Adaptive Range Evolution of a Species with Phenotype-Optimal Dispersal

Farshad Shirani* Judith R. Miller†

Abstract

Phenotype-dependent optimal dispersal or matching habitat habitat choice strategies can play pivotal roles in range evolution and persistence of a species. Such strategies facilitate rapid adaptation, increase range expansion abilities, and create directed gene flow that promotes genetic differentiation, reproductive isolation, and speciation. However, despite the potential for such crucial consequences, solid empirical evidence that confirms the evolution of phenotype-optimal dispersal in nature is limited. To understand the reasons, it is important to identify the major eco-evolutionary impacts of optimal dispersal, the conditions under which it is sufficiently beneficial to evolve, and the factors that can be used to reliably detect it. To contribute to such understanding, we develop and computationally study a spatially-structured continuum model of a species' range evolution in a heterogeneous environment. The individuals of the species in our model follow the environmental gradient in a fitness-related phenotypic trait and settle in habitat locations that best match their phenotype. Our results identify presence of a steep environmental gradient—possibly steeper than the gradient the majority of species typically experience in nature—as a key factor in making phenotype-optimal dispersal sufficiently consequential to evolve. In a steep gradient, we show that optimal dispersal facilitates rapid adaption in single-generation time scales, reduces within-population trait variation, increases range expansion speed, and enhances the chance of survival in rapidly changing environments. Moreover, it creates a directed gene flow that reverses the maladpative core-to-edge effects of random gene flow caused by random movements. Therefore, we suggest adaptive gene flow to range margins, as well as substantially reduced trait variation at central populations, as two key factors to be considered for detecting phenotype-optimal dispersal in natural populations.

*School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA
†Department of Mathematics and Statistics, Georgetown University, Washington, DC 20057, USA

Introduction

A species' range evolves through complex eco-evolutionary processes that vary in time and space. These processes involve numerous factors, such as, species' niche limitations, intra- and inter-specific competitions, Allee effects, species's dispersal strategies and dispersal limitations, genetic structure of the species's population, gene flow, genetic drift, landscape heterogeneity and fragmentation, and environmental stress gradients (Angert et al., 2020; Bridle and Vines, 2007; Case et al., 2005; Duckworth and Badyaev, 2007; Fronhofer and Altermatt, 2015; Gaston et al., 2003; Godsoe et al., 2017; Haddad et al., 2015; Holt and Keitt, 2005; Louthan et al., 2015; Miller et al., 2020; Ponchon and Travis, 2022; Sexton et al., 2009; Shirani and Miller, 2022). Understanding the mechanisms through which the convoluted interactions between these factors affect a species' range dynamics is crucial in predicting how the species' population distribution will respond to climate change. It is also crucial in identifying conditions that result in biological invasions. Therefore, the knowledge gained by studying causes and consequences of species' range evolution is critical in developing strategies for controlling invasive species, and informing management and conservation efforts for preserving biodiversity (Rafajlović et al., 2022).

Dispersal: A key factor in range evolution

One of the key ecological factors that determine a species' range dynamics is dispersal. Obviously, if a species does not disperse to unoccupied habit locations, then it cannot expand its range. However, the contribution of dispersal to the speed and pattern of range expansion is far beyond the trivial generation of pure displacements. In fact, the species' establishment, evolution, and persistence are all tightly linked to its dispersal (Baguette et al., 2013; Bonte et al., 2012; Bowler and Benton, 2005; Ronce, 2007). Hence, it is crucial for the survival and range expansion of a species' population to develop adaptive dispersal strategies based on its habitat structure, its intera- and interspecific interactions, and climatic changes in its environment. This makes dispersal a complex multi-dimensional phenotype that evolves jointly with multiple morphological, physiological, and behavioral traits in response to spatiotemporal variations in the environment (Baguette et al., 2013; Baines et al., 2019; Bonte et al., 2012; Clobert et al., 2009; Cote et al., 2017; Lustenhouwer et al., 2023; Ronce, 2007; Saastamoinen et al., 2018). Such convoluted evolution of dispersal is driven by the balance between the cost incurred at each of the three interdependent stages of dispersal—departure (emigration), transfer, and settlement (immigration)—and the overall benefit acquired from the dispersal (Bonte et al., 2012; Bowler and Benton, 2005; Cote et al., 2017; Garant et al., 2007). In animals, the cost-benefit balance of the individuals' dispersal strategy is often informed by a set of external cues as well as internal phenotypic traits, making their evolution of informed dispersal both *condition-dependent* and *phenotype-dependent* (Bonte et al., 2012; Clobert et al., 2009; Cote et al., 2017; Ponchon and Travis, 2022).

Impacts of dispersal on fitness

Dispersal is a major component of an individual's fitness in a heterogeneous landscape (Baguette et al., 2013). In some species of animals, such as herbivorous insects, individuals have no prior information about the quality of the surrounding habitat and make non-directional movements based only on the information they have about their current (possibly degraded) local habitat (Armsworth and Roughgarden, 2008). In general, the fitness of individuals decreases if they randomly disperse from their natal habitat locations, where they are likely better adapted to, to other locations wherein their genotypes are not tested. Therefore, the *migration load* imposed by random dispersal usually decreases the mean fitness of a population (Bolnick and Otto, 2013; Bonte et al., 2012; Edelaar and Bolnick, 2012; Lenormand, 2002) However, there is empirical evidence in a wide range of animal species confirming that individuals often bias their dispersal towards preferred habitat locations (Armsworth and Roughgarden, 2008; Bolnick and Otto, 2013; Cote et al., 2017). Most generally, a population's mean fitness increases when informed nonrandom dispersal strategies evolve in its individuals—that is, when the directed dispersal of individuals to new habitat locations provides them with a fitness (benefit) increase that is sufficiently exceeding their dispersal cost (Bowler and Benton, 2005; Cote et al., 2017; Edelaar and Bolnick, 2012, 2019; Jacob et al., 2017; Nicolaus and Edelaar, 2018). To make such fitness-associated dispersal, some animals can efficiently orient their movements by actively using different sources of information, such as abiotic cues, landscape landmarks, and presence and behavior of conspecifics. For example, nocturnal snakes sense the temperature and physical structure of rocks to move to thermally preferable habitats. (Bowler and Benton, 2005; Clobert et al., 2009; Ponchon and Travis, 2022).

Matching habitat choice: An adaptive dispersal strategy

Ideally, when individuals can perceive all components of their absolute fitness and are able to move freely, they can direct their dispersal to climb local fitness gradients and achieve their maximum expected fitness. Using individual-based models, the impacts of such *fitness-dependent dispersal* on adaptation and population dynamics of a species have been studied in a number of theoretical works (Armsworth, 2009; Armsworth and Roughgarden, 2005, 2008; Ravigné et al., 2009; Ruxton and Rohani, 1999). However, acquiring information about all major components of fitness is rather impossible. Instead, a more realistic, yet fairly idealized, informed dispersal strategy has been conceptualized in which dispersing individuals choose habitat locations that best match their fitness-related phenotype(s). For example, it has been suggested that medium ground finches with deeper bills tend to settle in areas richer in large-seeded plants, as their bill can crack larger seeds and thereby increase their food intake (Edelaar and Bolnick, 2019). Hence, with this type of phenotype- and condition-dependent adaptive dispersal—which is often referred to as *matching habitat choice* (Edelaar et al., 2008)—the expected performance of the individuals is maximized. Most likely, due to the correlation between individuals' phenotype and fitness, this maximized performance also leads to a significant increase in the individuals' fitness. (Edelaar and Bolnick, 2019).

Although strong evidence for the evolution of matching habitat choice in nature is still limited,

a growing number of experimental and empirical studies have identified it in different species. For instance, using microcosms of a ciliate species which shows genetic variability in performance along a thermal gradient, Jacob et al. (2017) have experimentally shown that local adaptation to the upper margin of the species' thermal niche is favored by dispersal with matching habitat choice, whereas it is hindered under random dispersal. Similarly, in a semi-natural warming experiment using a model species of reptiles, Bestion et al. (2015) have shown that individuals disperse to warmer or cooler habitats based on their preferred living temperature. In another study, Karpestam et al. (2012) manipulated the dorsal color of a pygmy grasshopper species to black and white, and observed the individuals' distribution over a solar radiation mosaic. They have demonstrated that, on average, black-painted individuals tend to reside in habitats with less radiation, and white-painted females had more hatchlings than black-painted ones in increased radiation treatments. Color-dependent habitat choice has also been identified in dark and pale reddish-pheomelanic barn owls (Dreiss et al., 2012). Mark-recapture data on nomadic crossbill birds has also proposed matching habitat choice as a contributor to rapid diversification of this species' ecotypes Benkman (2017).

Eco-evolutionary impacts of matching habitat choice

With matching habitat choice, individuals sort themselves across the environment to minimize their phenotype-environment mismatch. This induces an evolutionary force that acts at the individual level and permits local adaptation in ecological (within-generation) timescales, even in th absence of natural selection. Such mode of rapid adaptation is essentially different form adaptation by natural selection—which occurs at the population level and progresses in evolutionary (between-generation) timescales. Importantly, the preferential sorting across the environment also leads to substantial inter-individual variability in the distance and direction of dispersal, which creates specifically directed nonrandom gene flow in the population (Bolnick and Otto, 2013; Edelaar and Bolnick, 2012, 2019).

Verbal and individual-based models have identified significant consequences of the non-random gene flow and local adaptation caused by matching habitat choice. If a sufficiently large number of individuals disperse preferentially, the resulting non-random gene flow effectively compensates for, or even reverses, the homogenizing (maladaptive) effects of gene flow caused by randomly dispersing individuals. As a result, the standing genetic variation within locally adapted populations is reduced, whereas the genetic divergence between the populations is promoted (Edelaar and Bolnick, 2019). The enhanced genetic differentiation at the meta-population scale can then indirectly cause assortative mating by decreasing reproductive interactions among local populations. The resulting reproductive isolation and phenotypic segregation can eventually lead to sympatric speciation and increased biodiversity (Berner and Thibert-Plante, 2015; Bolnick and Otto, 2013; Edelaar et al., 2008; Karpestam et al., 2012; Nicolaus and Edelaar, 2018; Ravigné et al., 2009).

Despite the supporting experimental and empirical evidence—although still relatively scare—and the theoretically-predicted evolutionary consequences, the impacts of matching habits choice on species' fitness and adaptive evolution have largely been overlooked by researchers. In particular,

it is still not well-understood how the phenotype-dependent spatial assortment of individuals and the non-random gene flow created by matching habitat choice will influence the adaptive range expansion of a species. Importantly, further investigation is needed to understand how the increased level of competition between phenotypically localized individuals compromises the adaptive effects of directed gene flow on populations' mean fitness and phenotypic variation. In spatially structured populations, such interactions can be essentially different in central versus marginal populations. Due to such complexities, intuitive predictions of the consequences of these interacting processes on range expansion dynamics can potentially be misleading. Theoretical models are hence helpful in providing more reliable predictions, and guiding empirical and experimental studies. In particular, the prediction of the models can be used to identify conditions under which adaptive habitat choice is sufficiently consequential to be observable in nature.

Present work: Range evolution with phenotype-optimal dispersal

We develop and analyze a continuum model of a species' range evolution in a heterogeneous environment. The individuals' in our model disperse based on a special form of matching habitat choice strategy. In the general form of matching habitat choice formalism, it is assumed that individuals are able to collect information about the entire available habitat and directly disperse to the location (patch) that provides them with the best phenotype-environment match compared with any other habitat locations. This idealized assumption is rather unrealistic, as acquiring such global information about the habitat, even if possible, is very costly. Instead, we assume that individuals only locally assess their immediate surrounding environment and direct their movement towards the neighboring locations that maximally match their phenotype. Once moved to a new location, the individuals make assessments of the new surrounding areas, and make further movements to better matching locations. By continuously repeating this assessment-and-movement processes, the individuals eventually settle in a location which best matches their phenotype compared with other locations in their perceivable surroundings. Equivalently, in this mode of adaptive dispersal individuals direct their movements by following the direction of the gradient in environmental optimum phenotype, until they reach an optimally matching location. Depending on the spatial profile of the optimum phenotype, the eventually reached location by each individual can be the globally optimal habitat location which provides the individual with best possible phenotype match, or it can only be a locally optimal location. We refer to this form of matching habitat choice by tracking the environmental gradient as *phenotype-optimal dispersal*, and use our model to predict its evolutionary impacts on a species' range expansion dynamics.

We include a combination of both random and phenotype-optimal dispersal in our model. The random (diffusive) dispersal component can be considered as a model of local exploratory movements that the individuals undertake to assess their surrounding area and develop a perception of the direction and magnitude of environmental gradient. Such exploratory movements prior to dispersal have indeed been observed in nature (Armsworth and Roughgarden, 2005; Selonen and Hanski, 2006). The random dispersal can also incorporate the effects of other uniformed

movements, such as short-range movements to escape kin competition or inbreeding (Clobert et al., 2009). Moreover, the inclusion of random dispersal further provides the possibility for individuals to leave a locally optimal location and eventually settle in a location that globally optimizes their phenotype-environment performance.

The phenotype-optimal dispersal in our model is made by individuals who perceive both a phenotype-environment mismatch and a non-zero environmental gradient. If the environmental gradient is zero, or the individuals are not sensitive to it, then there will be no directed dispersal as all surrounding areas appear to be equally preferable to the individuals. In addition to each individual's direct sensory assessment, the individuals can perceive the gradient in the environmental optimum phenotype through their collective movement behavior. Moreover, tracking the environmental gradient can also occur through taxis, that is, in response to environmental stimuli (guides) that are strongly correlated with the optimum gradient. For instance, changes in temperature, chemical concentration, topography, wind strength and direction, water flow, or light intensity can effectively canalize the movement of individuals in the direction of the environmental gradient (Baguette et al., 2013).

We develop our model in a quantitative genetic framework as a system of partial differential equations (PDEs), presenting joint evolution of a species' population density, mean value of a fitness-related quantitative phenotypic trait, and the variance of the trait. We consider a heterogeneous environment with spatially varying optimal value for the phenotypic trait. In addition to random and phenotype-optimal dispersal, we incorporate directed and stabilizing selection, frequency-dependent selection generated by phenotype-dependent competition, and mutational changes in phenotype frequencies. We do not directly include any dispersal cost. However, the presence of intraspecific competition indirectly imposes a cost on the optimal dispersal, as phenotypically assorted individuals will experience intensified competition due to their close phenotype values. Also, we do not model the evolution of optimal dispersal. We assume the propensity to disperse optimally is already evolved in the individuals. However, we allow for the optimal dispersal propensity to be an adjustable parameter, so that we can make observations at different levels of evolved optimal dispersal.

