eq. E-28 of Lehmann et al., 2016). The main difference between the approximation and the exact expression is that the latter depends on reproductive values while the approximation $s_{\mathrm{w},i}(\boldsymbol{\theta})$ (eq. 15) 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 in different generations interact indirectly, through heritable ecological modifications. Selection due to these indirect interactions is captured by the second term of the selection gradient, $s_{\mathrm{e},i}(\boldsymbol{\theta})$ (eq. 14). When the evolving species has constant demography and interacts with only one other species, the term $s_{\mathrm{e},i}(\boldsymbol{\theta})$ reduces to eq. A39 of Lehmann (2008). More generally, $s_{\mathrm{e},i}(\boldsymbol{\theta})$ (eq. 14) 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_{Pr}, eq. 27 of Rousset and Ronce, 2004, eq. E-29 of Lehmann et al., 2016).

4.2.4 Summary

In summary, we have shown that the selection gradient on a trait, $s_i(\boldsymbol{\theta})$, depends on how a trait-change in a focal individual affects: (1) its own fitness and the fitness of its current relatives through direct intra-generational interactions ($s_{\mathrm{w},i}(\boldsymbol{\theta})$, eq. 16); and (2) the fitness of its downstream relatives living in the focal patch through heritable modifications of the ecological environment ($s_{\mathrm{e},i}(\boldsymbol{\theta})$, eqs. 17-18). This reveals that under limited dispersal, selection on intra- and inter-species interactions can generally be interpreted in terms of inter-generational inclusive fitness effects, i.e., 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 inter-generational effects and their roles in the moulding of functional traits. We illustrate more concretely the potential importance of inter-generational effects when we apply our approximation to specific models in the next section.

5 Applications

Here, we use our approximation to study the evolution of two traits that underlie intra- and inter-species interactions under limited dispersal. The first is the evolution of helping within species, which has received considerable scientific attention. This will allow us to contextualize our approach to study intra-specific 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.

5.1 Helping within a species

5.1.1 The biological scenario

We focus on a single species and study the evolution of a social behavior that increases the fitness of patch neighbours, but comes at a fitness cost to self. We consider the following life cycle. (1) Adults reproduce. A focal individual has mean fecundity $f(\tau_\bullet, \tilde{\tau})/(1+\gamma n)$, where $f(\tau_\bullet, \tilde{\tau})$ is its fecundity in the absence of density-dependent competition. The latter has intensity γ. Maximal fecundity $f(\tau_\bullet, \tilde{\tau})$ decreases with the level $\tau_\bullet$ of helping of the focal individual, but increases with the average level $\tilde{\tau}$ of helping among its neighbours in the focal patch ($\partial f(\tau_\bullet, \tilde{\tau})/\partial \tau_\bullet < 0$ and $\partial f(\tau_\bullet, \tilde{\tau})/\partial \tilde{\tau} > 0$). (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 that the transformation of these resources into offspring depends on the level of helping within the patch (for instance because individuals share resources).

5.1.2 Necessary components

We first specify the components necessary to compute the selection gradient (i.e., the terms that appear in eqs. 14–21). According to the life-cycle assumptions for the model of helping, the fitness of a focal individual that expresses a level of helping $\tau_\bullet$ in a patch of size n, when its average neighbour expresses level $\tilde{\tau}$, is

$$w(\tau_\bullet, \tilde{\tau}, n, \hat{n}(\theta)) = \frac{f(\tau_\bullet, \tilde{\tau})}{1+\gamma n}. \tag{24}$$

Note that here, fitness does not depend on species abundance in patches other than the focal ($\hat{n}(\theta)$). This is because we have assumed that competition occurs locally for material resources (see eq. C.1 in the appendix for an example of a fitness function that depends on $\hat{n}(\theta)$). Following the same argument used to derive eq. (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 generation was n and that other patches are at equilibrium $\hat{n}(\theta)$) can be written as

$$F(\bar{\tau}, n, \hat{n}(\theta)) = (1-m)\left(\frac{f(\bar{\tau}, \bar{\tau})}{1+\gamma n}\right) n + m\hat{n}(\theta). \tag{25}$$

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

$$\hat{n}(\theta) = \frac{f(\theta,\theta)-1}{\gamma}. \tag{26}$$

This equilibrium population size further allows us to obtain the pairwise relatedness $r(\theta)$, which is given by substituting eq. (26) into eq. (10), i.e.,

$$r(\theta) = \frac{(1-m)^2}{1+\left[1-(1-m)^2\right]\left[(f(\theta,\theta)-1)/\gamma-1\right]}. \tag{27}$$

This shows that pairwise relatedness increases as intra-specific competition γ increases because this leads to smaller patch size (eq. 26). As expected, relatedness increases as dispersal becomes limited ($m \to 0$). From here, it is straightforward to obtain the other necessary relatedness coefficient, $\bar{r}(\theta)$ (see eq. 20 for definition).

5.1.3 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 $s(\theta) = s_{\mathrm{w}}(\theta) + s_{\mathrm{e}}(\theta)$, where $s_{\mathrm{w}}(\theta)$ captures the intra-generational effects, and $s_{\mathrm{e}}(\theta)$, the inter-generational indirect effects.

Intra-generational effects of helping. Let us first study selection on helping according to its intra-generational effects (i.e., by looking at $s_{\mathrm{w}}(\theta)$, eq. 15). These effects can be expressed as

$$\begin{aligned} s_{\mathrm{w}}(\theta) &= \frac{\partial w(\tau_\bullet, \tilde{\tau}, n, \hat{n})}{\partial \tau_\bullet} + \frac{\partial w(\tau_\bullet, \tilde{\tau}, n, \hat{n})}{\partial \tilde{\tau}} r(\theta) \\ &= \frac{1}{f(\theta,\theta)} \left(\frac{\partial f(\tau_\bullet, \tilde{\tau})}{\partial \tau_\bullet} + \frac{\partial f(\tau_\bullet, \tilde{\tau})}{\partial \tilde{\tau}} r(\theta) \right), \end{aligned} \tag{28}$$

where we used eqs. (24) and (26). Note that since helping is individually costly but increases the fecundity of neighbours, the direct and indirect fitness effects of helping are negative and positive, respectively (i.e., $\partial w/\partial \tau_\bullet < 0$ and $\partial w/\partial \tilde{\tau} > 0$). Hence, the helping trait in our model is altruistic *sensu* evolutionary biology (e.g., Hamilton, 1964, Rousset, 2004, West and Gardner, 2010). In line with many previous models, eq. (28) reveals that altruistic helping is favoured by high relatedness. From the relatedness eq. (27), we therefore expect limited dispersal and intra-specific competition to favour the evolution of helping, owing to its intra-generational effects. However, selection on helping also depends on its inter-generational effects, which we investigate in the next paragraph.

Inter-generational effects of helping. When a single species is under scrutiny, selection on inter-generational effects (i.e., $s_{\mathrm{e}}(\theta)$, eq. 16) can be expressed as

$$s_{\mathrm{e}}(\theta) = \frac{\partial \overline{n}}{\partial \tau} \times \frac{\partial w(\tau_\bullet, \tilde{\tau}, n, \hat{n})}{\partial n}. \tag{29}$$

Using eq. (21), the effect of helping on the lineage-experienced equilibrium abundance can be written as

$$\begin{aligned} \frac{\partial \overline{n}}{\partial \tau} &= \frac{\partial F(\bar{\tau}, n, \hat{n})}{\partial \bar{\tau}} (1-m) \bar{r}(\theta) \left[1 - (1-m) \frac{\partial F(\bar{\tau}, n, \hat{n})}{\partial n} \right]^{-1} \\ &= r(\theta) \frac{[f(\theta,\theta) - 1]/\gamma}{f(\theta,\theta) - (1-m)^2} \left(\frac{\partial f(\tau_\bullet, \tilde{\tau})}{\partial \tau_\bullet} + \frac{\partial f(\tau_\bullet, \tilde{\tau})}{\partial \tilde{\tau}} \right). \end{aligned} \tag{30}$$

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

$$\frac{\partial w(\tau_\bullet, \tilde{\tau}, n, \hat{n})}{\partial n} = -\frac{\gamma}{f(\theta,\theta)} < 0 \tag{31}$$

(using eq. 24). This is because greater abundance leads to stronger intra-specific competition (according to γ). As a result, the selective inter-generational fitness effects of helping,

$$s_{\mathrm{e}}(\theta) = -\frac{r(\theta)}{f(\theta,\theta)} \underbrace{\left(\frac{f(\theta,\theta)-1}{f(\theta,\theta)-(1-m)^2}\right)}_{=\beta(\theta)>0} \left(\frac{\partial f(\tau_\bullet,\tilde{\tau})}{\partial \tau_\bullet} + \frac{\partial f(\tau_\bullet,\tilde{\tau})}{\partial \tilde{\tau}}\right) < 0, \tag{32}$$

are negative (found by substituting eqs. 30-31 into 29).

Balance between intra- and inter-generational effects. Summing eqs. (28) and (32), we find that the selection gradient is proportional to

$$s(\theta) \propto \frac{\partial f(\tau_\bullet,\tilde{\tau})}{\partial \tau_\bullet} + \kappa(\theta)\frac{\partial f(\tau_\bullet,\tilde{\tau})}{\partial \tilde{\tau}}, \tag{33}$$

where

$$\kappa(\theta) = \frac{(1-m)^2}{\hat{n}(\theta) f(\theta,\theta) - (1-m)^2(\hat{n}(\theta)-1)} = r(\theta)\frac{1-\beta(\theta)}{1-r(\theta)\beta(\theta)}, \tag{34}$$

is an effective (or scaled) relatedness coefficient, which decreases with dispersal (m, see Fig. 4 and eq. 32 for the definition of $\beta(\theta) > 0$). In the right hand side of eq. (34), the relatedness coefficient $r(\theta)$ in the numerator reflects selection on helping due to its positive intra-generational indirect effects (eq. 28). This positive effect, however, is discounted by a factor $[1-\beta(\theta)]/[1-r(\theta)\beta(\theta)] < 1$, due to the negative inter-generational indirect effects of helping (eq. 32). The effective relatedness coefficient $\kappa(\theta)$ thus reflect how selection on helping depends on the balance between the positive intra-generational indirect effects of helping, and its negative inter-generational indirect effects owing to increased competition.

To understand the balance between intra- and inter-generational effects better, it is noteworthy that relatedness among individuals decreases with the number of generations that separate them (eq. 20). As a result, selection on fitness effects becomes increasingly weak over generations. This is reflected in the fact that the effective relatedness coefficient is non-negative (i.e., $\kappa(\theta) \geq 0$, eq. 34, 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

$$f(\tau_\bullet,\tilde{\tau}) = f_{\mathrm{b}} \times (1 - C\tau_\bullet^2 + B\tilde{\tau}), \tag{35}$$

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 eq. (35) into eqs. (33)-(34) and solving $s(\theta^*) = 0$ allows us to find the singular strategy θ^*. When both C and B are small (of the order order of a parameter $\epsilon \ll 1$), the singular strategy θ^* 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,

$$\theta^* = \frac{B}{2C}\kappa(\theta) = \frac{B}{2C} \times \frac{1}{1+\hat{n}(0)\left(\frac{f_\mathrm{b}}{(1-m)^2}-1\right)} \geq 0. \tag{36}$$

Eq. (36) makes it straightforward to see that helping can evolve in spite of its negative inter-generational indirect effects. It further shows how the equilibrium level of helping, θ^*, decreases with dispersal m and local abundance in the absence of helping, $\hat{n}(0) = (f_\mathrm{b}-1)/\gamma$.