We use our model to investigate how phenotype-optimal dispersal affects local adaptation rate, core-to-edge asymmetric gene flow, trait variance, population density, and expansion wave speed. In addition to computationally provided predictions, the mean-field PDE-based framework of our model allows for gaining useful mechanistic insight by directly inspecting each of the terms associated with the optimal dispersal in the equations of the model. We further analyze the performance of optimally dispersing populations under abrupt environmental fluctuations, as well as in fragmented habitats. We specifically aim to identify conditions that result in sufficiently strong impacts of phenotype-optimal dispersal, to suggest possible explanation for relative infrequency of matching habitat choice in nature.

Model Description

We build our model on our previous work (Shirani and Miller, 2022), which was an extension of the seminal works of Kirkpatrick and Barton (1997) and Case and Taper (2000) and aimed to study the adaptive range dynamics of a community of interacting species over an environmental gradient. Since in the present work we investigate the range evolution of a solitary species, we use a reduced version of our previous multi-species model to a single species case. We incorporate into the model new components that additionally capture the effects of phenotype-optimal dispersal. We describe only those details of our previous work that are necessary for a clear description of the new model. We refer the reader to our previous work for further details.

We model the m-dimensional habitat of the species by an open rectangle $\Omega \subset \mathbb{R}^{\mathrm{m}}$. Although in general the habitat can be 3-dimensional, in this study we only consider 1- and 2-dimensional habitats. We model the adaptive range dynamics of the species at each location $x = (x_1, \ldots, x_{\mathrm{m}}) \in \Omega$ and time $t \in [0, T]$, over an evolution time horizon of $T > 0$, by deriving equations that govern the joint evolution of three population quantities: $n(x,t)$ denoting the population density of the species, $q(x,t)$ denoting the mean value of a fitness-related quantitative phenotypic trait within the population, and $v(x,t)$ denoting the intraspecific variance of the trait.

The derivation of the equations of the model relies on a basic equation that specifies, over a small interval of time, the variation in population density of individuals with a quantitative phenotypic trait value of p. To present this equation, we first denote by $\phi(x,t,p)$ the relative frequency of phenotype value $p \in \mathbb{R}$ among all individuals of the species' population. Moreover, we denote by $g(x,t,p)$ the intrinsic growth rate of the population of individuals with phenotype value p. This intrinsic growth rate, as given below in (5), includes a Lotka-Volterra model of intraspecific competition. We denote by $\alpha(p,p')$ the competition kernel that captures the per capita effect of individuals with phenotype p' on the frequency of individuals with phenotype p within the population. In addition, we denote by $\partial_t^{(\mathrm{M})}\phi(x,t,p)$ the rate of mutational changes in the frequency of phenotype p. The basic equation underlying the derivation of the model can then be presented as

$$
\begin{aligned}
& n(x,t+\tau)\phi(x,t+\tau,p) - n(x,t)\phi(x,t,p) \\
&\qquad = \tau \operatorname{div}\Big(\mathrm{D}(x)\nabla_x\big(n(x,t)\phi(x,t,p)\big)\Big) && \text{(1a)} \\
&\qquad\quad - \tau \operatorname{div}\Big(\mathrm{A}(x)n(x,t)\phi(x,t,p)\big(-\nabla_x\theta(p,x)\big)\Big) && \text{(1b)} \\
&\qquad\quad + \tau g(x,t,p)n(x,t)\phi(x,t,p) && \text{(1c)} \\
&\qquad\quad + \tau n(x,t)\partial_t^{(\mathrm{M})}\phi(x,t,p), && \text{(1d)}
\end{aligned}
$$

where ∇_x denotes the gradient with respect to x, div denotes the divergence with respect to x, and ∂_t denotes the partial derivative with respect to t. In writing (1), we assume that the amount of change in the population density of individuals with phenotype p over a small time interval of $\tau \to 0$ results from the contributions of four factors: random dispersal of individuals to and from neighboring locations, modeled by (1a); directed (optimal) dispersal of individuals in the direction

that gives them maximum environmental match, modeled by (1b); the intrinsic growth of the population, modeled by (1c); and the mutational changes in the relative frequency of p, modeled by (1d). Parameter D denotes the diffusion coefficient of the species' random dispersal. Parameter A and the term $-\nabla_x \theta$ in (1b) can be interpreted, respectively, as individuals' propensity and individuals' perceived force to disperse optimally. Further descriptions on these terms are provided in the Optimal Dispersal section below.

Below, we first provide the list of key assumptions that we have made in writing (1) and the other formulations in the rest of the paper. We then give the formulation of the species' intrinsic growth rate $g(x,t,p)$ in (1c), and describe in detail how we model the optimal dispersal in (1b). The diffusion term (1a) that models species' random dispersal is standard, and the formulation for the effect of mutational changes in (1d), based on Assumption (viii) below, is given in Section A.3 in our previous work (Shirani and Miller, 2022).

Model Assumptions

To derive the equations of our model, we make the following major assumptions on the populations' dispersal and reproduction, as well as the elements of the intrinsic growth rate and the intraspecific competition kernel.

(i) Random dispersal of the individuals in the habitat is made by diffusion.

(ii) An individual's environmental potential for making optimal dispersal is proportional to the square of the difference between its phenotype value and the environment's optimum phenotype.

(iii) Nonlinear environmental selection for an optimal phenotype Q is stabilizing for all $x \in \Omega$. The optimal phenotype can vary over space and time.

(iv) The frequency of phenotype values within the species is normally distributed at each occupied point in space at all times. That is,

$$\phi(x,t,p) := \frac{1}{\sqrt{2\pi v(x,t)}} \exp\left(-\frac{(p-q(x,t))^2}{2v(x,t)}\right), \quad x \in \Omega, \quad t \in [0,T]. \tag{2}$$

(v) The reproduction rate of the individuals with phenotype p depends (predominantly) on the population density of the individuals with the same phenotype p.

(vi) The environmental resources vary continuously along a resource axis. After identifying the resource axis by phenotype axis, as described in Section A.2 of our previous work (Shirani and Miller, 2022), the *phenotype utilization distribution* by an individual with phenotype p is assumed to be normal, given by

$$\psi_p(\tilde{p}) := \frac{1}{\sqrt{2\pi \mathrm{V}}} \exp\left(-\frac{(\tilde{p}-p)^2}{2\mathrm{V}}\right), \quad i = 1, \ldots, \mathrm{N}, \tag{3}$$

where V denotes the variance of phenotype utilization be any individuals of the species.

(vii) The strength of intraspecific competition between the individuals is determined by the overlap between their phenotype utilization curves, which give the competition kernel as

$$\alpha(p, p') = \exp\left(-\frac{(p-p')^2}{4\mathrm{V}}\right). \tag{4}$$

The details of the derivation of this kernel function are available in Sections A.2 and A.4 of our previous work (Shirani and Miller, 2022). Here, we only consider symmetric competition between the individuals of different phenotypes, that is, $\kappa = 0$ in the formulation given in our previous work.

(viii) The probability of mutational changes from one phenotype p to another phenotype p' depends on the difference between the phenotypes $\delta p = p - p'$. Letting $\nu(\delta p)$ denote the probability density of such mutational changes, we further assume that ν follows a distribution with constant zero mean and constant variance $\mathrm{V_M}$. See Section A.3 of our previous work (Shirani and Miller, 2022) for the formulation of our model of mutational changes.

Intrinsic Growth Rate

In the absence of dispersal and genetic mutations, the local population dynamics of the species is determined by the intrinsic growth rate of the individuals, which we model as (Shirani and Miller, 2022: Equ. (17)),

$$g(x,t,p) := \mathrm{R}(x)\left(1 - \frac{1}{\mathrm{K}(x)} n(x,t) \int_{\mathbb{R}} \alpha(p,p')\phi(x,t,p')\mathrm{d}p'\right) \tag{5a}$$

$$- \frac{\mathrm{S}}{2}(p - \mathrm{Q}(x))^2, \tag{5b}$$

where R denotes the maximum growth rate of the species, K denotes the carrying capacity of the environment, S denotes the strength of stabilizing selection, and Q denotes the environment's optimal trait value. The phenotype distribution ϕ and the competition kernel α are given by (2) and (4), respectively. The convolution term in the Lotka-Volterra model (5a) captures the effect of intraspecific phenotypic competition on the frequency of phenotype p. The quadratic term (5b) incorporates the effects of directional and stabilizing selection on individuals with phenotype p, by penalizing the phenotypes that are far from the optimal phenotype $\mathrm{Q}(x)$.

Optimal Dispersal

We model the species' phenotype- and environment-dependent optimal dispersal by the advection term (1b). The parameter A is analogous to the *mobility* parameter that is often used in drift-diffusion models of flowing particles in a liquid. In our model, we can reasonably interpret A as a simplified model of individuals' *propensity* to disperse optimally. The evolution of a species' dispersal propensity depends on many factors, such as costs and benefits of different stages of dispersal, environmental conditions, and frequency-dependent eco-evolutionary processes (Bonte et al., 2012; Clobert et al., 2009). Moreover, the dispersal propensity of the individuals of a species

is often plastic, and can change in an ecological time scale. We do not include the evolution of the species's optimal dispersal propensity in our model though, as the required information is lacking. Instead, we perform our studies with different values of A to observe if different degrees of dispersal propensity will be sufficiently beneficial to evolve in a species. Although the equations of our model allow for A to be dependent on space, in the results we present in this work we assume A to be constant in space and time.

The term $-\nabla_x \theta(p, x)$ in (1b) is analogous to the force applied to the particles of drift-diffusion models due to the presence of an external potential energy $\theta(p, x)$. In our model, $\theta(p, x)$ is interpreted as individuals' perceived *dispersal potential*, which depends both on phenotype p of each individual and the (perceived) environmental trait optimum Q. As a result, the advective optimal dispersal (1b) in our model represents an informed dispersal strategy that is both phenotype-dependent (dependence on p) and condition-dependent (dependence on $\mathrm{Q}(x)$), as defined by Clobert et al. (2009). In this informed dispersal context, the phenotype value of an individual is an internal state of the individual, developed by the individual's self perception. The environmental trait optimum and its gradient, as described below, are external factors whose information can initiate and direct the species' optimal dispersal through the evolution of sensory and cognitive processing mechanisms in the individuals.

We consider the following simplified yet meaningful model for the informed dispersal potential function of an individual with phenotype p at habitat location x,

$$\theta(p, x) := \frac{(p - \mathrm{Q}(x))^2}{2\mathrm{V}} - 1. \tag{6}$$

This potential function is a first-order approximation of the function $-\sqrt{2\pi \mathrm{V}}\psi_p(\mathrm{Q}(x))$, where ψ_p is the phenotype utilization distribution (3). That means, using (6), the dispersal potential of an individual with phenotype p is determined by how poorly the individual can utilize the environment's optimal phenotype[1] at its current location. If the individual's phenotype matches the optimum phenotype perfectly, then there is no phenotypically and environmentally induced force on the individual to disperse. If the individual's phenotype differs significantly from the environment's optimum—measured relative to the species' phenotype utilization variance V—then the individual perceives a high potential to disperse to habitat locations of better quality that match its phenotype. Yet, a high dispersal potential $\theta(p, x)$ does not necessarily generate a significant dispersal force $-\nabla_x \theta(p, x)$ on the individual to disperse, unless a sufficiently large gradient in the dispersal potential is perceived by the species' movement. As the following discussion shows, such potential gradient is present if the environmental gradient $\nabla_x \mathrm{Q}$ is sufficiently steep and the individual is sufficiently sensitive to it. In this case, the species experiences a considerably better-matching habitat after making the directed dispersal.

The dispersal potential (6) gives the *dispersal force* term $-\nabla_x \theta$ used in (1b) as

$$-\nabla_x \theta(p, x) = \frac{p - \mathrm{Q}(x)}{\mathrm{V}} \nabla_x \mathrm{Q}(x). \tag{7}$$

[1]Note that, as stated in Assumption (vi), in writing the phenotype utilization distribution (3) we assume an identification between the resource axis and the phenotype axis.

If the magnitude of the environmental gradient $\|\nabla_x \mathrm{Q}(x)\|_{\mathbb{R}^\mathrm{m}}$ is zero or is perceived as equal to zero due to insensitivity of the individual to the gradient, then the perceived dispersal force on the individual to disperse non-randomly is zero—regardless of the presence of a phenotype-environment mismatch $p \neq \mathrm{Q}(x)$. In this case, the individual only disperses randomly due to the diffusion term (1a). If $\|\nabla_x \mathrm{Q}(x)\|_{\mathbb{R}^\mathrm{m}} > 0$, where $\|.\|_{\mathbb{R}^\mathrm{m}}$ denotes the Euclidean norm in $\mathbb{R}^\mathrm{m}$, then an individual whose phenotype does not perfectly match the optimum phenotype will perceive a positive force to disperse optimally. If $p > \mathrm{Q}(x)$, the optimal dispersal will be in the direction of the environmental gradient so that the individual disperses to habitat locations with larger optimum phenotype values, which give the individual a better phenotype-environment match. Similarly, if $p < \mathrm{Q}(x)$, the optimal match will be achieved by the individual through dispersal in the opposite direction of the environmental gradient.

The magnitude of a dispersal force perceived by an individual is unlikely to remain directly proportional to the magnitude of the environmental gradient $\|\nabla_x \mathrm{Q}(x)\|_{\mathbb{R}^\mathrm{m}}$ when the gradient becomes increasingly steep. When there exists a phenotype-environment mismatch $p-\mathrm{Q}$, a sufficiently steep environmental gradient should be enough to generate a maximal force for the individual to disperse. Developing proportional sensitivity to steeper gradients would then be unnecessarily costly for the dispersal purpose of the individual as it does not provide further information. To approximately incorporate such information saturation in individual's perception of the environmental gradient, we replace $\nabla_x \mathrm{Q}$ in (7) by the following *perceived gradient* whose magnitude will saturate to a maximum value Π when $\|\nabla_x \mathrm{Q}(x)\|_{\mathbb{R}^\mathrm{m}}$ becomes exceedingly large:

$$\widetilde{\nabla_x \mathrm{Q}(x)} := \frac{\Pi}{\Pi + \|\nabla_x \mathrm{Q}(x)\|_{\mathbb{R}^\mathrm{m}}} \nabla_x \mathrm{Q}(x), \qquad x \in \Omega_\delta. \tag{8}$$

When $\|\nabla_x \mathrm{Q}(x)\|_{\mathbb{R}^\mathrm{m}}$ is sufficiently smaller than Π, the perceived gradient is approximately the same as the actual gradient. When $\|\nabla_x \mathrm{Q}(x)\|_{\mathbb{R}^\mathrm{m}}$ is significantly larger than Π, the magnitude of the perceived gradient approximately saturates to the maximum value Π, but its direction will always be the same as the direction of the actual gradient.

The smaller habitat Ω_δ considered in (8) includes all points of Ω except those that are closer than a constant δ to the boundary of Ω. This is a rather technical consideration we make to remove the indefiniteness of the gradient at boundary points and to avoid complicated boundary conditions. In Appendix A, we provide the details on how to smoothly extend $\widetilde{\nabla_x \mathrm{Q}(x)}$ in (8) to the whole habitat Ω, based on the assumption that the individuals can sense the habitat boundary and avoid crossing it. Specifically, for the one- and two-dimensional habitats that we simulate in this work, we use the definitions (17) and (18) given in the Appendix A for $\widetilde{\nabla_x \mathrm{Q}(x)}$ over the entire habitat. In all of the simulations, we assume the individuals can perceive the habitat boundary when they get as close as $\delta = 2\,\mathtt{X}$ to the boundary. Moreover, we set the maximum perceived gradient to be $\Pi = 1\,\mathtt{Q}/\mathtt{X}$, which still corresponds to a relatively steep environmental gradient, as we discussed in Section 3.2 of our previous work (Shirani and Miller, 2022). In our simulations, we use the dispersal propensity parameter A as an adjustment parameter for the species' total rate of optimal dispersal.

Replacing the actual environmental gradient in (7) with its perceived value, we write the individual's perceived force to disperse optimally as

$$-\nabla_x \theta(p,x) = \frac{p - \mathrm{Q}(x)}{\mathrm{V}}\, \widetilde{\nabla_x \mathrm{Q}(x)}. \tag{9}$$

Although not possessing a true interpretation as a potential—unlike the dispersal potential $\theta(p,x)$ given by (6)—for ease of reference we refer to the term $(p-\mathrm{Q})/\mathrm{V}$ in (9) as *phenotypic potential* for optimal dispersal. We also note that the perceived gradient $\widetilde{\nabla_x \mathrm{Q}}$ can be interpreted, in some sense, as the sensitivity of the individuals' dispersal force to changes in the habitat quality. In nature, the individuals of a species may develop a perception of the environmental gradient not only by directly sensing their environment's conditions, but also by collecting information through exploration and collective behavior of their conspecifics. The diffusive movement term (1a) that we have included in our model can additionally capture such exploratory movements of the individuals for the purpose of gaining a better perception of their environment's gradient.

Model Equations

The basic equation (1) and its components as described above provide the ingredients that we need for deriving the equations of our model for joint evolution of a species's population density $n(x,t)$ and the mean value $q(x,t)$ and variance $v(x,t)$ of a quantitative fitness-related trait within the species's population. Assumption (iv) is a key assumption in developing our model as it allows for an exact moment closure in deriving the equations for trait mean and trait variance. In this section, we only present the ultimate equations of our model. The derivation of the equations is provided in detail in the Appendix B.

For all $x \in \Omega$ and $t \in [0,T]$, we derive the equation for the evolution of the species' population density $n(x,t)$ as

$$\partial_t n(x,t) = \operatorname{div}\left(\mathrm{D}(x)\nabla_x n(x,t)\right) \tag{10a}$$

$$- \operatorname{div}\left(\mathrm{A}(x) n(x,t) \frac{q(x,t) - \mathrm{Q}(x)}{\mathrm{V}} \widetilde{\nabla_x \mathrm{Q}(x)}\right) \tag{10b}$$

$$+ \left(\mathrm{R}(x) - \frac{\mathrm{R}(x)}{\mathrm{K}(x)} \sqrt{\frac{\mathrm{V}}{v(x,t)+\mathrm{V}}}\, n(x,t) - \frac{\mathrm{S}}{2}\left[\left(q(x,t)-\mathrm{Q}(x)\right)^2 + v(x,t)\right]\right) n(x,t). \tag{10c}$$

Moreover, letting $\langle \cdot\,, \cdot \rangle_{\mathbb{R}^{\mathrm{m}}}$ denote the standard inner product in $\mathbb{R}^{\mathrm{m}}$, we derive the equations for

the population's trait mean $q(x,t)$ and trait variance $v(x,t)$ as

$$\partial_t q(x,t) = \operatorname{div}\left(\mathrm{D}(x)\nabla_x q(x,t)\right) \tag{11a}$$
$$+ 2\left\langle \nabla_x \log n(x,t)\,,\, \mathrm{D}_i(x)\nabla_x q(x,t)\right\rangle_{\mathbb{R}^{\mathrm{m}}} \tag{11b}$$
$$- \operatorname{div}\left(\mathrm{A}(x)\frac{v(x,t)}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}(x)}\right) \tag{11c}$$
$$- \left\langle \nabla_x \log n(x,t)\,,\, \mathrm{A}(x)\frac{v(x,t)}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}(x)}\right\rangle_{\mathbb{R}^{\mathrm{m}}} \tag{11d}$$
$$- \left\langle \nabla_x q(x,t)\,,\, \mathrm{A}(x)\frac{q(x,t)-\mathrm{Q}(x)}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}(x)}\right\rangle_{\mathbb{R}^{\mathrm{m}}} \tag{11e}$$
$$- \mathrm{S}\left(q(x,t)-\mathrm{Q}(x)\right)v(x,t), \tag{11f}$$

and

$$\partial_t v(x,t) = \operatorname{div}(\mathrm{D}(x)\nabla_x v(x,t)) \tag{12a}$$
$$+ 2\left\langle \nabla_x \log n(x,t)\,,\, \mathrm{D}(x)\nabla_x v(x,t)\right\rangle_{\mathbb{R}^{\mathrm{m}}} \tag{12b}$$
$$+ 2\left\langle \nabla_x q(x,t)\,,\, \mathrm{D}(x)\nabla_x q(x,t)\right\rangle_{\mathbb{R}^{\mathrm{m}}} \tag{12c}$$
$$- 2\left\langle \nabla_x q(x,t)\,,\, \frac{\mathrm{A}(x)}{\mathrm{V}}v(x,t)\widetilde{\nabla_x \mathrm{Q}(x)}\right\rangle_{\mathbb{R}^{\mathrm{m}}} \tag{12d}$$
$$- \left\langle \nabla_x v(x,t)\,,\, \frac{\mathrm{A}(x)}{\mathrm{V}}\left(q(x,t)-\mathrm{Q}(x)\right)\widetilde{\nabla_x \mathrm{Q}(x)}\right\rangle_{\mathbb{R}^{\mathrm{m}}} \tag{12e}$$
$$+ \frac{\mathrm{R}(x)}{\mathrm{K}(x)}\sqrt{\frac{\mathrm{V}}{v(x,t)+\mathrm{V}}}\frac{n(x,t)v^2(x,t)}{2\left(v(x,t)+\mathrm{V}\right)} - \mathrm{S}v^2(x,t) + \mathrm{U}. \tag{12f}$$

Definitions of the model parameters and their plausible ranges of values are given in Table 1. Note that $\mathrm{D}(x) \in \mathbb{R}^{\mathrm{m}\times\mathrm{m}}$, whereas the rest of the parameters are scalar-valued. Moreover, S, U, and V are assumed to be constant all over the habitat, whereas D, A, K, R, and Q can be variable in space. Although their dependence on t is not explicitly shown in the equations, all these model parameters can also vary in time. A discussion on the choice of parameter units and their plausible values is provided below.

For our simulations of a one-dimensional habitat $\Omega = (a,b)$, we assume no phenotypic flux through the habitat boundary. Since the perceived environmental gradient $\widetilde{\nabla_x \mathrm{Q}}$ given by (17) in Appendix A has no component normal to the boundary, the advection term (1b) does not result in any phenotype flux through the boundary. Therefore, our no-flux boundary assumption simply implies the following homogeneous Neumann (reflecting) boundary conditions

$$\partial_x n = 0, \quad \partial_x q = 0, \quad \partial_x v = 0, \qquad \text{on } \{a,b\}\times[0,T], \tag{13}$$

which is the same boundary condition we discussed in Remark 1 of our previous work (Shirani and Miller, 2022). For the two-dimensional habitat $\Omega = (a_1,b_1)\times(a_2,b_2)$ that we simulate in this work, we set reflecting boundary condition at boundary lines $x_1 = a_1$ and $x_1 = b_1$, and we assume that the habitat is extended periodically in the x_2-direction.

Table 1: Definition and plausible range of values of the parameters of the model (10)–(12)**.** Except for $\mathrm{A}(x)$, the range of parameter values and their choice of units are the same as those given by Shirani and Miller (2022). Unless otherwise stated, the typical values given here are the values used in the numerical studies of the Results section.

Parameter	Definition	Range	Typical	Unit
m	Spatial dimension of the geographic space	$\{1, 2, 3\}$	1, 2	—
$\mathrm{D}(x)$	Diffusion coefficient of the species' random dispersal	$[0, 25]^*$	1^*	$\mathtt{X}^2/\mathtt{T}$
$\mathrm{A}(x)$	Measure of the species' propensity to disperse optimally	$[0, 10]$	4	$\mathtt{X}^2/\mathtt{T}$
$\mathrm{K}(x)$	Carrying capacity of the environment	$(0, 10]$	1	$\mathtt{N}/\mathtt{X}^{\mathrm{m}}$
$\mathrm{R}(x)$	Maximum population growth rate of the species	$[0.1, 10]$	2	$1/\mathtt{T}$
V	Variance of the specie's phenotype utilization curve	$[0.25, 25]$	4	$\mathtt{Q}^2$
S	Measure of the strength of stabilizing selection	$[0, 2]$	0.2	$\mathtt{Q}^{-2}/\mathtt{T}$
U	Rate of increase in trait variance due to mutation	$[0, 0.2]$	0(0.02)	$\mathtt{Q}^2/\mathtt{T}$
$\mathrm{Q}(x)$	Optimal trait value for the environment	$[0, \infty)$	Linear†	$\mathtt{Q}$
$\|\nabla_x \mathrm{Q}(x)\|_{\mathbb{R}^{\mathrm{m}}}$	Magnitude of the gradient of the optimal trait	$[0, 10]$	0.2	$\mathtt{Q}/\mathtt{X}$

*When $\mathrm{m} > 1$, the range of values specified for $\mathrm{D}(x)$ can be considered for each entry of $\mathrm{D}(x) \subset \mathbb{R}^{\mathrm{m}\times\mathrm{m}}$. Typically, D is assumed to be diagonal.
†The typical value "Linear" specified for Q means that Q is typically considered to be a linear function of x over Ω.

In comparison with the equations of the single-species model that we studied in Section 4 of our previous work, the inclusion of the optimal dispersal strategy in the present work results in the additional term (10b) in the equation of population density, the terms (11c)–(11e) in the equation of trait mean, and the terms (12d) and (12e) in the equation of the trait variance. As a result, all quantitative measures of the population's dynamics, such as the population's density and speed of range expansion, the local adaptation rate of the population both at its core and its range margin, the asymmetric core-to-edge gene flow, and the dynamics of the intraspecific trait variance are expected to be affected by individuals' ability to disperse optimally to matching habitats. We demonstrate such impacts under different evolutionary regimes in our computational studies presented in the Results section.

Units and Parameter Values

We use the same units as we discussed in our previous work for the quantities included in our model. Specifically, we denote the unit of time by $\mathtt{T}$ and we set $1\,\mathtt{T}$ to be equal to the mean generation time of the species. For a one-dimensional habitat, we choose the unit of space so that the diffusion coefficient D of the population becomes unity. That is, denoting the unit of space by

$\mathtt{X}$, we set $1\,\mathtt{X}$ to be the root mean square of the (random) dispersal distance of the population in $1\,\mathtt{T}$, divided by $\sqrt{2}$. Estimates of the random component of the species' dispersal can be obtained, for example, by measuring dispersal distance of a sub-population of individuals that are well-adapted to the environment at the core of the population. Due to their negligible phenotype-environment mismatch, such individuals do not perceive a significant force to make directed dispersal. For multi-dimensional habitats, the same approach can be used to set the unit of space for each spatial dimension independently. Moreover, we denote the unit of measurement for population abundances by $\mathtt{N}$. Having set the unit of space, we set $1\,\mathtt{N}$ to be equal to the carrying capacity of the environment for $1\,\mathtt{X}^{\mathrm{m}}$ unit of habitat volume. This results in the carrying capacity to become unity. Finally, we denote the unit of measurement for the quantitative trait by $\mathtt{Q}$, and we set $1\,\mathtt{Q}$ to be equal to one standard deviation of the trait values at the core of the population. Further discussions on our choices of units are available in Section 3.1 of our previous work (Shirani and Miller, 2022).

Our suggestions of plausible ranges of values given in Table 1 are discussed in detail in Section 3.2 of our previous work, except for the new parameter A. The values specified as 'typical' in Table 1 are the values we consider as default parameter values in our simulations when otherwise is not stated. Due to the level of abstraction that is inevitably present in our model of optimal dispersal, finding a biologically reasonable range of values for the dispersal propensity parameter A based on empirical measurements available in the literature does not sound practical. Instead, we could suggest a range form 0 to $10\,\mathtt{X}^2/\mathtt{T}$ by simulating the model with different values of A and observing a reasonable range of variation in the population density, speed of range expansion waves, and magnitude of the trait variance. In an extreme situation, for example, when the mean value of the trait in (10b) differs from Q by one phenotype utilization variance and the magnitude of the environmental gradient is sufficiently larger than $\Pi = 1\,\mathtt{Q}/\mathtt{X}$, then a maximum dispersal (advection) rate of approximately $10\,\mathtt{X}/\mathtt{T}$ is created in a one-dimensional habitat at propensity value $\mathrm{A} = 10\,\mathtt{X}^2/\mathtt{T}$.

Interpretation of the Model Equations

Before presenting our results obtained by numerically solving the equations of our model, we discuss the interpretation of each of the terms involved in the equations. Inspection of the equations provides particularity useful mechanistic insight into the less-intuitive impacts of phenotype-optimal dispersal.