More generally, by solving eqs. (33) and (34) with eq. (35) numerically, we find that predictions generated from our approximation fits qualitatively and quantitatively well with observations from individual-based simulations, as much for the value of the convergence stable level of helping θ^* (Fig. 5, top panel) as for the concomitant equilibrium group size $\hat{n}(\theta^*)$ this generates (Fig. 5, bottom panel).

5.1.4 Connections to previous results on the evolution of altruism

Our finding that altruism decreases as dispersal and local abundance increase is a standard result of evolutionary biology. 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, El Mouden and Gardner, 2008, Lion and Gandon, 2009, 2010, Rodrigues and Gardner, 2012, Johnstone and Cant, 2008, Wild, 2011, Bao and Wild, 2012, Johnstone et al., 2012, Kuijper and Johnstone, 2012). First, the vast majority of previous analyses assumes that density-dependent competition occurs for space after dispersal (i.e., space or "open sites" is the limiting factor, e.g., Tilman, 1982, Chapter 8). In this case, intra-generational kin competition effects strongly inhibit the benefits of interacting among relatives (e.g., Taylor, 1992). Here, we have assumed that competition occurs for resources before dispersal. In this situation, we found that intra-generational kin competition effects do not abate the selective advantage of interacting with relatives (this can be seen from eq. 28, which only depends on pairwise relatedness). Rather, by increasing abundance, altruism increases kin competition for future generations (eq. 32). 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 and previous models is that the latter have had to rely exclusively on numerical approaches to compute the selection gradient (in the island model of dispersal with endogenous patch dynamics, e.g., Lehmann et al., 2006, Alizon and Taylor, 2008, Wild et al., 2009, Wild, 2011 for models of altruism evolution, and Metz and Gyllenberg, 2001, Cadet et al., 2003, Parvinen et al., 2003, Rousset and Ronce, 2004 for models of dispersal evolution). This reliance on numerical analysis has made it difficult to understand

how selection on altruism varies 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. 33), which nonetheless fits well with simulation results (Fig. 5).

It is noteworthy that the selection gradient we have derived for this example (eq. 33) applies to any type of social interactions within the life-cycle given in 5.1.1. In fact, the selection gradient eq. (33) can be adjusted to study other social behaviours simply by changing the fecundity function (e.g., eq. 35). Such selection gradient written in terms of marginal fecundity effects of behaviour 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 is ignored here).

5.1.5 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 behaviour (e.g., Le Galliard et al., 2005, Purcell et al., 2012, Mullon et al., 2017). 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).

In order to test whether our approximation could capture the interplay between social behaviour, 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 analysis of selection are given in Appendix C.

We find that dispersal increases as survival during dispersal, s, increases (Fig. 6, top panel, grey 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., 2017). 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 that the approximate selection gradient (eqs. 14–17) can be used to model dispersal evolution, in particular when local demography, genetic structure, and social traits feedback 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., 2017). Such an investigation would require studying the second-order effects of selection, which is beyond the scope of this paper.

5.2 Harming between species

5.2.1 The biological scenario

In order to illustrate how the 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 $f_1(\tau_{\bullet,1})/(1+\gamma n_1+\eta n_2)$, which decreases with intra- and inter-specific competition (respectively measured by parameters γ and η). The maximal fecundity of a focal individual of species 1, $f_1(\tau_{\bullet,1})$, decreases with its investment $\tau_{\bullet,1}$ into harming (i.e., $f_1'(\tau_{\bullet,1})<0$). 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 γ is the level of intra-specific competition. Note that only species 1 experiences inter-specific competition. This would occur, for instance, because species 2 is a generalist consumer while species 1 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., $D'(\bar{\tau}_1)>0$). 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, Inderjit et al., 2011). (3) All adults of both species die. (4) Surviving offspring of each species disperse with probability m and all survive to adulthood.

5.2.2 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 $\tau_{\bullet,1}$ into harming in a patch with $\boldsymbol{n}=(n_1,n_2)$ individuals of species 1 and 2 respectively, has fitness

$$w_1(\tau_{\bullet,1},\boldsymbol{n},\hat{\boldsymbol{n}}(\boldsymbol{\theta}))=\frac{f_1(\tau_{\bullet,1})}{1+\gamma n_1+\eta n_2}. \tag{37}$$

The abundances of both species in the focal patch after one iteration of the life-cycle, given (1) that the average level of harming in the patch is $\bar{\tau}_1$; (2) that the abundance at the previous generation was $\boldsymbol{n}=(n_1,n_2)$; and (3)

that other patches are at equilibrium $\hat{\boldsymbol{n}}(\boldsymbol{\theta}) = (\hat{n}_1(\theta_1), \hat{n}_2(\theta_1))$, are given by

$$\begin{aligned} F_1(\bar{\tau}_1, \boldsymbol{n}, \hat{\boldsymbol{n}}(\boldsymbol{\theta})) &= (1-m)\left(\frac{f_1(\bar{\tau}_1)}{1+\gamma n_1 + \eta n_2}\right) n_1 + m\hat{n}_1(\theta_1) \\ F_2(\bar{\tau}_1, \boldsymbol{n}, \hat{\boldsymbol{n}}(\boldsymbol{\theta})) &= (1-m)\left(\frac{f_2}{1+\gamma n_2} - D(\bar{\tau}_1)\frac{n_1}{n_2}\right) n_2 + m\hat{n}_2(\theta_1). \end{aligned} \tag{38}$$

The resident ecological equilibrium, which is found by solving $\hat{\boldsymbol{n}}(\boldsymbol{\theta}) = (F_1(\theta_1, \hat{\boldsymbol{n}}(\boldsymbol{\theta}), \hat{\boldsymbol{n}}(\boldsymbol{\theta})), F_2(\theta_1, \hat{\boldsymbol{n}}(\boldsymbol{\theta}), \hat{\boldsymbol{n}}(\boldsymbol{\theta})))$ simultaneously, is too complicated to be presented here for the general case. Note, however, that when the resident level θ_1 of harming is small, a first-order Taylor expansion of the resident ecological equilibrium around $\theta_1 = 0$ gives

$$\begin{aligned} \hat{n}_1(\theta_1) &= \hat{n}_1(0) + \frac{\theta_1}{\gamma}\left(\frac{f_2}{f_2-1}\hat{n}_1(0)\eta D'(0) + f_1'(0)\right) + O\left(\theta_1^2\right) \\ \hat{n}_2(\theta_1) &= \hat{n}_2(0) - \theta_1\left(\frac{f_2}{f_2-1}\hat{n}_1(0)D'(0)\right) + O\left(\theta_1^2\right), \end{aligned} \tag{39}$$

where

$$\begin{aligned} \hat{n}_1(0) &= \frac{f_1-1}{\gamma} - \frac{\eta}{\gamma}\hat{n}_2(0) \\ \hat{n}_2(0) &= \frac{f_2-1}{\gamma}, \end{aligned} \tag{40}$$

are the abundances in the absence of harming (assuming that $D(0) = 0$). Eq. (39) reveals that harming of species 2 reduces its abundance (i.e., $\hat{n}_2(\theta_1) \leq \hat{n}_2(0)$). Species 1, however, may increase or decrease in abundance according to the degree of interspecific competition, η, and the cost of harming (which is captured by the derivative $f_1'(0) < 0$). The final components necessary to derive the selection gradient is the relatedness coefficient for species 1. It is given by eq. (10), with the resident ecological equilibrium for species 1, $\hat{n}_1(\theta_1)$, which is defined by eq. (38).

5.2.3 Selection on harming

Intra-generational effects. Substituting eq. (37) into eq. (15), we obtain that selection on the intra-generational effects of harming,

$$s_{\mathrm{w},1}(\theta_1) = \frac{f_1'(\theta_1)}{f_1(\theta_1)} < 0, \tag{41}$$

are always negative. This is because harming is intra-generationally costly to express at the individual level and does not provide any intra-generational indirect fitness benefits. Hence, the only way for harming to evolve in this model is if this intra-generational cost is compensated by future benefits received by downstream relatives, which we investigate below.

Inter-generational effects. Selection on harming due to its effects on the fitness of downstream relatives is captured by the inter-generational part of the selection gradient eq. (16). Substituting eqs. (37)-(39) into eqs. (16) and (21), we find that selection on harming due to its inter-generational effect is given by

$$s_{\mathrm{e},1}(\theta_1) = \left(\frac{1-m}{f_1(\theta_1)}\right)^2 \left(\eta f_1(\theta_1) D'(\theta_1) - \gamma f_1'(\theta_1)\right) + O\left((1-m)^4\right). \tag{42}$$

From eq. (42), we see that inter-generational effects favour harming (i.e., $s_{\mathrm{e},1}(\theta_1) > 1$ since $D'(\theta_1) > 0$ and $f_1'(\theta_1) < 0$). This is because harming decreases the abundance of species 2 in future generations, which in turn increases the fitness of downstream relatives by alleviating them from competition. In line with this argument, eq. (42) shows that harming is particularly favoured when dispersal is limited and competition (within and between-species) is strong.

Convergence stable equilibrium of harming. To test explicitly 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 $\tau_{\bullet,1}$ of harming is,

$$f_1(\tau_{\bullet,1}) = f_{1,\mathrm{b}} \times \left(1 - C\tau_{\bullet,1}^2\right) \tag{43}$$

where $f_{1,\mathrm{b}}$ 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,

$$D(\bar{\tau}_1) = \alpha \bar{\tau}_1, \tag{44}$$

where α is a parameter tuning the deleteriousness of harming. The convergence stable level of harming, which is found by solving $s_1(\theta_1^*) = s_{\mathrm{w},1}(\theta_1^*) + s_{\mathrm{e},1}(\theta_1^*) = 0$ for θ_1^*, are shown in Figure 7 as a function of dispersal.

In line with eq. (42), we find that harming does not evolve when dispersal is complete ($m = 1$, Figure 7, top panel). This is because in that case, downstream relatives can never benefit from a decrease of inter-species competition owing to harming. As dispersal becomes limited, this intergenerational benefit increasingly goes to relatives so that harming evolves (Figure 7, top panel). This evolution, in turn, causes a significant reduction in the abundance of species 2 and an increase of species 1 (Figure 7, bottom panel). These results were confirmed using individual-based simulations, which further supports the goodness of fit of our approximation (Figure 7).

6 Discussion

Due to the physical limitations of movement, a community of species is typically structured in space to form a meta-community (e.g., Tilman, 1982, Clobert et al., 2001, Urban et al., 2008). So far, understanding selection in such a meta-community has proven challenging due to the feedback between local ecology and trait composition that emerges under limited dispersal and stochastic demography. In order to better understand these eco-evolutionary dynamics, we have derived 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 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 ten individuals (Fig. 1 and Figs. 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 $m_i(\boldsymbol{\theta}) = 0.1$ and $\hat{n}_i(\boldsymbol{\theta}) = 10$, eq. 10 gives $F_{\mathrm{ST}} = 0.30$ for haploids; for diploids, eq. 6.23 of Hartl and Clark, 2007, gives $F_{\mathrm{ST}} = 0.20$; note that equivalently, this regime entails one migrant every generation, i.e., $\hat{n}_i(\boldsymbol{\theta})m_i(\boldsymbol{\theta}) = 1$; 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 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- (eq. 15) and inter-generational effects (eqs. 16–17). Inter-generational 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 downstream generations due to ecological interactions (eq. 17, Figure 3). These changes then feedback on the fitness individuals living in downstream generations, who potentially carry genes that are identical-by-descent to the focal (i.e., who are kin). In other words, inter-generational 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 multi-species interactions, our model generalises previous models of two-species evolution in the presence of relatives that ignored trait-induced demographic

changes, either altogether (Frank, 1994, Foster and Wenseleers, 2006, Wyatt et al., 2013, Akçay, 2017) or in the evolving species (Lehmann, 2008).