Population Density Equation

The diffusion term given by (10a) in our model gives the rate of change in the population density of the species due to random dispersal of the individuals. The advection term (10b) incorporates changes in population density due to the directed optimal dispersal. When the mean phenotypic potential for optimal dispersal, $(q - \mathrm{Q})/\mathrm{V}$, and the perceived environmental gradient $\widetilde{\nabla_x \mathrm{Q}}$ are both non-zero, the population density undergoes a directional change. If $(q - \mathrm{Q}) > 0$, the whole

population moves in the direction of the environmental gradient. If $(q - \mathrm{Q}) < 0$, the population moves in the opposite direction. In either case, the mean mismatch $|q - \mathrm{Q}|$ is reduced through the directed movements. The term $-\frac{\mathrm{S}}{2}(q - \mathrm{Q})^2$ in (10c) shows the effects of natural selection in reducing the population density when the trait mean q differs from the optimum Q.

The term $-\frac{\mathrm{S}}{2}v$ in (10c) captures the effect of the phenotypic load imposed by natural selection on the population growth, compared with a monomorphic population. Note that greater values of trait variance create stronger phenotypic loads. The term $\frac{\mathrm{R}}{\mathrm{K}}\sqrt{\mathrm{V}/(v + \mathrm{V})}n$ in (10c) represents the decrease in population fitness due to the average intraspecific competition between the individuals. Unlike the phenotypic load, the average intraspecific competition becomes weaker when trait variance becomes larger. This is because larger trait variance implies greater average difference between phenotypes and hence less competition load due to (4). Moreover, when individuals' phenotype (resource) utilization variance V becomes smaller, that means when the individuals become more specialist, the average competition becomes weaker. This is because there is lower chance that specialists will share utilizing same resources. Note that the competitive release at small values of V can allow for population density to increase significantly above K—which is the carrying capacity we define for sufficiently competitive (generalist) individuals with $\mathrm{V} \gg 1$.

Trait Mean Equation

The terms (11a) and (11b) show how the random gene flow caused by random dispersal affects the rate of change of trait mean, or equivalently, the local adaptation rate of the population. The divergence term in (11a) implicates the homogenizing effect of random gene flow. Since population density changes sharply at the range margin, $\nabla_x \log n$ in (11b) is significantly larger near the edge of the population, compared with the core. Therefore, (11b) effectively models the asymmetric core-to-edge gene flow. In the absence of optimal dispersal, such maladaptive gene flow can potentially result in gene swamping at marginal populations (Kirkpatrick and Barton, 1997; Lenormand, 2002).

The terms (11c)–(11e) capture the effects of individual-level optimal dispersal on the population-level local adaptation. For the divergence term (11c), we can write

$$-\operatorname{div}\left(\mathrm{A}(x)\frac{v(x,t)}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}(x)}\right) = -\left\langle \frac{\nabla_x v(x,t)}{\mathrm{V}}\,,\, \mathrm{A}(x)\widetilde{\nabla_x \mathrm{Q}(x)}\right\rangle_{\mathbb{R}^{\mathrm{m}}} \tag{14a}$$

$$-\frac{v(x,t)}{\mathrm{V}}\operatorname{div}\left(\mathrm{A}(x)\widetilde{\nabla_x \mathrm{Q}(x)}\right). \tag{14b}$$

The inner product term (14a) implies that, due to directed movements, trait mean in a local population is decreased when the gradient in population's trait variance is aligned, or makes an acute angle, with the gradient in trait optimum. Trait mean increases if the two gradients make obtuse angle or, in particular, are in opposite direction to each other. Our simulation results provided in our previous work (Shirani and Miller, 2022: Figure 2) and the results given in Figure 1 below, as well as empirical observations (Takahashi et al., 2016), show that trait variance during the range expansion of a species—over a continuous habit with linearly varying trait optimum—decreases from core to edge. That means, the trait variance gradient will be in th same direction of

the environmental gradient in one side of the population, whereas it will be in the opposite direction in th other side. As a result, changes in trait mean due to (14a) will be increasing at one side and decreasing in the other. Such changes are often adaptive. For instance, trait mean in Figure 1 will increase due to (14a) in the right half of the population's range, and will decrease in the left half.

The component (14b) of (11c) shows that divergence in the perceived environmental gradient can indeed result in changes in trait mean, even if population density and trait variance are uniformly distributed. When $\mathrm{A}(x)$ is constant, as we assume throughout this work, trait mean decreases if $\mathrm{div}\left(\widetilde{\nabla_x \mathrm{Q}(x)}\right) > 0$. An illustration of the movement rates and directions that result in such a decrease in trait mean is provided in Figure S1. Similarly, trait mean increases when $\mathrm{div}\left(\widetilde{\nabla_x \mathrm{Q}(x)}\right) < 0$. Depending on the value of the trait mean, whether it is below the trait optimum or above it, such changes can be adaptive or maladaptive to the population. However, even if the pure effect of divergence in perceived environmental gradient appears to be maladpative, its combined effect with the several other components of the optimal dispersal present in th model—which jointly affect population density, trait mean, and trait variance, can still be adaptive. In this work, however, we always assume that environmental gradient is constant. This implies that $\mathrm{div}\left(\widetilde{\nabla_x \mathrm{Q}(x)}\right) = 0$ in our simulations, except at the close vicinity of the habitat boundary. Therefore, (14b) does not have any impacts in the local adaptation that we observe in our results due to phenotype-optimal dispersal.

Similar to (11b), presence of the term $\nabla_x \log n$ in (11d) implies that this term mainly represents the asymmetric core-to-edge effects of gene flow caused by directed dispersal. Due to the population's adaptation to the environment during its range expansion, $\nabla_x q$ is expected to be aligned with $\widetilde{\nabla_x \mathrm{Q}}$. Since (11d) and (11b) have opposite signs, the core-to-edge directed gene flow cause by phenotype-optimal dispersal will indeed be adaptive at the range margin, unlike the maladaptive effects of the asymmetric gene flow created by random dispersal. Therefore, (11d) represents one of the main effects of adaptive habitat choice in facilitating adaptation at range margins. Note that, (11d) further implies that larger values of trait variance provides more fuel for such adaptive effects.

Th inner product term (11e) shows how the mean phenotypic potential for optimal dispersal, $(q - \mathrm{Q})/\mathrm{V}$, and the perceived environmental gradient directly cause local adaption. Due to the expected alignment between the directions of $\nabla_x q$ and $\widetilde{\nabla_x \mathrm{Q}}$, (11e) is negative when $(q - \mathrm{Q}) > 0$, and it is positive when $(q - \mathrm{Q}) < 0$. In either case, the resulting change in trait mean due to (11e) is adaptive, that is, it decreases the magnitude of the mismatch $|q - \mathrm{Q}|$. Importantly, (11e) explicitly shows that the rate of population-level adaptation is higher when the mean phenotypic potential of individuals for optimal dispersal is greater, and their perceived environmental gradient is steeper. Note that as the population gradually adapts to the environment, the mean phenotypic potential for dispersal decreases asymptotically. However, the presence of trait variation in the population can still provide additional adaptation force due to the terms (11c) and (11d) discussed above.

Finally, the reaction term (11f) represents local adaptation by natural selection. When $(q - \mathrm{Q}) > 0$, this term forces a decreasing change in q, which results in a decrease in $|q - \mathrm{Q}|$ and enhanced

local adaptation. Similarly, when $(q - \mathrm{Q}) < 0$, the trait mean q is increased due to (11f), leading to a decrease in $|q - \mathrm{Q}|$ and better adaptation. Importantly, (11f) demonstrates the crucial role of genetic variation in enabling adaption by natural selection. The adaptation rate controlled by (11f) is directly proportional to the level of trait variance v. Larger values of v imply greater amounts of genetic variation for natural section to act upon, and hence faster adaption rates by natural selection.

Trait Variance Equation

The terms (12a)–(12c) represent the effects of random dispersal on trait variance. Along with (11a), which tends to homogenize trait mean, the divergence term in (12a) tends to homogenize trait variance. Therefore, (11a) and (12a) together capture the homogenizing effects of random gene flow on population's phenotypes. Similar to what described above for the trait mean equation, the presence of the term $\nabla_x \log n$ in (12b) implies that (12b) mainly captures the effects of asymmetric core-to-edge random gene flow on trait variance. As our simulations results shown in Figure 1 below implies, $\nabla_x \log n$ and $\nabla_x v$ are often aligned with each other during the range expansion of a population. Therefore, (12b) is expected to be positive, and hence it tends to increase trait variance at range margins, as the range expands. This can explain a traveling-wave profile for trait variance, as we observe in our simulations. The inner product term (12c) is always positive, and it depends on the magnitude of the gradient in trait mean. Since the population typically gets well-adapted at its core, $\nabla_x q$ closely follows $\nabla_x \mathrm{Q}$ at central regions of the range. As a result, (12c) is the term which is responsible for inflating trait variance at the population's core, a well-known effect of random gene flow over an environmental gradient. The steeper th environmental gradient, the larger the trait variance at central populations.

The terms (12d) and (12e) give the effects of the non-random gene flow created by optimal dispersal on the rate of change of trait variance. Since $\nabla_x q$ and $\nabla_x \mathrm{Q}$ are expected to be in the same direction as the population adapts to the environment, (12d) typically takes negative values. Therefore, trait variance is reduced due to (12d). Since $\nabla_x q$ follows $\nabla_x \mathrm{Q}$ more closely at central regions, and since v is also larger at those regions, the reduction in trait variance caused by (12d) is significantly stronger at the well-adapted core of the population. The mismatch $(q - \mathrm{Q})$ and the gradient $\nabla_x v$ in (12e) both take larger magnitudes near the range margins, where the population is less-adapted. As a result, (12e) mainly influences the trait variance at marginal populations. As discussed in our previous work (Shirani and Miller, 2022: Sect. 2.1), and shown in Figure 1 below, the decreasing core-to-edge profile of trait variance during range expansion over a constant environmental gradient (linearly changing Q) is directly related to the spatial profile of trait mean. Near the peripheral regions over which maladpative gene flow causes q to fall below the optimum Q, trait variance decreases in the direction of the environmental gradient. That means, when $(q - \mathrm{Q}) < 0$, we expect $\left\langle \nabla_x v \,,\, \widetilde{\nabla_x \mathrm{Q}} \right\rangle_{\mathbb{R}^m}$ to be negative. This implies that (12e) will be negative when $(q - \mathrm{Q}) < 0$. Oppositely, near the peripheral regions over which $(q - \mathrm{Q}) > 0$, we observe v to be increasing in the direction of the environmental gradient. As a result, (12e) will also be negative

when $(q-\mathrm{Q})>0$. Therefore, in any case, (12e) tends to decrease the trait variance, predominantly within less-adapted marginal populations.

The above discussion implies that the directed gene flow generated by phenotype-optimal dispersal reduces trait variance both within marginal and central populations. As we discussed in the Introduction section, reducing phenotypic variation within local populations is in fact one of the important predicted effects of adaptive dispersal, or in particular, matching habitat choice. The terms (12d) and (12e) in our model show how this reduction is controlled by the interaction between mean phenotypic potential for optima dispersal $(q-\mathrm{Q})/\mathrm{V}$, trait variance, perceived environmental gradient, and gradient in trait mean and variance.

The first term in (12f) models the effects of intraspecific competition in changing trait variance. Note that this term is always nonnegative, implying that competition tends to increase trait variance. This is because competition reduces the fitness of individuals with close phenotype values, while it does not significantly affect the individuals with sufficiently different phenotypes. When $\mathrm{V}\to\infty$, that means when individuals become highly generalists, the completion term in (12f) vanishes to zero and causes no inflation in trait variance. This is because competition affects the fitness of highly generalist individuals almost uniformly, as these individuals almost equally utilizes all available resources regardless of their phenotype. As a result, no phenotype variation is generated by the competition between them. When $\mathrm{V}\to 0$, that means when individuals become highly specialist, the term $\sqrt{\mathrm{V}/(v+\mathrm{V})}$ in (12f) converges to zero. However, the competitive release gained by the population when $\mathrm{V}\to 0$ allows for the maximum steady-state population density n, controlled by (10c), to take arbitrarily large values. As a result, the competition term in (12f) remains positive (non-zero) when $\mathrm{V}\to 0$. Our computations of spatially homogeneous steady-state values of trait variance imply that, when $\mathrm{V}\to 0$ trait variance increases to a finite value due to this non-zero competition term; see Figures 2b and S4b for $\mathrm{A}=0$.

The term $-\mathrm{S}v^2$ in in (12f) represents the well-known effect of natural selection in eroding genetic variation by eliminating less fit individuals. Note that, the larger the trait variance v, the stronger the effect of natural selection, and hence the higher rate of reduction in v. Finally, the presence of the constant term U in (12f) shows the effect of mutation, as a perpetual source of genetic variation within the population.

Results

In the Interpretation of the Model Equations section we discussed, separately, the interpretation of each of the terms in equations of the model, and how they will potentially impact population density, adaptation, and trait variations in a species. However, the degree to which each of these terms contribute to the range evolution of the species is largely dependent on the complicated interactions between the terms through the couplings between the equations. Intuitive interpretation of the results of such interactions carry the risk of creating misleading predictions, knowing the fact that the terms interact at significantly different timescales. In particular, whether or not phenotype-optimal dispersal will be sufficiently consequential in range expansion dynamics of the species is

hard to be predicted intuitively by inspecting the equations.

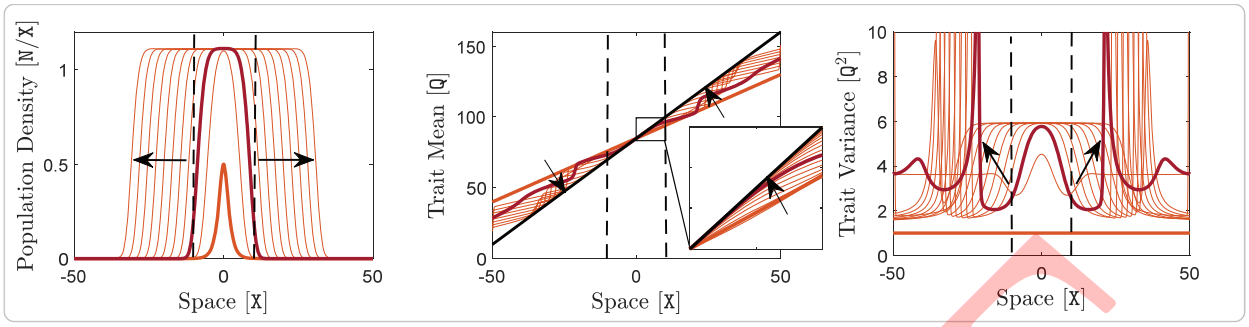
Since a rigorous analytical study of our model is rather impractical, due to its level of complexity, we numerically solve the equations (10)–(12) of the model using the biologically plausible ranges of parameter values given in Table 1. Thereby, we make observations on the evolutionary impacts of phenotype-optimal dispersal on adaptation, phenotype variation, and range evolution of a species. In all except one of our numerical studies we consider a one-dimensional continuous habitat with linearly changing environmental optimum phenotype. Our only study in a two-dimensional habitat aims to investigate the effects of phenotype-optimal dispersal in presence of habitat fragmentation. We use the same numerical scheme as we used in our previous work to solve the equations of our model. The details of the numerical scheme used to compute the solutions are given in Appendix B.