Remarkably, the eco-evolutionary, inter-generational, feedbacks that emerge in our model are captured mathematically by analogues of press perturbations (eqs. 21-22). 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 one due to ecological interactions (Yodzis, 1989, Case, 2000). Here, 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 relatives experience it (Figure 3). Because this time increases as dispersal becomes more limited, inter-generational 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-centred perspective: because limited dispersal ties in the fate of a trait-changing mutation with its inter-generational ecological effects, these effects become more important for how selection targets this mutation.

One clear prediction that emerges from our analysis is that traits within communities that are strongly limited spatially are more likely to have long-term effects that have been shaped by selection. Plant communities would be ideal to test this notion. For instance, many plants are engaged in inter-specific chemical warfare, with lasting effects on soil composition (Inderjit et al., 2011). In the light of our results, it would be interesting to study how these inter-generational effects vary with the degree of dispersal (or gene flow), which can be estimated from Fst values.

We applied our model to the evolution of two behavioural traits with demographic and ecological consequences, and in line with the above considerations, we found an important influence of inter-generational effects due to limited dispersal. First, we studied the evolution of altruistic helping within species. Our model follows a rich literature on this topic (Eshel, 1972, Rogers, 1990, Taylor, 1992, Frank, 1998, Rousset, 2004, Gardner and West, 2006, Lion and Gandon, 2009, for e.g. in patch-structured populations). This literature 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 intra-generational cost of kin competition (e.g., Taylor, 1992, Gardner and West, 2006, Lion and Gandon, 2009). By contrast, we assumed that competition occurs for material resources before dispersal and that helping can influence patch size. In this case, the cost of kin competition are delayed and paid by downstream generations. Because these inter-generational costs are weaker than intra-generational ones, we found that selection favours altruism for a large range of parameters, in particular when dispersal is limited (Figure 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 costly harming between species. We found that harming evolution strongly depends on inter-generational effects. Specifically, we found that harming evolves when it alleviates inter-species competition for downstream relatives by reducing the other species' abundance. This 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 (Figure 7). Previous theory has focused on understanding how "altruism" or mutualism between species can evolve (Frank, 1994, Wyatt et al., 2013), ignoring the influence of behaviour on abundance. These studies have highlighted 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 inter-specific competition.

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 (like when populations are close to extinction). In fact, it would be useful to analyse our model with greater mathematical rigour in order to obtain a sharper understanding of the conditions under which ecological stochasticity can be neglected (for instance 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 meta-community stochastic demography, precludes the consideration of limit cycles or spatio-temporal 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, individual fitness by individual growth rates, and calculating inter-temporal relatedness coefficients in continuous time, e.g., Sozou, 2009).

To conclude, our heuristic approximation is a step further towards the integration of multi-species 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 understand how selection moulds intra- and inter-species interactions when dispersal is limited.

Acknowledgements

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 a previous version of this paper. CM is funded by Swiss NSF grant PP00P3-123344 to LL.

Appendix A: Individual-based simulations

Here, we describe how we implemented the individual-based simulations for the life cycle described in section 2.1 (using Mathematica 10.0.1.0, Wolfram Research, 2016, the code is available on request from the corresponding author).

The simulated population is divided among a finite number $n_{\mathrm{d}} = 1000$ of patches (instead of an infinite number of patches like in the analytical model). At the beginning of a generation, each individual j of species i is characterized by its phenotype θ_{ji} taken from the type space Θ_i. Calculations proceed as follows. (1) We first calculate the mean fecundity of each individual according to its phenotype and possibly that of its neighbours (depending on the model of interaction). Each individual then produces a Poisson distributed number of offspring according to its mean fecundity. (2) Each offspring disperses according to a Bernoulli trial (whose mean may depend on phenotype if dispersal evolves). During dispersal, death occurs according to a Bernoulli trial (whose mean may depend on evolving phenotypes). If an offspring disperses and survives dispersal, it is allocated to a non-natal patch according to discrete uniform distribution. (3) If local density-dependent competition takes place, each offspring survives according to a Bernoulli trial with mean depending on the number of individuals entering competition in the patch. (4) Mutation occurs in each offpsring with probability μ (Bernouilli trial). If no mutation occurs, an offspring has the same phenotypic values as its parent. If a mutations occurs, we add a perturbation to the parental phenotypic value that is sampled from a multivariate Normal distribution with mean zero for each trait, variance σ^2 and no covariance among traits. The resulting phenotypic values are controlled to remain in a given range if necessary.

The model tracks the phenotype(s) θ_{ji} of each individual of all species, and the number of individuals in each patch. From these data, we can evaluate various statistics such as mean patch size, average trait value, and relatedness. Relatedness under the neutral model (i.e., when vital rates are the same within each species, see Fig. 2) is calculated as the ratio of average phenotypic covariance among individuals of the same patch (averaged over all patches) to the total phenotypic variance in the population. This gives the probability of identity-by-descent within patches (Frank, 1998, eqs. 3.9-3.10).

Appendix B: Evolutionary analysis

In this appendix, we derive the approximate selection gradient on a mutant τ_i in species i (eqs. 14–22 of the main text).

B.1 Local evolutionary invasion analysis

Under the full stochastic model (life cycle described in section 2.1), the selection gradient $S_i(\boldsymbol{\theta})$ is equal to the sensitivity of the invasion fitness $\rho_i(\tau_i, \boldsymbol{\theta})$ of a mutant allele with phenotype $\tau_i \in \mathbb{R}$ in species i in a resident population $\boldsymbol{\theta} \in \mathbb{R}$; namely

$$S_i(\boldsymbol{\theta}) = \left.\frac{\partial \rho_i(\tau_i, \boldsymbol{\theta})}{\partial \tau_i}\right|_{\tau_i = \theta_i} \tag{B.1}$$

(Lehmann et al., 2016, Box 2). Invasion fitness $\rho_i(\tau_i, \boldsymbol{\theta})$ is defined as the per capita number of mutant copies produced asymptotically over a time step of the reproductive process by a whole mutant lineage descending from a single initial mutant (and when the mutant is rare in the population, i.e., $\rho_i(\tau_i, \boldsymbol{\theta})$ is the geometric growth rate of the mutant, Cohen, 1979, Tuljapurkar, 1989, Caswell, 2000, Tuljapurkar et al., 2003).

Instead of using invasion fitness, $\rho_i(\tau_i, \boldsymbol{\theta})$, which is complicated to manipulate, we will use an invasion fitness proxy (i.e., a quantity that is sign equivalent to invasion fitness, e.g., Stearns, 1992, Charlesworth, 1994, Case, 2000, Metz and Gyllenberg, 2001, Lehmann et al., 2016). In the homogeneous island model (i.e., no exogenous differences among islands), one invasion fitness proxy is the basic reproductive number, $R_{0,i}(\tau_i, \boldsymbol{\theta})$, which is the expected number of successful offspring produced by a randomly sampled individual from a local mutant lineage of species i during the (expected) sojourn time of this lineage in a single patch (and when the mutant is rare in the population, see Lehmann et al., 2016, eq. 14 for a formal definition of $R_{0,i}(\tau_i, \boldsymbol{\theta})$).

To see that the selection gradient can be inferred from $R_{0,i}(\tau_i, \boldsymbol{\theta})$, consider a Taylor expansion of $\rho_i(\tau_i, \boldsymbol{\theta})$ and $R_{0,i}(\tau_i, \boldsymbol{\theta})$ around $\delta_i = \tau_i - \theta_i = 0$,

$$\begin{aligned} \rho_i(\tau_i, \boldsymbol{\theta}) &= 1 + \delta_i \left.\frac{\partial \rho_i(\tau_i, \boldsymbol{\theta})}{\partial \tau_i}\right|_{\tau_i = \theta_i} + O_\rho\left(\delta_i^2\right) \\ R_{0,i}(\tau_i, \boldsymbol{\theta}) &= 1 + \delta_i \left.\frac{\partial R_{0,i}(\tau_i, \boldsymbol{\theta})}{\partial \tau_i}\right|_{\tau_i = \theta_i} + O_{R_0}\left(\delta_i^2\right), \end{aligned} \tag{B.2}$$

where $O_\rho\left(\delta_i^2\right)$ and $O_{R_0}\left(\delta_i^2\right)$ are second-order remainders. Since $\rho_i(\tau_i, \boldsymbol{\theta})$ and $R_{0,i}(\tau_i, \boldsymbol{\theta})$ are sign equivalent around one (i.e., $\rho_i(\tau_i, \boldsymbol{\theta}) \leq 1 \iff R_{0,i}(\tau_i, \boldsymbol{\theta}) \leq 1$, Lehmann et al., 2016, eq. 15), eq. (B.2) reveals that the sensitivities of $\rho_i(\tau_i, \boldsymbol{\theta})$ and $R_{0,i}(\tau_i, \boldsymbol{\theta})$, are proportional, i.e.,

$$S_i(\boldsymbol{\theta}) = \left.\frac{\partial \rho_i(\tau_i, \boldsymbol{\theta})}{\partial \tau_i}\right|_{\tau_i = \theta_i} = \epsilon(\boldsymbol{\theta}) \left.\frac{\partial R_{0,i}(\tau_i, \boldsymbol{\theta})}{\partial \tau_i}\right|_{\tau_i = \theta_i}, \tag{B.3}$$

where $\epsilon(\boldsymbol{\theta}) > 0$ is a positive factor of proportionality that depends only on the resident trait $\boldsymbol{\theta}$ (since the left and right hand side of eq. B.3 must only depend on $\boldsymbol{\theta}$). The sensitivity of $R_{0,i}(\tau_i, \boldsymbol{\theta})$ is therefore sufficient to determine the sign of the selection gradient, in line with previous results derived for class-structured population in continuous time (Metz and de Kovel, 2013, differentiating their eq. 2.4 produces our eq. B.3, see also Metz

and Leimar, 2011, p.175; but note that these authors use the average lifespan of an individual as a timescale to evaluate fitness, whereas we use one life cycle iteration).

B.2 Basic reproductive number

We now specify the basic reproductive number for our deterministic approximation, i.e., when local abundance change is deterministically given by eq. (11). In order to distinguish among the exact basic reproductive number, $R_{0,i}(\tau_i, \boldsymbol{\theta})$, and its approximation, we will denote the latter by $\tilde{R}_{0,i}(\tau_i, \boldsymbol{\theta})$.

We first let $M_{i,t} \in I_{i,t} = \{0, 1, 2, ..., n_{i,t}\} \subset \mathbb{N}_+$ be a random variable for the number of mutant individuals of species i at generation $t = 0, 1, 2, ...$ in the focal patch (assuming that the mutant arose as a single copy in that patch a generation $t = 0$). Note that even if the one-generational change in abundance is assumed to be deterministic in our model, its dynamics are stochastic when there is more than one allele in the focal patch. This is because the one-generational change in abundance depends on the stochastic genetic composition in the focal patch, i.e., on the random variable $M_{i,t}$ (see eq. 11). We write $\boldsymbol{n}_t(\boldsymbol{\tau}_i) \in \mathcal{N}_{i,t}$ as the random variable for the community vector in the focal patch at time t, where $\mathcal{N}_{i,t} \subset \mathbb{R}_+^S$ denotes the state space of the community demography at time t when we consider a mutant in species i. We assume that the community vector is bounded (i.e., $n_{i,t} < \infty$ for all i). The couple $\{M_{i,t}, \boldsymbol{n}_t(\boldsymbol{\tau}_i)\}_{t \geq 0}$ is then a random dynamical system that occurs on $\Omega_{i,t} = I_{i,t} \times \mathcal{N}_{i,t}$, which is the state space of the genetic-demographic process in the focal patch at time t, conditional on there being at least one mutant in the patch (conditional on $M_{i,t} \geq 1$). We denote by $\Pr(\boldsymbol{s}_{i,t})$ the probability that the realized state at time t is $\boldsymbol{s}_{i,t} = (M_{i,t}, \boldsymbol{n}_t(\boldsymbol{\tau}_i)) \in \Omega_{i,t}$.

Next, we observe that in each generation, $M_{i,t}$ has a single absorbing state: the local extinction of the lineage ($M_{i,t} = 0$). This can be understood by observing that because we have assumed that all offspring have a non-zero dispersal probability there can be no immigration of mutants into the focal group when the mutant is rare in the population (e.g., Metz and Gyllenberg, 2001, Lehmann et al., 2016). Hence, the lineage of the mutant will eventually go extinct locally, i.e., $\lim_{t\to\infty} \Pr(M_{i,t} = 0) = 1$, and as a consequence, $\lim_{t\to\infty} \Pr(\boldsymbol{s}_{i,t} = 0) = 1$ (for a product space like $\Omega_{i,t}$, the finite dimensional distribution $\Pr(\boldsymbol{s}_{i,t})$ is defined for any realization $\boldsymbol{s}_{i,t} \in \Omega_{i,t}$, and it can be extended to events in $\Omega_i = \lim_{t\to\infty} \Omega_{i,t}$, Meyn and Tweedie, 2009, p. 60).

With the above definitions, we can write the basic reproductive number for our deterministic approximation of population dynamics as

$$\tilde{R}_{0,i}(\tau_i, \boldsymbol{\theta}) = \sum_{t=0}^{\infty} \sum_{\boldsymbol{s}_t \in \Omega_t} w_i(\tau_i, \boldsymbol{s}_t) q_i(\boldsymbol{s}_t), \tag{B.4}$$

where $w_i(\tau_i, \boldsymbol{s}_t)$ is the individual fitness of an individual (i.e., the expected number of successful offspring produced by this individual over one life cycle iteration, possibly including itself through survival) of species i that

carries mutant phenotype τ_i, when the patch is in a genetic-demographic state $\boldsymbol{s}_t$ at time t. Further,

$$q_i(\boldsymbol{s}_t) = \frac{M_{i,t}\Pr(\boldsymbol{s}_t)}{N_{\mathrm{L}}(\tau_i,\boldsymbol{\theta})}, \tag{B.5}$$

is defined as the probability that a randomly sampled mutant lineage member is sampled exactly at time t on realization $\boldsymbol{s}_t$ of the stochastic process ($N_{\mathrm{L}}(\tau,\theta)$ is the expected total size of the mutant lineage over its lifetime in a single patch: $N_{\mathrm{L}}(\tau_i,\boldsymbol{\theta}) = \sum_{t=0}^{\infty}\sum_{\boldsymbol{s}_t\in\Omega_t} M_{i,t}\Pr(\boldsymbol{s}_t)$). Note that owing to the fact that extinction of the mutant lineage in the focal patch is certain ($\lim_{t\to\infty}\Pr(M_{i,t}=0)=1$), the infinite sum in eq. (B.4) always converges.

B.3 Selection gradient

We now derive the sensitivity,

$$s_i(\boldsymbol{\theta}) = \frac{\partial \tilde{R}_{0,i}(\tau_i,\boldsymbol{\theta})}{\partial \tau_i}, \tag{B.6}$$

which determines the approximate selection gradient presented in the main text (eq. 14). First, we write individual fitness explicitly in terms of the genetic and demographic processes as

$$w_i(\tau_i,\boldsymbol{s}_t) = w_i(\tau_i,\tilde{\boldsymbol{\tau}}_{i,t},\boldsymbol{n}_t,\hat{\boldsymbol{n}}(\boldsymbol{\theta})), \tag{B.7}$$

where $\tilde{\boldsymbol{\tau}}_{i,t} = (\theta_1,\ldots,\theta_{i-1},\tilde{\tau}_{i,t},\theta_{i+1},\ldots,\theta_S)$ is the vector of average phenotype among the neighbours of the focal individual. In this vector, the average phenotype among the neighbours of the focal individual of the same species i at time t is given by

$$\tilde{\tau}_{i,t} = \frac{n_{i,t}(\boldsymbol{\tau}_i)\bar{\tau}_{i,t} - \tau_i}{n_{i,t}(\boldsymbol{\tau}_i) - 1}, \tag{B.8}$$

i.e., the average phenotype in species i excluding the focal. The individual fitness function on the right hand side of eq. (B.7)

$$w_i : \mathbb{R}\times\mathbb{R}^S\times\mathcal{N}_{i,t}\times\mathbb{R}_+^S \to \mathbb{R}_+, \tag{B.9}$$

maps the phenotype τ_i of that individual, the average phenotype configuration $\tilde{\boldsymbol{\tau}}_{i,t}$ among its patch neighbours, and the local $\boldsymbol{n}_t$ and global $\hat{\boldsymbol{n}}(\boldsymbol{\theta})$ population states into a positive real number of expected offspring.

To obtain the approximate selection gradient, we substitute eq (B.7) into (B.4), which is in turn substituted into (B.6). Using the chain rule, this operation yields

$$s_i(\boldsymbol{\theta}) = \sum_{t=0}^{\infty}\sum_{\boldsymbol{s}_t\in\Omega_t}\left[\frac{\partial w_i(\tau_i,\tilde{\boldsymbol{\tau}}_{i,t},\boldsymbol{n}_t,\hat{\boldsymbol{n}})}{\partial\tau_i}q_i^{\circ}(\boldsymbol{s}_t) + w_i(\theta_i,\boldsymbol{\theta},\hat{\boldsymbol{n}},\hat{\boldsymbol{n}})\frac{\partial q_{i,t}(\boldsymbol{s}_t)}{\partial\tau_i}\right] \tag{B.10}$$

where here and throughout all derivatives are evaluated at the resident value phenotypic value $\boldsymbol{\theta}$ and ecological

equilibrium $\hat{\boldsymbol{n}}(\boldsymbol{\theta})$. We use $q_i^\circ(\boldsymbol{s}_t)$ to denote the genetic-demographic state distribution in the resident population (i.e., in a monomorphic $\boldsymbol{\theta}$ population). The second summand of eq. (B.10), in fact, vanishes. This follows from the fact that,

$$\sum_{\boldsymbol{s}_t\in\Omega_t}\frac{\partial q_i(\boldsymbol{s}_t)}{\partial\tau_i}=\frac{\partial}{\partial\tau_i}\left[\sum_{\boldsymbol{s}_t\in\Omega_t}q_i(\boldsymbol{s}_t)\right]=\frac{\partial}{\partial\tau_i}(1)=0, \tag{B.11}$$

since $q_i(\boldsymbol{s}_t)$ is a probability density function over Ω_t. The selection gradient thus reduces to

$$s_i(\boldsymbol{\theta})=\sum_{t=0}^{\infty}\sum_{\boldsymbol{s}_t\in\Omega_t}\frac{\partial w_i(\tau_i,\tilde{\boldsymbol{\tau}}_{i,t},\boldsymbol{n}_t,\hat{\boldsymbol{n}})}{\partial\tau_i}q_i^\circ(\boldsymbol{s}_t). \tag{B.12}$$

In order to evaluate the fitness derivative in eq. (B.12), we use the chain rule to obtain

$$\frac{\partial w_i(\tau_i,\tilde{\boldsymbol{\tau}}_{i,t},\boldsymbol{n}_t,\hat{\boldsymbol{n}})}{\partial\tau_i}=\frac{\partial w_i(\tau_{\bullet,i},\tilde{\boldsymbol{\tau}}_i,\boldsymbol{n},\hat{\boldsymbol{n}})}{\partial\tau_{\bullet,i}}+\frac{\partial w_i(\tau_{\bullet,i},\tilde{\boldsymbol{\tau}}_i,\boldsymbol{n},\hat{\boldsymbol{n}})}{\partial\tilde{\tau}_i}\frac{\partial\tilde{\tau}_{i,t}}{\partial\tau_i}+\sum_{j=1}^{S}\frac{\partial w_i(\tau_{\bullet,i},\tilde{\boldsymbol{\tau}}_i,\boldsymbol{n},\hat{\boldsymbol{n}})}{\partial n_j}\frac{\partial n_{j,t}}{\partial\tau_i}, \tag{B.13}$$

where $w_i(\tau_{\bullet,i},\tilde{\boldsymbol{\tau}}_i,\boldsymbol{n},\hat{\boldsymbol{n}}(\boldsymbol{\theta}))$ is the fitness of a focal individual in species i with phenotype $\tau_{\bullet,i}$, when its neighbours have average phenotype given by the vector $\tilde{\boldsymbol{\tau}}_i=(\theta_1,...,\theta_{i-1},\tilde{\tau}_i,\theta_{i+1},...,\theta_S)$ (where $\tilde{\tau}_i$ is the average phenotype in species i among the neighbours of the focal individual). The fitness function on the right hand side of eq. (B.13) no longer depends explicitly on time, in contrast to the function on the left hand side (which depends on time through $\tilde{\boldsymbol{\tau}}_{i,t}$ and $\boldsymbol{n}_t$). This is owing to our use of the chain rule, which allows us to capture time dependence only in the derivatives of $\tilde{\boldsymbol{\tau}}_{i,t}$ and $\boldsymbol{n}_t$ with respect to τ_i in the right hand side of eq. (B.13). In this equation, the derivative of $\tilde{\boldsymbol{\tau}}_{i,t}$ can be specified by inserting eq. (12) into eq. (B.8), to get

$$\left.\frac{\partial\tilde{\tau}_{i,t}}{\partial\tau_i}\right|_{\substack{\tau_i=\theta_i\\ \boldsymbol{n}=\hat{\boldsymbol{n}}}}=\frac{M_{i,t}-1}{\hat{n}_i-1}. \tag{B.14}$$