Adaptive Range Expansion Dynamics

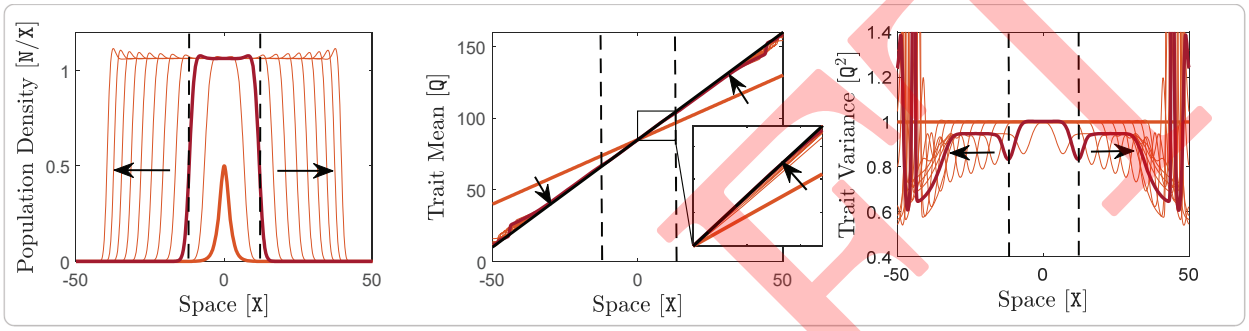
To demonstrate general effects of phenotype-optimal dispersal on the range dynamics and trait variations of a species, we perform two closely related simulations. In one of the simulations we only consider random dispersal, that means we set $\mathrm{A} = 0\,\mathtt{X}^2/\mathtt{T}$. In the other simulation we additionally include strong optimal dispersal, by setting $\mathrm{A} = 10\,\mathtt{X}^2/\mathtt{T}$. The other parameters of the model remain the same in both simulations. Other than the trait optimum Q, which is considered to be linearly increasing over Ω, the rest of the parameters are assumed to be constant. We consider a one-dimensional habitat $\Omega = (-50\ \mathtt{X}, 50\ \mathtt{X}) \subset \mathbb{R}$ with the reflecting boundary conditions (13). To make the effects of optimal dispersal strong enough to be clearly visible in our graphs, we consider a steep environmental gradient of $\mathrm{d}_x\mathrm{Q} = 1.5\ \mathtt{Q}/\mathtt{X}$. We initially introduce the species at the center of the habitat with a density given as $n(x,0) = 0.5\,\mathrm{sech}(|x|/\sqrt{2})$. We assume that the trait mean in this initial population varies linearly in space, with the constant gradient of $\nabla_x q(x,0) = 0.6\,\nabla_x\mathrm{Q}$. We further assume that the initial population is perfectly adapted to the environment at the center, $q(0,0) = \mathrm{Q}(0)$, and has a constant trait variance of $v(x,0) = 1\ \mathtt{Q}^2$. The results of our simulations over the computation time horizon of $T = 40\ \mathtt{T}$ are shown in Figure 1.

In both cases, with and without optimal dispersal, the species' population density initially grows to an upper limit set by environment's carrying capacity and level of competitive release. Afterwards, the population expands its range in the form of a traveling wave. The populations trait mean converges to the optimum trait, due to the adaptation caused both by natural selection and by phenotype-optimal dispersal, when it exists. The population's trait variance evolves a spatial profile that is decreasing from core to edge of the population. The maximum trait variance at the population core reaches a constant upper bound.

In the absence of optimal dispersal, significant maladaptation is observed in Figure 1a near the edge of the population. This maladaptation is mainly caused by asymmetric core-to-edge random gene flow, which decreases the trait mean below the optimum near the right edge and increases it above the optimum near the left edge. At the core, population density is almost uniform and hence gene flow is symmetric. This implies that random gene flow does not significantly affect the trait mean at central locations and local adaptation is maintained by natural selection. Since



(a)



(b)

Figure 1: Adaptive range dynamics of a species in a one-dimensional habitat with steep environmental gradient. Here, m = 1, Q(x) is linear in x with a relatively steep gradient of $\nabla_x \mathrm{Q} = 1.5;$, Q/X, and A(x) is constant taking different values in panels (a) and (b). The rest of the model parameters take their typical values given in Table 1. In each of the panels, evolution of the population density $n(x,t)$ is shown on the left, evolution of the trait mean $q(x,t)$ is shown in the middle, and evolution of the trait variance $v(x,t)$ is shown on the right. Panel (a) shows the range expansion dynamics of a species without optimal dispersal, A = 0 X^2/T, whereas panel (b) shows the range expansion dynamics of a species with strong optimal dispersal, A = 10 X^2/T. In all graphs, curves are shown at every 4 T, and the thick orange curves indicate the initial curves at $t = 0$ T. In the insets of trait mean graphs, curves are shown at every 1 T. Arrows show the direction of evolution in time. In each graph, a sample curve at $t = 8$ T is highlighted in red. Dashed lines indicate the effective edges of the population at $t = 8$ T, associated with the inflection points on the highlighted curve of population density. The solid black lines in the graphs of trait mean show the environmental trait optimum Q. The curves of trait variance take large values outside the effective range of species. Such values are not biologically meaningful as they do not occur within the range of the species, and have been cut to smaller values for better visualization of the meaningful parts of the graphs.

environmental gradient is steep and population is well-adapted to it at its core, random gene flow from adjacent areas generates large phenotypic variations among central individuals. Near the edges, however, trait mean fails to follow the steep gradient in the optimum trait. As a results,

trait variance decreases from core to edge in parallel with the decrease in gradient of the trait mean. Further descriptions on the profiles of trait mean and variance under random dispersal are available in our previous work (Shirani and Miller, 2022: Sect. 4.1).

In the presence of strong phenotype-optimal dispersal, as Figure 1b shows, the traveling-wave-based range expansion dynamics of the species remains generally similar to the random dispersal case. Nevertheless, important differences can be observed. Comparing the the insets in the graphs of trait mean in Figures 1a and 1b shows that phenotype-optimal dispersal significantly increases the local adaptation rate of the population. Convergence to the environmental optimum at the population's core occurs in almost one generation time (ecological time-scale) with optimal dispersal, whereas same level of adaptation takes more than ten generations to occur in the absence of optimal dispersal. The level of maladaptation at range margins is also substantially decreased by the directed gene flow created by phenotype-optimal dispersal. These are all consistent with the general predictions on the adaptive effects of matching habitat choice, as discussed in the Introduction section, and as predicted by inspecting the equations of our model. Importantly, by comparing the curves of trait variance in Figures 1a and 1b, we also observe that the phenotypic assortment resulting from the preferential movements under optimal dispersal significantly reduces the trait variance within the population. Moreover, the curves of population density in Figures 1a and 1b show that the reinforced adaptation by optimal dispersal increases the overall range expansion speed of the population, especially at earlier stages of the population's establishment in the habitat. The maximum population density, however, is not notably impacted by optimal dispersal at the environmental gradient we simulated in Figure 1.

The pronounced effects of phenotype-optimal dispersal we observed in Figure 1 were obtained at the environmental gradient $\nabla_x \mathrm{Q} = 1.5\ \mathtt{Q/X}$. The steepness of a the environmental gradient is in fact a key factor in determining whether or not phenotype-optimal dispersal will be sufficiently consequential in a species' range evolution. The perceived forced for optimal dispersal (9) is directly proportional to the (perceived) magnitude of the environmental gradient. Moreover, the phenotypic potential $(p - \mathrm{Q})/\mathrm{V}$ in (9) is directly influenced by the level of phenotypic variation in the population, which in turn can be substantially inflated by random gene flow under steep environmental gradients. Estimates of realistic values for the slope of environmental optimum gradient in nature, however, are not widely available in the literature—noting that our choices of units for $\nabla_x \mathrm{Q}$ requires joint measurements of optimal trait values, dispersal distance, and generation time. Yet, based on some available data, in our previous work we discussed that a plausible range of values for $\|\nabla_x \mathrm{Q}\|_{\mathbb{R}^m}$ can approximately be between 0 and 2 $\mathtt{Q/X}$, (Shirani and Miller, 2022: Sect. 3.2). That means, the gradient $\nabla_x \mathrm{Q} = 1.5\ \mathtt{Q/X}$ we used for the results shown 1 is very steep; possibly associated with long-range dispersal of birds over elevation gradients, as an example. In supplementary Figure S2, we show our simulation results for a reasonably shallower environmental gradient of $\nabla_x \mathrm{Q} = 0.2\ \mathtt{Q/X}$, which might be more typically observed in nature. We see that phenotype-optimal dispersal does not significantly impact the species's range dynamics in this case.

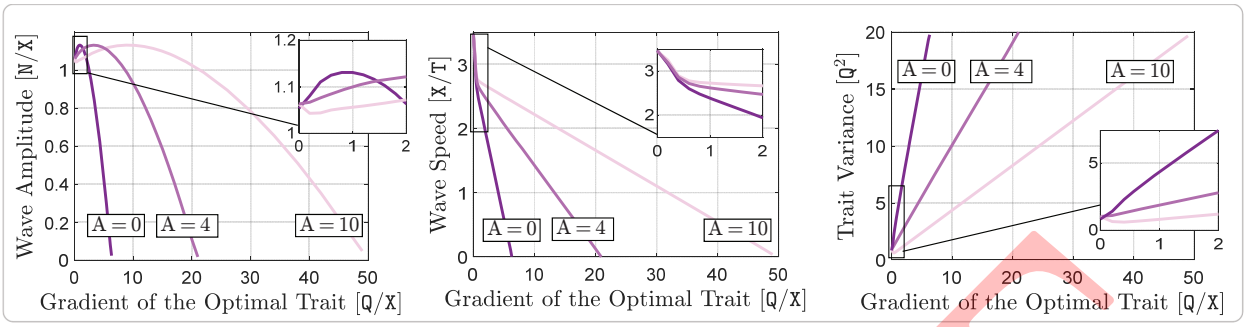
The phenotypic potential $(p - \mathrm{Q})/\mathrm{V}$ for optimal dispersal in (9) is stronger when phenotype utilization variance V is smaller, that means when individuals are specialists. To see if phenotype-

optimal dispersal by specialist individuals can significantly influence the range dynamics when environmental gradient is not very steep, we simulate our model with the relatively small value of $\mathrm{V} = 1\ \mathtt{Q}^2$ under the typical gradient of $\nabla_x \mathrm{Q} = 0.2\ \mathtt{Q/X}$. The results are shown in supplementary Figure S3. We observe similar effects to those shown in Figure 1. That is, phenotype-optimal dispersal facilitates local adaptation of specialists in ecological time-scales, reduces their average maladaptation at range margins, and reduces their within-population trait variance. Unlike what we observed in Figure 1, the population's range expansion speed is not considerably increased though. In the absence of optimal dispersal, we see in Figure S2a that the population density rises significantly above $\mathrm{K} = 1\ \mathtt{N/X}$. This is due to the population's ecological release gained by less-competitive specialists. However, phenotype-optimal dispersal limits the level of competitive release, as it counteracts the effects of smaller phenotype utilization variance (less competition) by significantly reducing phenotype variation within the population. Phenotypically close individuals can still remain sufficiently competitive even when they are specialists and utilize less amount of common resources. As a result, we see in Figure S2b that population density is still approximately bounded by the carrying capacity $\mathrm{K} = 1\ \mathtt{N/X}$ when the specialist individuals disperse optimally.

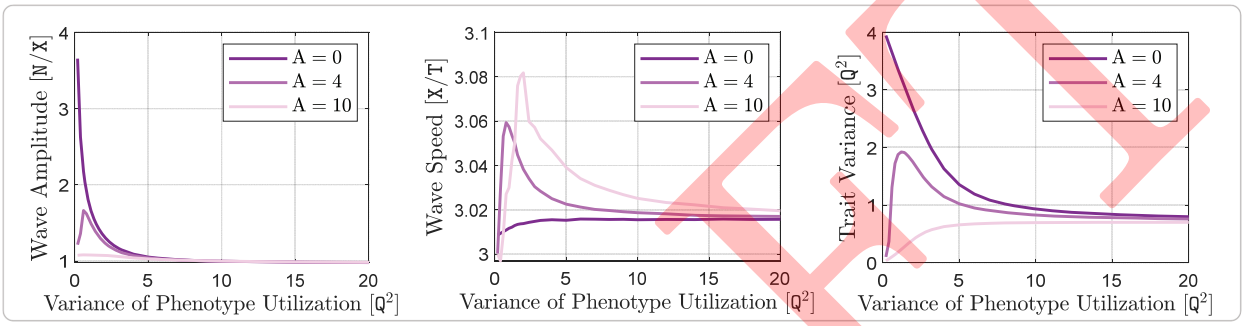
To further investigate whether or not phenotype-optimal dispersal can have sufficiently consequential impacts on a species' range expansion dynamics, we repeat our simulations for different magnitudes (slopes) of the environmental gradient, as well as different values of the phenotype utilization variance. In each case, we consider three different levels of optimal dispersal propensity, $\mathrm{A} = 0\ \mathtt{X}^2\mathtt{/T}$ (no optimal dispersal), $\mathrm{A} = 4\ \mathtt{X}^2\mathtt{/T}$ (medium optimal dispersal), and $\mathrm{A} = 10\ \mathtt{X}^2\mathtt{/T}$ (strong optimal dispersal). We run each simulation for a sufficiently long period of time, so that the initial transient states pass. We uses the computed curves near the end of each simulation to measure approximate speed and amplitude (peak value) of the traveling waves of population density, as well as maximum trait variance attained at the population's center. The results are shown in Figure 2.