Substituting eq. (B.14) into eq. (B.13), which is in turn substituted into eq. (B.12), gives

$$s_i(\boldsymbol{\theta})=\sum_{t=0}^{\infty}\sum_{\boldsymbol{s}_t\in\Omega_t}\left[\frac{\partial w_i(\tau_{\bullet,i},\tilde{\boldsymbol{\tau}}_i,\boldsymbol{n},\hat{\boldsymbol{n}})}{\partial\tau_{\bullet,i}}+\frac{\partial w_i(\tau_{\bullet,i},\tilde{\boldsymbol{\tau}}_i,\boldsymbol{n},\hat{\boldsymbol{n}})}{\partial\tilde{\tau}_i}\left(\frac{M_{i,t}-1}{\hat{n}_i-1}\right)+\sum_{j=1}^{S}\frac{\partial w_i(\tau_{\bullet,i},\tilde{\boldsymbol{\tau}}_i,\boldsymbol{n},\hat{\boldsymbol{n}})}{\partial n_j}\frac{\partial n_{j,t}}{\partial\tau_i}\right]q_i^\circ(\boldsymbol{s}_t), \tag{B.15}$$

for the selection gradient. Next, we decompose eq. (B.15) as $s_i(\boldsymbol{\theta})=s_{\mathrm{w},i}(\boldsymbol{\theta})+s_{\mathrm{e},i}(\boldsymbol{\theta})$, where

$$s_{\mathrm{w},i}(\boldsymbol{\theta})=\sum_{t=0}^{\infty}\sum_{\boldsymbol{s}_t\in\Omega_t}\left[\frac{\partial w_i(\tau_{\bullet,i},\tilde{\boldsymbol{\tau}}_i,\boldsymbol{n},\hat{\boldsymbol{n}})}{\partial\tau_{\bullet,i}}+\frac{\partial w_i(\tau_{\bullet,i},\tilde{\boldsymbol{\tau}}_i,\boldsymbol{n},\hat{\boldsymbol{n}})}{\partial\tilde{\tau}_i}\left(\frac{M_{i,t}-1}{\hat{n}_i-1}\right)\right]q_i^\circ(\boldsymbol{s}_t) \tag{B.16}$$

$$s_{\mathrm{e},i}(\boldsymbol{\theta})=\sum_{t=0}^{\infty}\sum_{\boldsymbol{s}_t\in\Omega_t}\left[\sum_{j=1}^{S}\frac{\partial w_i(\tau_{\bullet,i},\tilde{\boldsymbol{\tau}}_i,\boldsymbol{n},\hat{\boldsymbol{n}})}{\partial n_j}\frac{\partial n_{j,t}}{\partial\tau_i}\right]q_i^\circ(\boldsymbol{s}_t). \tag{B.17}$$

We will now show that $s_{\mathrm{w},i}(\boldsymbol{\theta})$ as defined by eq. (B.16) reads as eq. (15) of the main text (which captures the intra-generational effects), and $s_{\mathrm{e},i}(\boldsymbol{\theta})$ as defined by eq. (B.17) reads as eqs. (16)–(22) of the main text (which captures

the inter-generational effects).

B.3.1 Intra-generational effects, $s_{\mathrm{w},i}(\boldsymbol{\theta})$

To show that eq (B.16) equals eq. (15), we simply use the fact that $\sum_{\boldsymbol{s}_t \in \Omega_t} q_i^\circ(\boldsymbol{s}_t) = 1$, so that

$$s_{\mathrm{w},i}(\boldsymbol{\theta}) = \frac{\partial w_i(\tau_{\bullet,i}, \tilde{\boldsymbol{\tau}}_i, \boldsymbol{n}, \hat{\boldsymbol{n}})}{\partial \tau_{\bullet,i}} + \frac{\partial w_i(\tau_{\bullet,i}, \tilde{\boldsymbol{\tau}}_i, \boldsymbol{n}, \hat{\boldsymbol{n}})}{\partial \tilde{\tau}_i} r_i(\boldsymbol{\theta}) \tag{B.18}$$

where

$$r_i(\boldsymbol{\theta}) = \sum_{t=0}^{\infty} \sum_{\boldsymbol{s}_t \in \Omega_t} \frac{M_{i,t} - 1}{\hat{n}_i - 1} q_i^\circ(\boldsymbol{s}_t). \tag{B.19}$$

From the definition of $q_i^\circ(\boldsymbol{s}_t)$ (eq. B.5), $r_i(\boldsymbol{\theta})$ as defined by eq. (B.19) is the probability that two randomly sampled individuals from species i in the same patch carry an allele identical-by-descent, i.e., the pairwise coefficient of relatedness. Hence, eq. (B.18) corresponds to eq. (15), as required.

B.3.2 Inter-generational effects, $s_{\mathrm{e},i}(\boldsymbol{\theta})$

Here, we derive eqs. (16)–(22), which underpin selection on a trait due to its inter-generation effects.

Lineage-experienced abundance. We first write eq. (B.17) as eq. (16) of the main text,

$$s_{\mathrm{e},i}(\boldsymbol{\theta}) = \sum_{j=1}^{S} \frac{\partial \overline{n}_j}{\partial \tau_i} \times \frac{\partial w_i(\tau_{\bullet,i}, \tilde{\boldsymbol{\tau}}_i, \boldsymbol{n}, \hat{\boldsymbol{n}})}{\partial n_j}, \tag{B.20}$$

where

$$\overline{n}_j = \sum_{t=0}^{\infty} \sum_{\boldsymbol{s}_t \in \Omega_t} n_{j,t} q_i^\circ(\boldsymbol{s}_t), \tag{B.21}$$

is thus the abundance of species j, which is experienced by a mutant of species i that is randomly sampled from its local lineage under neutrality (i.e., from the distribution $q_i^\circ(\boldsymbol{s}_t)$).

Inter-generational effects of a mutant on its own ecology. We now derive eqs. (17)–(19) of the main text, which reveal the inter-generational nature of the effect of a mutant on its lineage-experienced abundance. We begin by writing the effect of a mutant on lineage-experienced abundance in vector form,

$$\frac{\partial \overline{\mathbf{n}}}{\partial \tau_i} = \sum_{t=0}^{\infty} \sum_{\boldsymbol{s}_t \in \Omega_t} \frac{\partial \boldsymbol{n}_t}{\partial \tau_i} q_i^\circ(\boldsymbol{s}_t), \tag{B.22}$$

where $\overline{\mathbf{n}} = (\overline{n}_1, \overline{n}_2, \ldots, \overline{n}_S)$. The derivative on the right hand side of eq. (B.22) can be evaluated by differentiating both sides of $n_{j,t}(\boldsymbol{\tau}_i) = F_j\big(\bar{\boldsymbol{\tau}}_{i,t-1}, \boldsymbol{n}_{t-1}(\boldsymbol{\tau}_i), \hat{\boldsymbol{n}}(\boldsymbol{\theta})\big)$ (i.e., eq. 11 of the main text) with respect to τ_i using the chain rule. The result of this operation can be written in vector form as

$$\frac{\partial \boldsymbol{n}_t}{\partial \tau_i} = \mathbf{C}(\boldsymbol{\theta}) \frac{\partial \boldsymbol{n}_{t-1}}{\partial \tau_i} + \boldsymbol{\Psi}_i(\boldsymbol{\theta}) \hat{n}_i \frac{M_{i,t-1}}{\hat{n}_i}, \tag{B.23}$$

where $\mathbf{C}(\boldsymbol{\theta})$ is the community matrix (eq. 6), and

$$\boldsymbol{\Psi}_i(\boldsymbol{\theta}) = \frac{1}{\hat{n}_i} \left(\frac{\partial F_1(\bar{\boldsymbol{\tau}}_i, \boldsymbol{n}, \hat{\boldsymbol{n}})}{\partial \bar{\tau}_i}, \frac{\partial F_2(\bar{\boldsymbol{\tau}}_i, \boldsymbol{n}, \hat{\boldsymbol{n}})}{\partial \bar{\tau}_i}, \ldots, \frac{\partial F_S(\bar{\boldsymbol{\tau}}_i, \boldsymbol{n}, \hat{\boldsymbol{n}})}{\partial \bar{\tau}_i} \right), \tag{B.24}$$

is a vector of S elements (all evaluated at $\boldsymbol{\theta}$) with $\bar{\boldsymbol{\tau}}_i = (\theta_1, \ldots, \theta_{i-1}, \bar{\tau}_i, \theta_{i+1}, \ldots, \theta_S)$ where $\bar{\tau}_i$ is the average phenotype in the focal patch (i.e., the k-th element of $\boldsymbol{\Psi}_i(\boldsymbol{\theta})$ is given by eq. 18 of the main text).

Eq. (B.23) is a linear recurrence equation in $\partial \boldsymbol{n}_t / \partial \tau_i$ with initial condition $\partial \boldsymbol{n}_0 / \partial \tau_i = \mathbf{0}$ (as the first mutant appears at generation 0). The solution to a linear vector recurrence of the form $\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{b}_{t-1}$, with initial condition $\mathbf{x}_0 = \mathbf{0}$, is $\mathbf{x}_t = \sum_{h=0}^{t-1} \mathbf{b}_h \mathbf{A}^{t-h-1}$ (e.g., Sydsaeter et al., 2008). Hence, the solution of eq. (B.23) is

$$\frac{\partial \boldsymbol{n}_t}{\partial \tau_i} = \sum_{h=0}^{t-1} \boldsymbol{\Psi}_i(\boldsymbol{\theta}) \hat{n}_i \frac{M_{i,h}}{\hat{n}_i} \mathbf{C}(\boldsymbol{\theta})^{t-h-1} = \boldsymbol{\Psi}_i(\boldsymbol{\theta}) \sum_{h=1}^{t} \mathbf{K}_h(\boldsymbol{\theta}) \hat{n}_i \frac{M_{i,t-h}}{\hat{n}_i}, \tag{B.25}$$

where recall $\mathbf{K}_h(\boldsymbol{\theta}) = \mathbf{C}(\boldsymbol{\theta})^{h-1}$ (see eq. 19). Substituting the right hand side of eq. (B.25) into eq. (B.22) gives

$$\begin{aligned} \frac{\partial \overline{\mathbf{n}}}{\partial \tau_i} &= \sum_{t=0}^{\infty} \sum_{\boldsymbol{s}_t \in \Omega_t} \boldsymbol{\Psi}_i(\boldsymbol{\theta}) \sum_{h=1}^{t} \mathbf{K}_h(\boldsymbol{\theta}) \hat{n}_i \frac{M_{i,t-h}}{\hat{n}_i} q_i^{\circ}(\boldsymbol{s}_t) \\ &= \boldsymbol{\Psi}_i(\boldsymbol{\theta}) \sum_{h=1}^{\infty} \mathbf{K}_h(\boldsymbol{\theta}) \hat{n}_i \left[\sum_{t=h}^{\infty} \sum_{\boldsymbol{s}_t \in \Omega_t} \frac{M_{i,t-h}}{\hat{n}_i} q_i^{\circ}(\boldsymbol{s}_t) \right], \end{aligned} \tag{B.26}$$

where the second line is obtained from the first by exchanging dummy variables.