The curves of wave amplitude versus environmental gradient shown in Figure 2a show that the population density and chance of survival at extreme environmental gradients is drastically increased when propensity for optimal dispersal is increased. Similarly, population's expansion at extremely steep gradients is much faster when the individuals disperse optimally. The intraspecific trait variance is also controlled to much lower, and much more reasonable, levels by the assortment effects of optimal dispersal at extreme gradients. Note that the relatively sharp decline in the amplitude of the population density waves at extreme gradients is predominantly due to the phenotypic load $-\frac{\mathrm{S}}{2}v$ in (10c) which increases as v increases with gradient. That is why significantly slower decline in wave amplitude is observed when optimal dispersal is strong and effectively controls the rise in trait variance to more moderate values.

The profound improvements to species' range expansion capacity that we observe in Figure 2a in presence of phenotype-optimal dispersal come with an important caveat. Based on our previous discussions on plausible ranges of values of environmental gradients (Shirani and Miller, 2022: Sect. 3.2), such improvements mainly occur at exceedingly steep gradients which are unlikely to be biologically realistic. As the insets in the graphs of Figure 2a show, when the gradient in



(a)



(b)

Figure 2: Effects of optimal dispersal on range expansion waves and maximum intraspecific trait variance of a species. Here, $\mathrm{m} = 1$, and $\mathrm{A}(x)$ takes three different constant values $0\,\mathtt{X}^2/\mathtt{T}$, $4\,\mathtt{X}^2/\mathtt{T}$, and $10\,\mathtt{X}^2/\mathtt{T}$. The trait optimum $\mathrm{Q}(x)$ is linear in x, with variable gradient in panel (a) and the constant gradient of $\nabla_x \mathrm{Q} = 0.2\,\mathtt{Q}/\mathtt{X}$ in panel (b). The phenotype utilization variance takes the constant value $\mathrm{V} = 4\,\mathtt{Q}^2$ in panel (a), and is variable in panel (b). The rest of the model parameters take their typical values given in Table 1. In each panel, variations in the speed of the traveling waves of population density are shown on the left, variations in the amplitude of the traveling waves are shown in the middle, and variations in the maximum intraspecific trait variance is shown on the right. Panel (a) shows the effects of different levels of optimal dispersal at different magnitudes of the environmental gradient $\nabla_x \mathrm{Q}$. Panel (b) shows the effects of different levels of optimal dispersal at different values of the species' phenotype utilization variance V.

trait optimum takes more reasonable values between 0 and 2 $\mathtt{Q}/\mathtt{X}$, the improvements are much less pronounced. We observe sufficiently significant changes in the expansion speed only at steep gradients greater than 1 $\mathtt{Q}/\mathtt{X}$. At more typically observed gradients below 1 $\mathtt{Q}/\mathtt{X}$, phenotype-optimal dispersal still facilitates local adaptation and significantly reduces within-population trait variance, but such impacts appear to be less consequential to population's range expansion capacity.

In a typically shallow environmental gradient of $\nabla_x \mathrm{Q} = 1\ \mathtt{Q}/\mathtt{X}$, Figure 2b shows that phenotype-optimal dispersal does not significantly increase range expansion speed, even when the phenotype utilization variance in the population is small. In the absence of phenotype-optimal dispersal, that is when $\mathrm{A} = 0\ \mathtt{X}^2/\mathtt{T}$, the competitive release at small values of V results in significant increase in

both population density and trait variance as $\mathrm{V} \to 0$. The directed gene flow created by optimal dispersal, however, substantially depresses the effects of such competitive release. It controls the level of increase in trait variance to much lower values. As a result, competition remains sufficiently strong even though small values of V tend to release the individuals from competition. When optimal dispersal propensity is very strong, that is $\mathrm{A} = 10\ \mathtt{X}^2/\mathtt{T}$, the effect of competitive release is fully depressed and the amplitude of population density waves remains close to the carrying capacity $\mathrm{K} = 1\ \mathtt{N}/\mathtt{X}$ for all values of V. When n is controlled to almost constant values, decreasing V in the first (competition) term in (12f) to zero will vanish the effect of competition in inflating trait variance. As a result, with strong optimal dispersal, steady-state trait variance will converge to a value mainly controlled by random gene flow (term (12c) in our model) and mutation-section balance (term $-\mathrm{S}v^2 + \mathrm{U}$ in (12f)). Since both U and the environmental gradient are relatively small in Figure 2b, this value is relatively small. This explains the sharp decline we observe in the curves of trait variance in Figure 2b as $\mathrm{V} \to 0$, both for $\mathrm{A} = 10\ \mathtt{X}^2/\mathtt{T}$ and for $\mathrm{A} = 4\ \mathtt{X}^2/\mathtt{T}$.

We note that the curves of wave amplitude and trait variance in Figure 2 were computed approximately, by running the simulations for sufficiently long time and measuring the (almost steady) values of these quantities at the center of the population at the end of the simulation. These curves can more accurately be computed by solving the equations of the spatially homogeneous equilibrium of the model, similar to our computations in our previous work (Shirani and Miller, 2022: Sect. 4.2). Although we avoid repeating such analysis here, we give the resulting curves in supplementary Figure S4. We observe very close agreement between the approximate curves shown in Figure 2 and the more accurate ones shown in Figure S4.

Directed Gene Flow and Local Adaptation at Rang Margins

In previous section, we made general observations on the effects of phenotype-optimal dispersal on local adaptation of the species, using curves of trait mean. Here, we take a closer look at how the directed gene flow generated by optimal dispersal influences the overall asymmetric core-to-edge gene flow and adaptation at range margins. We use equation of the trait mean (11) to see how the mean adaptation rate, $\partial_t q$, is affected by random and non-random (directed) components of gene flow, particularly at range margins. The sum of the terms (11a) and (11b) presents the contribution of random gene flow in determining the rate of change of trait mean, whereas the sum of the terms (11c)–(11e) presents the contribution of directed gene flow. The sum of all terms (11a)–(11e) determines how the trait mean is changed due to individuals' dispersal, both random and directed. We use the same simulation layout as we used for the range dynamics results shown in Figure 1b, and the same solutions of the model we computed in there, to illustrate each of these three contributions separately. The resulting curves, as the population expands its range over time, is shown in Figure 3. Due to symmetry, the curves are shown only for the right half of the habitat. Positive values in each curve at a point x imply that the corresponding component of gene flow represented by the curve tends to increase the trait mean q at x. Oppositely, negative values imply tendency to decrease q. Therefore, since the initial profile of trait mean at $t = 0\ \mathtt{T}$, as shown

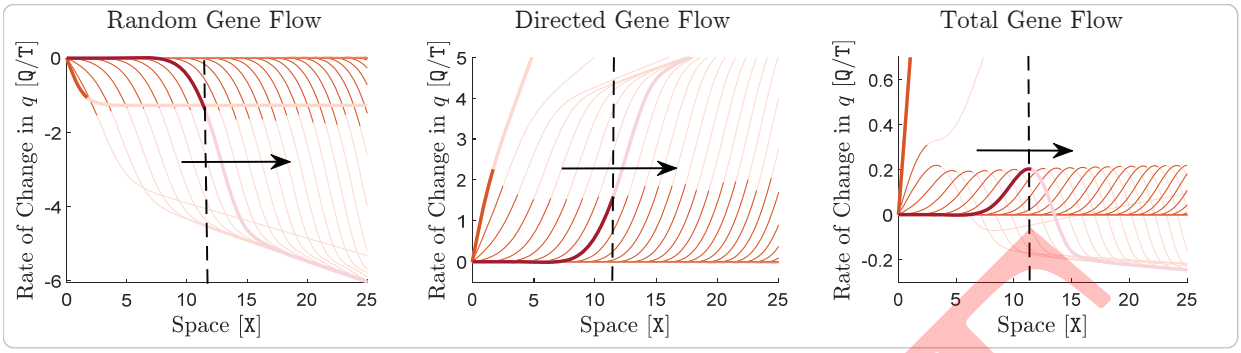
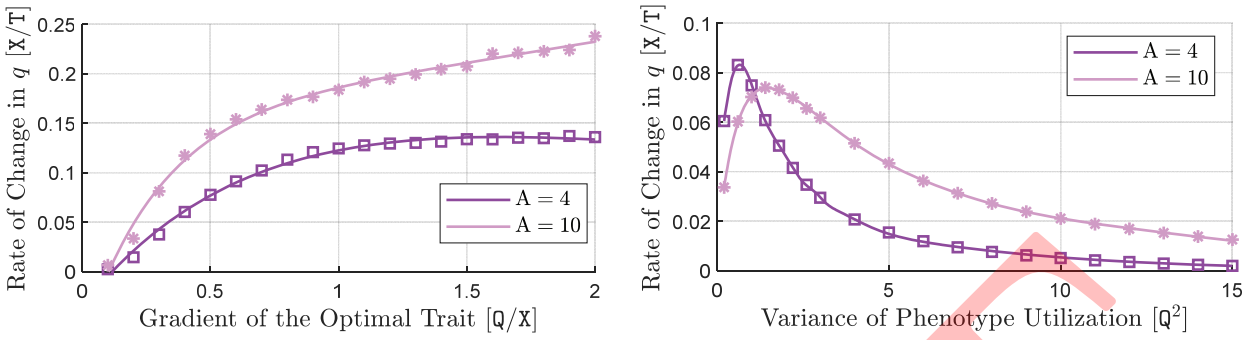


Figure 3: Effects of gene flow on local adaption of a species. Contribution of random gene flow to the rate of change of the trait mean, $\partial_t q$, is shown on the left. The curves of this graph are computed as the sum of the terms (11a) and (11b), which capture the effects of random dispersal on $\partial_t q$. In the middle, contribution of directed gene flow to $\partial_t q$ is shown. The directed gene flow is generated by optimal dispersal and is computed as the sum of the terms (11c)–(11e). Contribution of the total (net) gene flow to $\partial_t q$, that is the sum of the curves shown in the middle and the left graphs, is shown on the right. The graphs correspond to the same simulation of species' range dynamics given in Figure 1b, that is, when $\nabla_x \mathrm{Q} = 1.5\,\mathtt{Q/X}$ and $\mathrm{A} = 10\,\mathtt{X}^2/\mathtt{T}$. In all graphs, the evolution of the computed curves are shown only on the right half of the habitat. The curves extend symmetrically about the origin to the left half of the habitat. Moreover, the portion of each curve that lies outside the effective range of the species, that means over the regions where the population density is approximately zero, has been made transparent. In all graphs, curves are shown at every 1 $\mathtt{T}$. The descriptions of the highlighted curves, arrows, and dashed lines are the same as those provided in Figure 1b.

in Figure 1b, is below the trait optimum Q over the right half of the habitat, positive values of the curves imply adaptive effects (increasing q toward Q), and negative values imply maladaptive effects.

As we saw in Figure 1b, convergence of q to Q at population's core occurs approximately in ecological time scales, due to the strong effect of phenotype-optimal dispersal at the steep gradient $\nabla_x \mathrm{Q} = 1.5\,\mathtt{Q/X}$. This convergence can also be clearly observed in Figure 3. The rate of change in q quickly approaches to zero at the core of the population after a couple of generations. At range margins, however, Figure 1b shows that adaptation never occurs perfectly. The core-to-edge random gene flow created by random dispersal is always maladaptive, whereas the directed gene flow generated bt phenotype-optima dispersal is always adaptive. Importantly, the total gene flow to range margins appears to be always adaptive, implying that phenotype-optimal dispersal not only compensates for the maladaptive effects of random movements, but also reverses their effects in local adaptation of marginal populations. Similar observations are made when we analyze the adaptive/maladaptive effects of different components of gene flow associated with the simulation results shown in supplementary Figure S3b, that is, when environmental gradient is shallow ($\nabla_x \mathrm{Q} = 0.2\,\mathtt{Q/X}$) but individuals are highly specialists ($\mathrm{V} = 1\ \mathtt{Q}^2$). The resulting curves are shown in



(a) **(b)**

Figure 4: Effects of directed gene flow on local adaption of a species at the edge of its range. Contribution of the total (net) gene flow to the rate of change of trait mean, $\partial_t q$, at the edge of the species' range is shown in each graph, for two different values of A. In (a), the species' phenotype utilization variance takes its typical value of $\mathrm{V} = 4\,\mathtt{Q}^2$, and the magnitude of $\nabla_x \mathrm{Q}$ is made variable. In (b), the environmental gradient takes its typical value of $\nabla_x \mathrm{Q} = 0.2\,\mathtt{Q}/\mathtt{X}$, and V is made variable. The rest of the model parameters take their typical values given in Table 1, with $\mathrm{m} = 1$. To obtain each data points of the graphs (a) and (b), the contribution of the total gene flow is first computed for the right half of the species' range, as the sum of all terms (11a)–(11e) in (11). The graphs shown on the right of Figure 3 show samples of such results. At each value of $\nabla_x \mathrm{Q}$ in (a) and each value of V in (b), the simulations are performed for sufficiently long periods of time so that the species' range expansion dynamics reaches an approximate steady-state. The value of the total gene flow contribution to $\partial_t q$ at the edge of the species' range, obtained at the end of each simulation, is then shown in graphs (a) and (b).

Figure S5. The difference in this case—possibly because of the complicated interaction between the effects of optimal dispersal and competition—is that convergence at the core shows a small overshooting dynamics, during which the curves of directed gene flow take negative values (but still adaptive) to decrease the overshot q back to Q.

To further investigate if the adaptive effects of directed gene flow remain sufficiently significant at shallower (more typical) environmental gradients or when individuals are non-specialist, we repeat similar analysis as shown in Figures 3 and S5 for different slopes of the environmental gradient and different values of V. The sample curves in Figure 3 show that the maximum adaptive (or maladaptive) effects of directed (or random) gene flow occur at the edge of the population. Therefore, we use the value of the total contribution of gene flow to $\partial_t q$, computed at the edge of the population, as our reference for measuring the significance of phenotype-optimal dispersal in facilitating adaptation at range margins. The resulting curves are shown in Figure 4.

We see in Figure 4 that the total gene flow remains adaptive to the range margins, even when environmental gradient is fairly shallow or individuals are generalists. This implies that phenotype-optimal dispersal is quite effective in compensating for the maladpative effects of core-to-edge random gene flow. Yet, the curves shown in Figure 4, along with our previous observations through

Figure 2, suggest that the adaption facilitated by phenotype-optimal dispersal at range margins is sufficiently consequential to enhancing range expansion capacity of the population primarily when environmental gradient is steep (slopes greater than 1 $\mathtt{Q}/\mathtt{X}$).