Next, we define

$$r_{i,h}(\boldsymbol{\theta}) = \sum_{t=h}^{\infty} \sum_{\boldsymbol{s}_t \in \Omega_t} \frac{M_{i,t-h}}{\hat{n}_i} q_i^{\circ}(\boldsymbol{s}_t), \tag{B.27}$$

as the relatedness coefficient between two individuals of species i sampled h generations apart in a monomorphic resident population ($r_{i,h}(\boldsymbol{\theta})$ sums the probability of sampling two mutants at h generations apart over the whole sojourn time of the lineage in a single group, see eq. B.5). Such relatedness coefficients can be computed using standard identity-by-descent arguments (Malécot, 1973, Epperson, 1999, Lehmann, 2008), and for $h \geq 1$, it takes the form given in eq. (20) of the main text. If we substitute eq. (B.27) into eq. (B.26), we obtain

$$\frac{\partial \overline{\mathbf{n}}}{\partial \tau_i} = \boldsymbol{\Psi}_i(\boldsymbol{\theta}) \sum_{h=1}^{\infty} \mathbf{K}_h(\boldsymbol{\theta}) \left[\hat{n}_i r_{i,h}(\boldsymbol{\theta}) \right], \tag{B.28}$$

which is eq. (17) of the main text in vector form, as required.

Dispersal-limited versions of press perturbations. Finally, we derive eqs. (21)–(22) of the main text. Substituting eqs. (20) and (19) into eq. (B.28), we have

$$\begin{aligned}\frac{\partial \overline{\mathbf{n}}}{\partial \tau_i} &= \mathbf{\Psi}_i(\boldsymbol{\theta})(1-m_i(\boldsymbol{\theta}))\hat{n}_i\bar{r}_i(\boldsymbol{\theta})\sum_{h=1}^{\infty}(1-m_i(\boldsymbol{\theta}))^{h-1}\mathbf{C}(\boldsymbol{\theta})^{h-1} \\ &= \mathbf{\Psi}_i(\boldsymbol{\theta})(1-m_i(\boldsymbol{\theta}))\hat{n}_i\bar{r}_i(\boldsymbol{\theta})\Lambda_i(\boldsymbol{\theta}),\end{aligned} \tag{B.29}$$

where

$$\Lambda_i(\boldsymbol{\theta}) = [\mathbf{I}-(1-m_i(\boldsymbol{\theta}))\mathbf{C}(\boldsymbol{\theta})]^{-1} \tag{B.30}$$

comes from the standard matrix summation result: $\sum_{h=1}^{\infty}\mathbf{A}^{h-1} = (\mathbf{I}-\mathbf{A})^{-1}$ (Sydsaeter et al., 2008), and reveals the dispersal-limited versions of press perturbations. As required, eq. (B.29) is eq. (21) in vector form.

Appendix C: Coevolution of helping and dispersal

C.1 Life-cycle

Here, we work out the example on the coevolution of helping and dispersal that is described in section 5.1.5. First, let us give more details on the life-cycle considered in this model. We assume a semelparous life-cycle whereby adults first interact socially (i.e., help each other), reproduce (in the absence of density-dependent competition), and then die. Offspring disperse according to a dispersal probability that is inheritable. Those that disperse survive dispersal with probability s. After dispersal, each offspring survives density-dependent regulation with probability $1/(1+\gamma J)$ where J is the total number of juveniles in a patch.

We consider the evolution of two traits: the level of helping and the probability of dispersing at birth. To capture this, we write the phenotype of a focal individual as $\tau_\bullet = (z_\bullet, d_\bullet)$, where $z_\bullet$ is its level of helping and $d_\bullet$ the dispersal probability of its own and of its offspring. Likewise, we write the average phenotype among neighbours of the focal as $\tilde{\tau} = (\tilde{z}, \tilde{d})$, where $\tilde{z}$ is the average level of helping among the neighbours of the focal, and $\tilde{d}$ is the dispersal probability of their offspring.

Two remarks are worth making about this model before proceeding to its analysis. First, our life cycle assumptions integrate those of Rousset and Ronce (2004), which studied the evolution of dispersal only, and those of Lehmann et al. (2006), which studied the evolution of helping only. Both of these works used the exact version of the selection gradient (Rousset and Ronce, 2004, eqs. 26-27, Lehmann et al., 2006, eqs. A.18-19), which required

extensive numerics to compute singular strategies.

Second, competition in this model occurs after dispersal, rather than before dispersal (like in the model of helping presented in the main text, section 5.1.1). In that latter case, selection in fact drives dispersal to zero (result not shown). This is because when competition occurs before dispersal, dispersal does not alleviate kin competition, which is the only driver of dispersal evolution here (i.e., in the absence of environmental extinction and inbreeding cost, Hamilton and May, 1977, Clobert et al., 2001).

C.2 Necessary components

In order to evaluate the approximate selection gradient, we need the fitness of a focal individual in a mutant patch with n individuals, when other patches are monomorphic for the resident traits $\theta = (z, d)$ and at ecological/demographic equilibrium, $\hat{n}(\theta)$. According to our assumptions, this is

$$w(\tau_\bullet, \tilde{\tau}, n, \hat{n}(\theta)) = F(\bar{\tau}, n, \hat{n}(\theta)) \left(\frac{(1-d_\bullet) f(z_\bullet, \tilde{z})}{\lambda(\bar{\tau}, n, \hat{n}(\theta))} \right) + \hat{n}(\theta) \left(\frac{d_\bullet s f(z_\bullet, \tilde{z})}{(1-d+ds) f(z,z) \hat{n}(\theta)} \right), \tag{C.1}$$

where $f(z_\bullet, \tilde{z})$ is the fecundity of the focal individual, which depends on the focal's own level of helping $z_\bullet$ and the average level of helping $\tilde{z}$ among its neighbors in its group (given by eq. 35), and

$$\lambda(\bar{\tau}, n, \hat{n}(\theta)) = (1-\bar{d}) f(\bar{z}, \bar{z}) n + d s f(z,z) \hat{n}(\theta), \tag{C.2}$$

is the expected number of offspring in the focal patch after dispersal (but before competition). The first summand of eq. (C.1) is thus the philopatric component of fitness (i.e., the expected number of offspring that settle in the focal patch), and the second line, the dispersal component (i.e., the expected number of offspring that settle in other patches).

Our analysis of selection also requires specifying the population map, $F(\bar{\tau}, n, \hat{n}(\theta))$, which is the expected number of offspring reaching adulthood in the focal patch, conditional on the focal patch being of size n in the parental generation and other patches being at the resident demographic equilibrium, $\hat{n}(\theta)$. To compute $F(\bar{\tau}, n, \hat{n}(\theta))$, note first that from our assumption of Poisson fecundity (section 2.1), the (random) number of offspring J entering competition in the focal patch is Poisson distributed with mean $\lambda(\bar{\tau}, n, \hat{n}(\theta))$ (eq. C.2). Second, note that each of these J offspring independently survives competition with probability $1/(1+\gamma J)$. Hence, the expected number of offspring surviving competition is obtained as the expected value of the Binomially distributed number of surviving offspring (with parameters J and $1/(1+\gamma J)$), over the Poisson distribution of J with parameter

λ, i.e.,

$$F(\bar{\tau}, n, \hat{n}(\theta)) = \sum_{J=0}^{\infty} \frac{J}{1+\gamma J} \frac{\lambda(\bar{\tau}, n, \hat{n}(\theta))^J \exp\Big(-\lambda(\bar{\tau}, n, \hat{n}(\theta))\Big)}{J!}. \tag{C.3}$$

This is a complicated expression, but it is well approximated by a first-order delta approximation to the mean (Lynch and Walsh, 1998),

$$F(\bar{\tau}, n, \hat{n}(\theta)) \approx \frac{\lambda(\bar{\tau}, n, \hat{n}(\theta))}{1+\gamma\lambda(\bar{\tau}, n, \hat{n}(\theta))}. \tag{C.4}$$

From eq. (C.4), the equilibrium patch size in a resident population (obtained by solving $\hat{n} = F(\theta, \hat{n}, \hat{n})$ for $\hat{n}$) is

$$\hat{n}(\theta) = \frac{1}{\gamma}\left(1 - \frac{1}{(1-d+ds)\, f(z,z)}\right), \tag{C.5}$$

which depends on the level of helping z (through fecundity, $f(z,z)$) and decreases with dispersal d. The final piece of information that is necessary to compute the selection gradient is the coefficient of relatedness. This is given by eq. (10) of the main text, with resident equilibrium $\hat{n}$ given by eq. (C.5) and

$$m(\theta) = \frac{ds}{1-d+ds}, \tag{C.6}$$

as the backward probability of dispersal in the resident population.

C.3 Analysis of selection

From eqs. (14)–(22) of the main text, the selection gradient on trait $u \in \{z, d\}$ (respectively helping and dispersal) is given by

$$s_u(\theta) = \frac{\partial w(\tau_\bullet, \tilde{\tau}, n, \hat{n})}{\partial u_\bullet} + \frac{\partial w(\tau_\bullet, \tilde{\tau}, n, \hat{n})}{\partial \tilde{u}} r(\theta) + \frac{\partial \overline{n}}{\partial u} \times \frac{\partial w(\tau_\bullet, \tilde{\tau}, n, \hat{n})}{\partial n} \tag{C.7}$$

where

$$\frac{\partial \overline{n}}{\partial u} = \frac{\partial F(\bar{\tau}, n, \hat{n})}{\partial \overline{u}} (1-m(\theta))\bar{r}(\theta) \left[1 - (1-m(\theta)) \frac{\partial F(\bar{\tau}, n, \hat{n})}{\partial n}\right]^{-1}. \tag{C.8}$$

Singular helping and dispersal strategies are found by substituting eqs. (C.1)–(C.6) into eq. (C.7) and solving

$$s_z(\theta^*) = 0 \qquad \text{and} \qquad s_d(\theta^*) = 0, \tag{C.9}$$

numerically fo $\theta^* = (d^*, z^*)$ (subject to the constraint that patch size is positive, $\hat{n}(\theta) > 0$, see eq. C.5). The singular strategies for specific parameter values are shown in Fig. (6) (which also shows the results of individual-based simulations for this model).

References

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-Biological Sciences* 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. In Clobert, J., E. Danchin, A. A. Dhondt, and J. D. Nichols (eds.), *Dispersal*, chap. 23, pp. 329–340. Oxford University Press.

Bohonak, A. J. 1999. Dispersal, Gene Flow, and Population Structure. *The Quarterly Review of Biology* 74:21–45.

Bonduriansky, R. 2012. Rethinking heretidy, again. *Trends in Ecology and Evolution* 27:330–6.

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* 162:427–41.

Case, T. J. 2000. An Illustrated Guide to Theoretical Ecology. Oxford University Press, Oxford.

Caswell, H. 2000. Matrix Population Models. Sinauer Associates, Massachusetts.

Charlesworth, B. 1994. Evolution in Age-Structured Populations. Cambridge University Press, Cambridge, 2th edn.

Charnov, E. L. 1976. Optimal foraging, the marginal value theorem. *Theoretical Population Biology* 9:129–36.

Chesson, P. L. 1978. Predator-prey theory and variability. *Annual Review of Ecology and Systematics* 9:323–47.

Chesson, P. L. 1981. Models for spatially distributed populations: the effect of within-patch variability. *Theoretical Population Biology* 19:288–325.

Chesson, P. L. 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.

Cohen, J. E. 1979. Long-run growth rates of discrete multiplicative processes in Markovian environments. *Journal of Mathematical Analysis and Applications* 69:243–251.

Comins, H. N., W. D. Hamilton, and R. M. May. 1980. Evolutionarily stable dispersal strategies. *Journal ot 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* 74:209–25.

Débarre, F., S. L. Nuismer, and M. Doebeli. 2014. Multidimensional (co)evolutionary stability. *Am. Nat.* 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.

Débarre, F., L. S., M. van Baalen, and S. Gandon. 2012. Evolution of host life-history traits in a spatially structured host-parasite system. *American Naturalist* 179.

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.

Epperson, B. K. 1999. Gene genealogies in geographically structured populations. *Genetics* 152:797–806.

Eshel, I. 1972. On the neighbor effect and the evolution of altruistic traits. *Theoretical Population Biology* 11:258–277.

Eshel, I. 1983. Evolutionary and continuous stability. *Journal ot Theoretical Biology* 103:99–111.

Eshel, I. 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.

Frank, S. A. 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.

Gardner, A. and S. A. West. 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. In Krebs, J. R. and N. Davies (eds.), *Behavioural Ecology*, pp. 5–31. Blackwell Scientific Publications, Oxford.

Grafen, A. 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, 1. *Journal of Theoretical Biology* 7:1–16.

Hamilton, W. D. 1966. The moulding of senescence by natural selection. *Journal of Theoretical Biology* 12:12–45.

Hamilton, W. D. 1971. Selection of selfish and altruistic behaviour in some extreme models. In Eisenberg, J. and W. Dillon (eds.), *Man and Beast: Comparative Social Behavior*, pp. 59–91. Smithsonian Institutions Press, 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.

Hartl, D. and A. G. Clark. 2007. Principles of Population Genetics. Sinauer, Massachusetts, 4th edn.

Hubbell, S. P. 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press, Princeton.

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.

Jansen, V. A. A. and A. L. Lloyd. 2000. Local stability analysis of spatially homogeneous solutions of multi-patch systems. *J. Math. Biol.* 41:232–252.

Johnstone, R. A., M. Cant, and F. J. 2012. Sex-biased dispersal, haplodiploidy and the evolution of helping in social insects. *Proceedings of the Royal Society B: Biological SciencesBiological Sciences* 279:787–793.

Johnstone, R. A. and M. A. Cant. 2008. Sex differences in dispersal and the evolution of helping and harming. *American Naturalist* 172:318–330.

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* .

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.

Le Galliard, J., R. Ferrière, and U. Dieckmann. 2003. The adaptive dynamics of altruism in spatially heterogeneous populations. *Evolution* 57:1–17.

Le Galliard, J., R. Ferrière, and U. Dieckmann. 2005. Adaptive Evolution of Social Traits: Origin, Trajectories. *American Naturalist* 165:206–224.

Lehmann, L. 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.

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* In press.

Lion, S. and S. Gandon. 2009. Habitat saturation and the spatial evolutionary ecology of altruism. *Journal of Evolutionary Biology* 22:1487–1502.

Lion, S. and S. Gandon. 2010. Life history, habitat saturation and the evolution of fecundity and survival altruism. *Evolution* 64:1594–606.

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.

Lion, S. and M. van Baalen. 2007b. Self-structuring in spatial evolutionary ecology. *Ecology Letters* 11:277–295.

Lynch, M. and B. Walsh. 1998. Genetics and Analysis of Quantitative Traits. Sinauer, Massachusetts.

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.

Maynard Smith, J. 1982. Evolution and the Theory of Games. Cambridge University Press, Cambridge.

Metz, J. and M. Gyllenberg. 2001. How should we define fitness in structured metapopulation models? including an application to the calculation of evolutionary stable dispersal strategies. *Proceedings of the Royal Society of London Series B-Biological Sciences* 268:499–508.

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. In van Strien, S. J. and S. M. Verduyn Lunel (eds.), *Stochastic and Spatial Structures of Dynamical Systems*, pp. 183–231. North-Holland, Amsterdam.

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.

Metz, J. J. and C. G. F. de Kovel. 2013. The canonical equation of adaptive dynamics for Mendelian diploids and haplo-diploids. *Interface Focus* 3:20130025–20130025.

Meyn, S. and R. L. Tweedie. 2009. Markov Chains and Stochastic Stability. Cambridge University Press, Cambridge, 2th edn.

Mullon, C., L. Keller, and L. Lehmann. 2016. Evolutionary stability of jointly evolving traits in subdivided populations. *American Naturalist* 188:175–195.

Mullon, C., L. Keller, and L. Lehmann. 2017. Social polymorphism is favoured by the co-evolution of dispersal with social behaviour. *Nature Ecology & Evolution* 2:1–11.

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-Verlag, 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.

Odling-Smee, F. J., K. L. Laland, and M. W. Feldman. 2003. Niche Construction. Princeton University Press, Princeton, NJ.

Parker, G. A. 1990. Sperm competition games: raffles and roles. *Proceedings of the Royal Society B-Biological Sciences* 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* 16:143–53.

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.

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.

Rodrigues, A. M. M. and A. Gardner. 2012. Evolution of helping and harming in heterogeneous populations. *Evolultion* 66:2065–2079.

Rogers, A. R. 1990. Group selection by selective emigration: the effects of migration and kin structure. *American Naturalist* 135:398–413.

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. *Proc. R. Soc. B Biol. Sci.* 276:2955–2962.

Stearns, S. 1992. The Evolution of Life Histories. Oxford University Press, Oxford.

Sydsaeter, K., P. Hammond, A. Seierstad, and A. Strøm. 2008. Further Mathematics for Economic Analysis. Prentice Hall, Essex, 2th edn.

Taylor, P. 1988a. Inclusive fitness models with two sexes. *Theoretical Population Biology* 34:145–168.

Taylor, P. D. 1988b. An inclusive fitness model for dispersal of offspring. *Journal of Theoretical Biology* 130:363–378.

Taylor, P. D. 1989. Evolutionary stability in one-parameter models under weak selection. *Theoretical Population Biology* 36:125–143.

Taylor, P. D. 1992. Altruism in viscous populations - an inclusive fitness model. *Evolutionary Ecology* 6:352–356.

Taylor, P. D. 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. *Am. Nat* 191:368–380.

Tilman, D. 1982. Resource Competition and Community Structure. Princeton University Press, New Jersey.

Tuljapurkar, S. 1989. An uncertain life: demography in random environments. *Theoretical Population Biology* 35:227–94.

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.

Urban, M. C., M. A. Leibold, P. Amarasekare, L. De Meester, R. Gomulkiewicz, M. E. Hochberg, C. A. Klausmeier, N. Loeuille, C. de Mazancourt, J. Norberg, J. H. Pantel, S. Y. Strauss, M. Vellend, and M. J. Wade. 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.

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.

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* 73:1028–51.

Wild, G., A. Gardner, and S. A. West. 2009. Adaptation and the evolution of parasite virulence in a connected world. *Nature* 459:983–986.

Wolfram Research, I. 2016. Mathematica. Wolfram Research, Inc., Champaign, Illinois.

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* 26:1854–65.

Yodzis, P. 1989. Introduction to Theoretical Ecology. Harper & Row, New York.

Figures

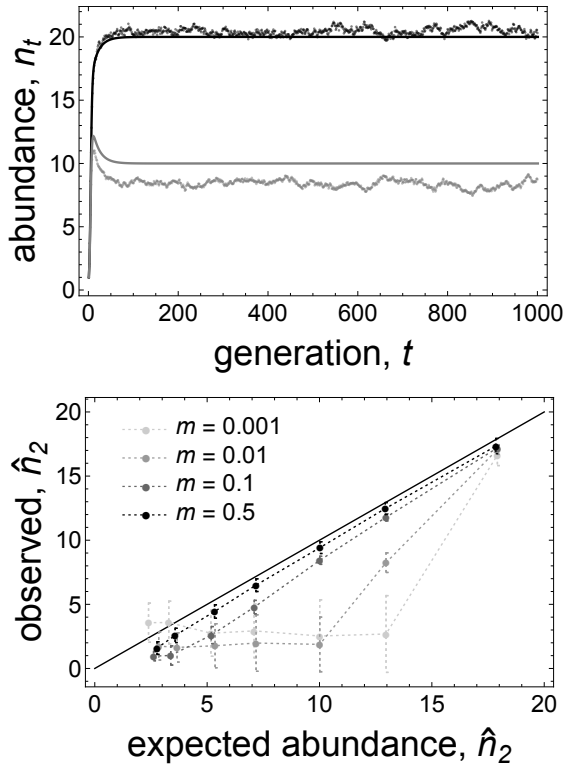

Figure 1: Stochastic dynamics of local abundance and their deterministic approximation. **Top panel**: deterministic (full line, from eq. 7, with $\hat{\boldsymbol{n}} = \boldsymbol{n}_{t-1}$) and stochastic (dots, with 1000 patches, see Appendix A.1 for details) dynamics for Leslie-Gower model with two species (species 1 in black and 2 in grey, with parameters $f_1 = 2.1$; $f_2 = 2$; $\gamma = 0.04$; $\eta = 0.03$; $m_1 = m_2 = 0.1$, starting with one individual of each species in each patch). **Middle panel**: Comparisons between the deterministic (x-axis, from eq. 8) and stochastic (y-axis) number of individuals of species 2, averaged over 1000 generations starting with one individual of each species in each patch (error bars give the standard deviation of the difference between deterministic and stochastic value, with parameters $f_1 = 2.1$; $f_2 = 2$; $(\gamma, \eta) = (0.05, 0.045)$, $(0.048, 0.043)$, $(0.045, 0.039)$, $(0.043, 0.035)$, $(0.04, 0.03)$, $(0.040, 0.025)$, $(0.05, 0.003)$; $m_1 = m_2 = 0.001, 0.01, 0.1, 0.5$). Departures from the diagonal indicate deviations between the exact process and the deterministic approximation.

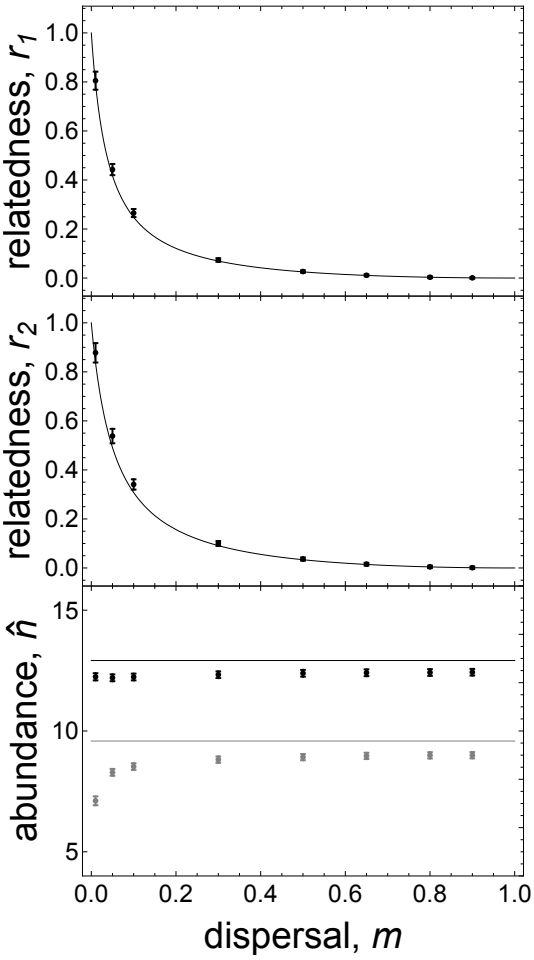

Figure 2: Relatedness and average local abundance under stochastic individual-based simulations and their deterministic approximation. The top two graphs show the relatedness for species 1 and 2, respectively, under the Leslie-Gower demographic model (section 3.2). Relatedness values obtained from the deterministic approximation are shown in full 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 5000 generations after burnin of 5000 generations, with $m = 0.01, 0.05, 0.1, 0.3, 0.5, 0.8, 0.9$, see Appendix A.1 for details on calculations of relatedness). 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 5000 generations after burnin of 5000 generations, with $m = 0.01, 0.05, 0.1, 0.3, 0.5, 0.8, 0.9$). Error bars show standard deviation.