Range Dynamics Under Abrupt Climate Fluctuations

Rapid adaptation of individuals within the time span of a single generation, as facilitated by adaptive dispersal strategies such as matching habitat choice, is predicted to be pivotal to survival of a population under climate change, particularly when changes are sharp and frequent (Bonte et al., 2012; Edelaar and Bolnick, 2019; Jacob et al., 2017; Nicolaus and Edelaar, 2018). To make observations on how phenotype-optimal dispersal will affect the range dynamics of a species under abrupt climatic changes, we simulate our model in a one-dimensional habitat with a linear trait optimum of steep gradient $\nabla_x \mathrm{Q} = 1.5\ \mathtt{Q}/\mathtt{X}$, which periodically fluctuates up and down without changing its gradient. For this, we initialize our simulation at $t = 0\ \mathtt{T}$ with a relatively established population at the center of the habitat. This initial population is obtained as the curves at $t = 4\ \mathtt{T}$ of a preliminary simulation similar to the one shown in Figure 1, with $n(x, 0) = 0.8\,\mathrm{sech}(|x|/\sqrt{2})$, $q(0, 0) = \mathrm{Q}(0)$, $\nabla_x q(x, 0) = 0.7\,\nabla_x \mathrm{Q}$, and $v(x, 0) = 1\ \mathtt{Q}^2$. The thick orange curves shown in Figure 5 indicate the initial populations. To simulate abrupt temporal fluctuations in the environment, we uniformly shift up the line of trait optimum Q by a certain fluctuation amplitude at the beginning of a fluctuation period, and then shift it back down by the same amplitude at the middle of the period. We repeat these fluctuations periodically, starting at $t = 0\ \mathtt{T}$, with a relatively short period of $2\ \mathtt{T}$.

Figure 5a shows the simulation results for fluctuation amplitude of $5\ \mathtt{Q}$, when dispersal is only random, $\mathrm{A} = 0\,\mathtt{X}^2/\mathtt{T}$. The high level of maladaptation that is abruptly induced in the population at $t = 0\ \mathtt{T}$, when the optimum gradient is shifted up, quickly reduces the population density and its expansion speed. Yet, directed natural selection acts on the large deviation of trait mean from the trait optimum, and the population gradually adapts to the new environmental optimum. However, since period of the fluctuations is relatively short, the population will not fully recover its peak density before experiencing another abrupt change in the environment. The periodically repeated loss-and-recovery dynamics of the population density eventually reaches a steady-state, at which the peak population density fluctuates between a fixed high and a fixed low value as the populations expands its range. The red and blue curves in Figure 5a show samples of population density profile at such high and low extremes.

Strong phenotype-optimal dispersal, $\mathrm{A} = 10\,\mathtt{X}^2/\mathtt{T}$, significantly enhances the range expansion capacity of the population under the abrupt environmental fluctuations that we simulate here, by increasing both the expansion speed and the peak population density. The results are shown in Figure 5b. This is because the large phenotype-environment mismatch perceived by the individuals, immediately after an abrupt shift in the optimum phenotype, creates a strong phenotypic potential for the individuals to disperse to better-matching locations. As a result, the population can rapidly adapt to the new environmental optimum when a change occurs, and hence it loses much less

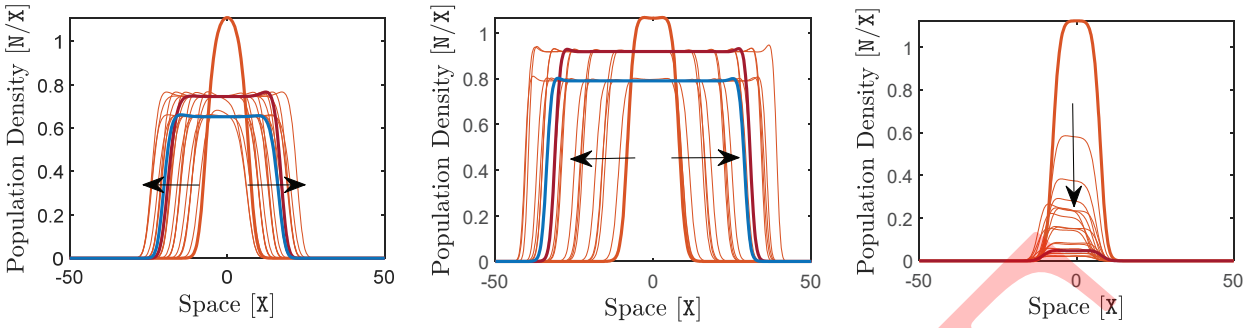
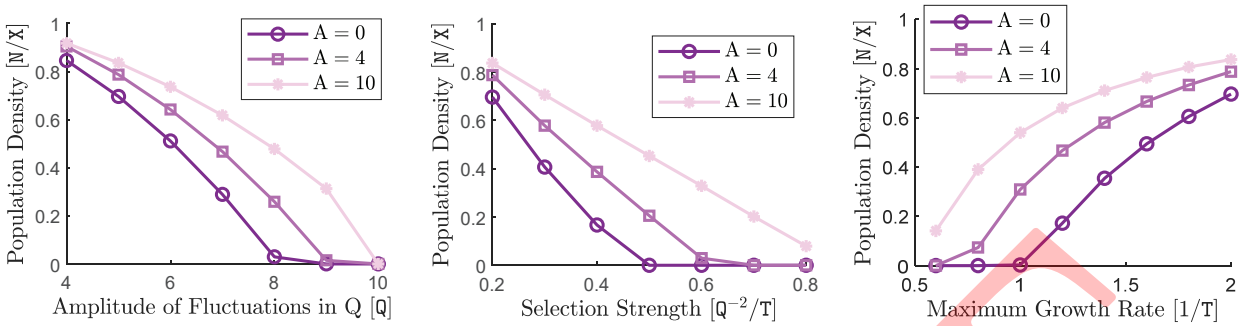


Figure 5: Range dynamics of a species under periodic abrupt fluctuations in the environmental trait optimum. Here, $m = 1$, $\nabla_x \mathrm{Q} = 1.5\,\mathrm{Q/X}$, and A takes different values in each graph. The rest of the model parameters take their typical values given in Table 1. Graph (a) shows range dynamics without optimal dispersal, that is $\mathrm{A} = 0\,\mathrm{X}^2/\mathrm{T}$. Graph (b) shows range dynamics with strong optimal dispersal, $\mathrm{A} = 10\,\mathrm{X}^2/\mathrm{T}$. Graph (c) shows extinction of a species with the typical value $\mathrm{A} = 4\,\mathrm{X}^2/\mathrm{T}$ due to high-amplitude environmental fluctuations. In all graphs, the period of abrupt fluctuations in the trait optimum is $2\,\mathrm{T}$. At the beginning of each period, the trait optimum Q is shifted up by a preset fluctuation amplitude and remains at this value for the first half of the period. Then, it is shifted down by the same amplitude to the initial value and remains at this value for the second half of the period. The fluctuation amplitude is set equal to $5\,\mathrm{Q}$ in (a) and (b), and equal to $9\,\mathrm{Q}$ in (c). The thick orange curves indicate the initial curves at $t = 0\,\mathrm{T}$ and arrows show the direction of evolution in time. In (a) and (b), curves are shown at every $1.5\,\mathrm{T}$, and two sample curves are highlighted at $t = 20\,\mathrm{T}$ (in red) and $t = 20.5\,\mathrm{T}$ (in blue). In (c), curves are shown at 20 logarithmically distributed time samples, with the first curve after the initial curve being shown at $t = 0.1\,\mathrm{T}$. A sample curve at $t = 20\,\mathrm{T}$ is also highlighted in red.

of its density. However, in agreement with our observations in the previous studies described above, the significant effects of optimal dispersal that we observe here are in the presence of a steep environmental gradient of $\nabla_x \mathrm{Q} = 1.5\ \mathrm{Q/X}$. In shallower (more typical) gradients, as the supplementary Figure S6 shows, the effects are much less significant.

When the amplitude of the environmental fluctuations is exceedingly large, the density loss due to the excessive level of maladaptation after each abrupt change will be too high to be fully recovered by the adaptation that occurs afterwards. As a result, the population will not be able to reach a persistent state and its density keeps decreasing due to the repeated changes in the environment. Eventually, the population becomes extinct. Figure 5c shows an example of such extinction dynamics, when phenotype-optimal dispersal is still fairly strong, $\mathrm{A} = 4\,\mathrm{X}^2/\mathrm{T}$, but environmental fluctuations occur with a large amplitude of $9\,\mathrm{Q}$. To see if phenotype-optimal dispersal increases the chance of survival when amplitude of the fluctuations is large, we repeat our simulation associated with Figure 5 for different values of fluctuation amplitudes. For each value, we compute the average value of the fluctuations in population density when it reaches a steady-state, and show it



(a) **(b)** **(c)**

Figure 6: Steady-state mean population density of a species under periodic abrupt fluctuations in the environmental trait optimum. Here, $m = 1$, $\nabla_x \mathrm{Q} = 1.5\,\mathtt{Q}/\mathtt{X}$, and A takes different values in each graph. The same periodic abrupt fluctuations as described in Figure 5 is applied to the trait optimum Q, with the typical amplitude of $5\,\mathtt{Q}$ in (b) and (c) and variable amplitude in (a). Except for S and R which are made variable in (b) and (c), respectively, the rest of the model parameters take their typical values given in Table 1. At each value of the variable parameter shown at the horizontal axis of each graph, the simulation is run for a sufficiently long period of time so that the amplitude of the fluctuations in population density of the species reaches a steady-state. The minimum and maximum values of such periodic fluctuations (peaks of the blue and red curves as shown in Figures 5a and 5b) are calculated near the end of the simulation, and their average value is shown by different markers for each value of A, along with an interpolated solid line. In (a), the steady-state mean value of the population density is shown with respect to changes in the amplitude of the abrupt fluctuations in Q. In (b), the steady-state mean value of the population density is shown with respect to changes in the strength of stabilization selection, S. In (c), the steady-state mean value of the population density is shown with respect to changes in the maximum growth rate of the species, R.

in Figure 6a. We see that when the amplitude of the environmental fluctuations is not very large, optimal dispersal does not significantly increase the sustained density of the population. At very large fluctuation amplitudes, however, optimal dispersal can largely increase the survival chance of the species.

Strong natural selection amplifies the effects of maladaptation induced by abrupt environmental changes, as implied from the term $\frac{\mathrm{S}}{2}(q - \mathrm{Q})^2$ in (10c). As a result, when S is large, the population suffers from a greater density loss after each shift in the trait optimum. However, larger values of S expedite adaptation to the new environment, due to (11f). To see how these contradicting effects of stronger natural selection affect population density under environmental fluctuations, and whether or not phenotype-optimal dispersal can be sufficiently advantageous to population survival, we repeat the simulation associated with Figure 5 for different values of S and measure the steady-state average value of the fluctuations in population density. The results are shown in Figure 6b. In general, we see that stronger natural selection reduces the sustained population density under

the abrupt fluctuations we simulated. When stabilizing section is weak, as we typically observe in nature (Kingsolver et al., 2001), phenotype-optimal dispersal does not significantly increase mean population density. However, under very strong stabilizing selection, we observe that the population's chance of survival is substantially increased if its individuals disperse optimally.

The maximum growth rate of the population is a key factor in accelerating population density recovery following density losses caused by environmental changes. Slowly growing populations will have lower chance of recovering their full density before suffering form another loss. To see how this impacts the population survival, we repeat our simulations for different values of R and show the steady-state average value of population density fluctuations in Figure 6c. We see that slowly-growing populations maintain a significantly lower density under environmental fluctuations, compared with the fast-growing populations. However, phenotype-optimal dispersal substantially increases the survival chance of slowly-growing species. In particular, with the relatively large fluctuation amplitude of 5 $\mathtt{Q}$ that we considered in our simulation, observe that a randomly dispersing population becomes extinct when R takes a relatively low value of approximately 1 $\mathtt{T}^{-1}$, whereas a population with strong optimal dispersal can persist with sufficiently large density. We note that, with our choice of generation time as the unit of time $\mathtt{T}$, the maximum intrinsic growth rate argued to be a demographic invariant within some homogeneous taxonomic groups (Niel and Lebreton, 2005). In particular, a slow growth rate of $\mathrm{R} = 1\ \mathtt{T}^{-1}$ is observed in a variety of taxa such as birds, sharks, and mammals (Dillingham et al., 2016; Niel and Lebreton, 2005; Shirani and Miller, 2022), which are often sufficiently mobile to evolve optimal dispersal strategies.

Fragmented Habitat

In highly fragmented environments, the cost of randomly moving from one habitat patch to reach another patch is particularly high. Nevertheless, optimal habitat selection can also be harder in such environments as optimal detection of matching habitat patches can be substantially imperfect and dispersal mortality can be higher (Cote et al., 2017). However, the enhanced local adaptation facilitated by optimal dispersal can still be imagined to be sufficiently beneficial to compensate for such costs. Therefore, it is argued that habitat choice behavior and adaptive dispersal strategies should also be selected in fragmented landscapes, and might be essential for the persistence of populations in such environments (Bonte et al., 2012; Cote et al., 2017). Here, we investigate whether or not phenotype-optimal dispersal can be of particular importance to a population's adaptation and range expansion capacity when the available habitat is highly fragmented.

We consider a two-dimensional habitat $\Omega = (-50, 50) \times (-50, 50)$, with a trait optimum profile that changes linearly in x_1-direction (horizontal axis) and is constant in x_2-direction (vertical axis). We set a relatively steep gradient of $\partial_{x_1}\mathrm{Q} = 1\,\mathtt{Q}/\mathtt{X}$ in x_1-direction. Since habitat loss directly affects the carrying capacity of the environment (Baguette et al., 2013), we simulate habitat fragmentation by setting a patchy profile for the carrying capacity parameter $\mathrm{K}(x)$. This profile is shown in supplementary Figure S7, and its specific pattern can also be approximately seen through the last frames (at $t = 40\ \mathtt{T}$) in Figure 7. If patch sizes are large in both directions, relative to

the average (random) dispersal distance per generation, the range dynamics inside each patch is expected to be similar to the continuous habitat cases we studied in previous cases. Therefore, to make the effect of fragmentation sufficiently pronounced, we consider relatively narrow patches of width 2 $\mathtt{X}$—noting that the random dispersal coefficient D takes the typical value of $1I_2\ \mathtt{X}^2/\mathtt{T}$, where I_2 denotes the 2×2 identity matrix. We let the length of these rectangular patches take values between 10 $\mathtt{X}$ and 15 $\mathtt{X}$. On the right-half of the habitat, we layout the patches horizontally, so that they are stretched in the direction of the environmental gradient. On the left-half of the habitat, we layout the patches vertically to be stretched perpendicular to the environmental gradient. We consider this specific layout to further observe if the alignments of the patches with the environmental gradient can have a particular impact on the range expansion dynamics. We arrange the rectangular patches side-by-side, with no gap between them. However, over each patch, we smoothly decrease the value of K from 1 $\mathtt{N}/\mathtt{X}^2$ at the patch center to 0.05 $\mathtt{N}/\mathtt{X}^2$ at the patch edge. This leaves fairly inhabitable regions of low carrying capacity between patch cores, as shown in Figure S7. Finally, we make an exception for the size of the patch located at the center of the habitat, and let it be sufficiently large. We initialize the population in this central patch, so that it can get well-established before expanding through the highly fragmented areas.