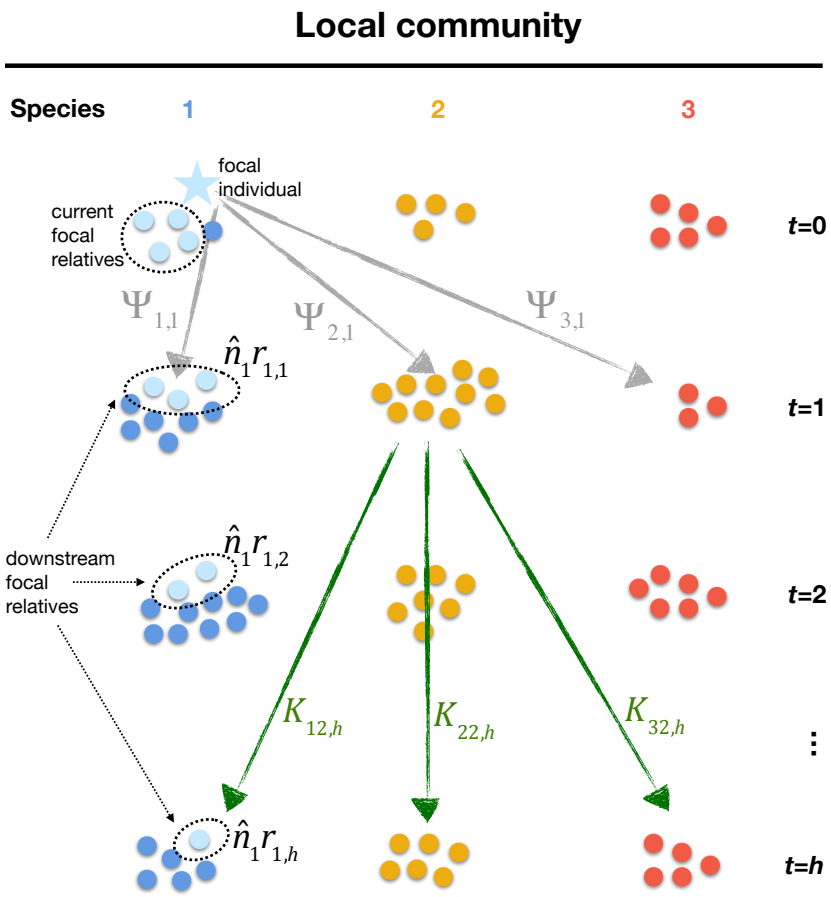

Figure 3: Inter-generational effects of a focal individual. As an example, consider a local community of three species, labelled 1 (blue), 2 (orange) and 3 (red). We look at the effect of a mutation in species 1. A focal carrier of a mutation τ_1 in species 1 living at generation $t = 0$ (denoted by a blue star) first directly influences the population dynamics of species 1, 2 and 3 at generation $t = 1$ according to $\Psi_{1,1}(\theta)$, $\Psi_{2,1}(\theta)$ and $\Psi_{3,1}(\theta)$ (grey arrows, eq. 18). This change in abundance is experienced by an expected number $\hat{n}_1(\boldsymbol{\theta})r_{1,1}(\boldsymbol{\theta})$ of relatives of the focal that live in generation $t = 1$ (eq. 20). This change in abundance at generation $t = 1$ affects the abundance of the other species through time due to ecological interactions. For e.g., a change in abundance of species 2 at generation $t = 1$ influences the abundance of species 1, 2 and 3 at generation $t = h$ according to $\mathrm{K}_{12,h}(\boldsymbol{\theta})$, $\mathrm{K}_{22,h}(\boldsymbol{\theta})$ and $\mathrm{K}_{32,h}(\boldsymbol{\theta})$, respectively (eq. 19). These changes are experienced by $\hat{n}_1(\boldsymbol{\theta})r_{1,h}(\boldsymbol{\theta})$ relatives of the focal individual.

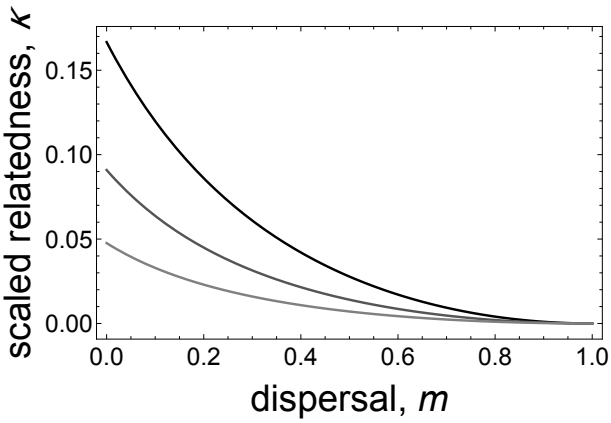

Figure 4: Effective relatedness for the evolution of social interactions within species. Eq. (34) with eq. (26) plotted against dispersal with $\gamma = 0.2$ (black), $\gamma = 0.1$ (dark grey) and $\gamma = 0.05$ (light grey). Other parameters: $f(\theta,\theta) = 2$. Effective relatedness therefore decreases with dispersal and patch size (since smaller values of γ lead to larger equilibrium patch size).

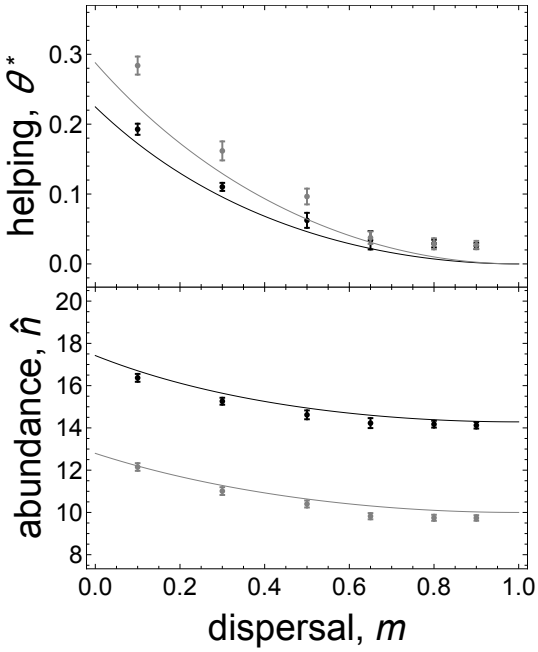

Figure 5: Convergence stable level of within-species helping and the concomitant local abundance it generates. Full lines are the convergence stable strategies (top) and concomitant local abundance (bottom) obtained from the selection gradient (obtained by finding the θ^* value solving eq. 33) under weak (black, $\gamma = 0.07$) and strong (grey, $\gamma = 0.1$) intra-specific competition (other parameters: $B = 0.05$, $C = 0.5$, $f = 2$). Points are the results obtained from individual based simulations (time average over 50000 generation after 50000 generations of evolution, error bars show standard deviation). Parameters for simulations: 1000 patches, $m = 0.1, 0.3, 0.5, 0.8, 0.9$, probability of a mutation = 0.01, standard deviation of the quantitative effect of a mutation = 0.005 (see Appendix A.1 for details on simulations).

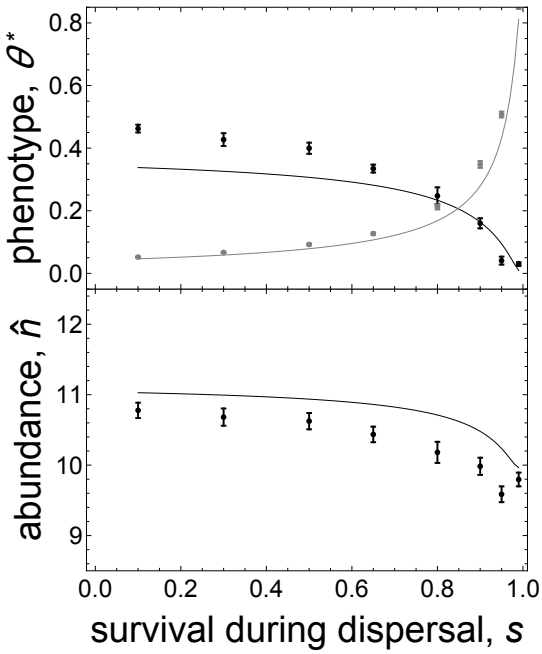

Figure 6: Co-evolutionary convergence stable level of helping and dispersal, and the concomitant local abundance it generates. Full lines are the convergence stable strategies (top) for helping (black) and dispersal (grey) and concomitant local abundance (bottom) obtained from the selection gradient (eq. C.7, with $B = 0.05$, $C = 0.5$, $f = 2$, $\gamma = 0.05$). Points are the results obtained from individual based simulations (time average over 50000 generation after 50000 generations of evolution, error bars show standard deviation). Parameters for simulations: 1000 patches, survival during dispersal $s = 0.1, 0.3, 0.5, 0.8, 0.9, 0.95, 0.99$, probability of a mutation = 0.01, standard deviation of the quantitative effect of a mutation on each trait = 0.005, and no covariance (see Appendix A.1 for details on simulations).

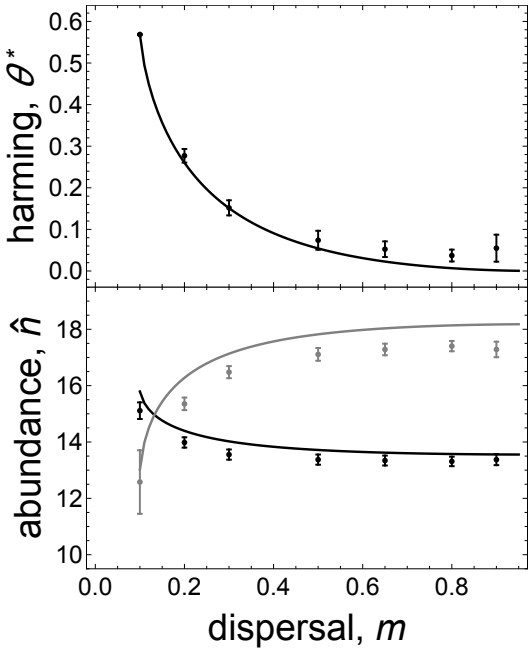

Figure 7: Convergence stable level of between-species harming and concomitant local abundance it generates. Full lines are the convergence stable strategies (top) and concomitant local abundance (bottom) in species 1 (black) and 2 (grey) obtained from the selection gradient (eqs. 14–21 along with eqs. 37–44, with $C_1 = 0.0085$, $f_1 = 2.2$, $f_1 = 2$, $\gamma = 0.055$, $\eta = 0.025$, $\alpha = 0.24$). Points are the results obtained from individual based simulations (time average over 50000 generation after 50000 generations of evolution, error bars show standard deviation). Parameters for simulations: 1000 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.