Figure 7 shows the simulation results for a species with only randomly dispersing individuals (upper panel) and for a species with phenotype-optimal dispersal (lower panel). To make comparison with range expansion in a continuous habitat, we also perform the simulations with constant carrying capacity of $\mathrm{K} = 1\ \mathtt{N}/\mathtt{X}^2$ and indicate the leftmost and rightmost edges of the population by blue bars in Figure 7. The population establishes itself in the central patch to its almost maximum capacity of approximately 1 $\mathtt{N}/\mathtt{X}^2$. However, it spreads over the fragmented regions with much lower density. This is mainly due to a significant level of maladaptation that is maintained by the homogenizing effect of random gene flow. Figure S8 shows the phenotype-environment mismatch $q - \mathrm{Q}$ at the end of the simulations, when the mismatch has approximately reached a steady state over the whole habitat. The steady-state level of mismatch is low at the center of the patches. However, it increases rather sharply towards the edges. Since patches are relatively narrow, this results in a relatively high level of overall maladaptation over a large area of the patches, and hence significant loss in population density.

Phenotype-optimal dispersal facilitates faster adaptation to new patches and thereby enhances the transient expansion dynamics of the population, as observed through the increased expansion speed in the lower panel of Figure 7. However, we observe that optimal dispersal does not significantly improves the steady-state level of adaptation (shown in Figure S8) and population density within the patches. The maximum population density that we observe in patches of the fragmented region at $t = 40\ \mathtt{T}$ is approximately $0.45\ \mathtt{N}/\mathtt{X}^2$ when $\mathrm{A} = 10\,\mathtt{X}^2/\mathtt{T}$, which is only slightly larger than the maximum density of $0.41\ \mathtt{N}/\mathtt{X}^2$ observed when $\mathrm{A} = 0\,\mathtt{X}^2/\mathtt{T}$. This is because, due to the narrow widths of the patches, the directed gene flow created by phenotype-optimal dispersal is not sufficiently strong to effectively compensate for the maladaptive core-to-edge random gene flow within the patches.

Although with much lower density, Figure 7 shows that the population expands its range over

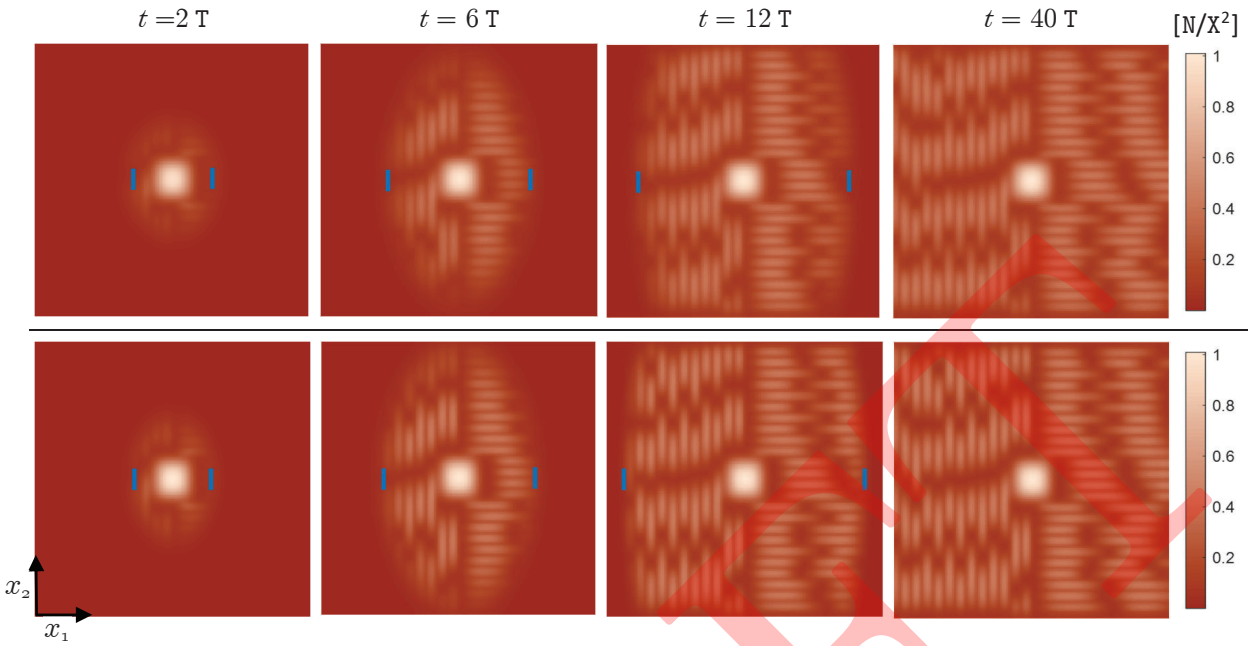


Figure 7: Range expansion of a species in a fragmented two-dimensional habitat. Here, m = 2, the environmental gradient along the x_2-axis is zero, and the environmental gradient along the x_1-axis is $\partial_{x_1} \mathrm{Q} = 1\,\mathtt{Q/X}$. Habitat fragmentation is simulated by considering a patchy profile for the carrying capacity K, as shown in the supplementary Figure S7. Parameter A takes different values for the results shown in the upper and lower panels. The rest of the parameters take their typical values given in Table 1. Four frames of the spatial profile of the species' population density are shown in each panel as the species' range evolves in time. The upper panel shows range expansion of a species with no optimal dispersal, $\mathrm{A} = 0\,\mathtt{X}^2\mathtt{/T}$, whereas the lower panel shows the range expansion of a species with strong optimal dispersal, $\mathrm{A} = 10\,\mathtt{X}^2\mathtt{/T}$. For comparison purposes, the same simulations but with constant carrying capacity of $\mathrm{K} = 1\ \mathtt{N/X}^2$ are performed and the leftmost and rightmost edges of the population are indicated by blue bars in each plot. It should also be noted that the final (appropriate steady-steady) profile of the phenotype-environment mismatch and trait variance at $t = 40\,\mathtt{T}$ are also shown in supplementary Figures S8 and S9.

the fragmented habitat almost with the same speed as it does it over a continuous habitat. This is implied by the locations of the blue bars in Figure 7, which also coincide with the leftmost and rightmost edges of the population in the fragmented habitat. Importantly, we do not observe any effects of phenotype-optimal dispersal that particularly enhance the range expansion capacity of the population when a habitat is fragmented, compared with a continuous habitat case.

When dispersal is only random, the particular layout of the patches—horizontally arranged on the right and vertically arranged on the left—does not result in any noticeable asymmetry in population's expansion to each side of the habitat. With phenotype-optimal dispersal, a relatively small asymmetry can be observed. The maximum steady-state ($t = 40\ \mathtt{T}$) population density at vertical patches is about 10 percent higher than the maximum density at horizontal patches.

Slightly better adaptation and lower trait variance is also observed within vertical patches. This can be explained by noting that the vertical patches in our simulation are stretched perpendicular to the gradient of the trait optimum. Therefore, optimal movements in the direction of the environmental gradient lead to a distribution of relatively close phenotypes along each of the vertical patches. Knowing that the patches are also relatively narrow, the overall closeness of phenotypes decreases the maladpative impacts of the random component of gene flow. As a result, population density can grow to higher levels within vertical patches.

Similar to our observations in the continuous habitat, a major impact of phenotype-optimal dispersal that we observe in the fragmented habitat is significant reduction in trait variation within the patches. Figure S9 shows the final profiles of trait variance at $t = 40$ T. In the presence of a steep environmental gradient, as we consider here, the reduced phenotypic variability within patches results in substantial increase in variability among the patches. Although not included in our model, it can be argued that this enhanced genetic differentiation between the patches contributes to reproductive isolation between the local populations. Large genetic differences between the patches can reduce the propensity of individuals to move between the patches, to avoid mating with individuals in a genetically novel population (Benkman, 2017; Bolnick and Otto, 2013; Cote et al., 2017; Garant et al., 2007). The reduced movements between the patches and reduced genetic variation within the patches can then promote the evolution of assortative mating, which further reinforces the reproductive isolation between the patches. This process can eventually lead to sympatric speciation in the habitat (Berner and Thibert-Plante, 2015; Bolnick and Otto, 2013; Edelaar and Bolnick, 2012; Garant et al., 2007; Lenormand, 2002; Nicolaus and Edelaar, 2018).

Discussion

Studies that identify and quantify the effects of adaptive dispersal strategies on the dynamics of spatially structured populations are still inadequate for developing a deep understanding of the evolution of these strategies. To contribute to such understanding, in the present work we developed and studied a deterministic mean-field model of adaptive range evolution of a population whose individuals can disperse optimally—by tracking the environmental gradient in a fitness-related phenotypic trait and moving to locations which minimize their phenotype-environment mismatch. In our model, the individuals' perceived force for phenotype-optimal dispersal depends both on their phenotype-environment mismatch and on the environmental gradient in trait optimum. Therefore, our model differs from non-optimal habitat choice mechanisms (Scheiner et al., 2022), in which a sufficiently strong phenotype-environment mismatch initiates dispersal of an individual to a new (randomly chosen) habitat patch, regardless of the presence or absence of an environmental gradient. The presence of an additional random-dispersal component in our model prevents the individuals from getting stuck in (small) non-matching habitat regions of zero environmental gradient.

We aimed to develop mechanistic insights into the effects of phenotype-optimal dispersal on joint evolution of the population density, trait mean, and trait variance. Thereby, we specifically aimed to make observations on the extent by which optimal dispersal can enhance range expansion

capacity and survival chance of a population. Importantly, through such observations we can identify conditions under which such enhancements are sufficiently strong to potentially evoke the evolution of the phenotype-optimal strategy.

Range Expansion

Our results confirm that sufficiently strong phenotype-optimal dispersal, or matching habitat choice in general, facilitates rapid adaptation of the individuals (and the population) within a single generation. In particular, the directed gene flow created by optimal dispersal effectively compensates for, or reverses, the swamping effects (migration load) of asymmetric core-to-edge random gene flow at population's range margins; see Figure 3. This enhances the population's range expansion capacity, both in continuous and in fragmented habitats. However, our results predict that the increase in range expansion speed is sufficiently significant only when the environmental gradient is steep; see Figure 2. The rapid adaptation caused by optimal dispersal also allows for faster growth in population density during transient states of population establishment. However, in the absence of environmental changes, we predict that the maximum density that the population can attain is not significantly changed by the individual's optimal dispersal strategy.

Genetic Swamping and Range Limit

Genetic swamping of peripheral populations by the asymmetric gene flow from central populations can destabilize the migration-selection equilibrium at range margins and limit the species' range. This has been one of the longstanding hypotheses proposed for explaining existence of stable range limits in the absence of dispersal limitations and gross environmental discontinuities (Haldane, 1956; Kirkpatrick and Barton, 1997; Lenormand, 2002; Mayr, 1963; Sexton et al., 2009). The seminal work of Kirkpatrick and Barton (1997) has established the theoretical foundation of this hypothesis. However, relaxing an unrealistic assumption on phenotypic variation (to be constant) in Kirkpatrick and Barton's model, the possibility of genetic swamping causing range limits has been challenged (Barton, 2001; Shirani and Miller, 2022). A review of empirical studies that tested the genetic swamping hypothesis has also identified little evidence for genetic swamping as a major cause of range limits (Kottler et al., 2021). In particular, Kottler et al. have found limited support for asymmetry in (total) center-to-edge gene flow, and very little evidence that such gene flow reduces mean fitness in edge populations. Their results suggest that gene flow often has neutral or positive effects on range edge populations.

In agreement with the observations made by Kottler et al., our results shown in Figures 3 and 4 predict that the total (net) gene flow to range margins has indeed positive (adaptive) effects on peripheral populations, even with moderate levels of optimal dispersal and at shallow environmental gradients. Therefore, our results confirm that genetic swamping is unlikely to be a major cause of species range limits, especially when some level of phenotype-optimal or phenotypically-directed dispersal is present in the population. Moreover, our results suggest phenotype-dependent habitat choice as a possible reason for the absence of significant asymmetry in center-to-edge gene flow in

the empirical studies reviewed by Kottler et al..

Environmental Changes and Species' Persistence

The significant enhancement of a species' ability to rapidly adapt to new environmental conditions is a key consequence of phenotype-optimal dispersal. It is mostly effective during transitions in the state of the population, for example, when the population is introduced to a new habitat, or when it responds to a temporal change in the environment. In particular, our results predict that phenotype-optimal dispersal can remarkably increase population density, range expansion speed, and chance of survival when changes in the environment occur sharply (within a fraction of the generation time of the species) and frequently (in every generation or couple of generations). The improvements are more substantial when natural selection is strong or fluctuations in the environmental trait optimum have large amplitudes. Slowly growing species, such as birds and mammals, can specifically benefit from such improvements. Without evolving rapid adaptation abilities, such species cannot recover sufficiently fast from their population density loss due to a sharp change in the environment. If changes occur frequently, they can eventually lead to species' extinction.

Phenotypic Variation, Competition, and Speciation

Phenotype-optimal dispersal has crucial impacts on the level of intraspecific phenotype variations, which can have consequences beyond adaption and range evolution of the species. Our results confirm that the spatial assortment of optimally dispersing individuals, based on the matching between their phenotype and the environment, substantially reduces within-population trait variance. This reduction in trait variation is a self-stabilizing (self-restricting) effect of phenotype-optimal dispersal. Stronger optimal dispersal propensity results in lower levels of phenotypic variation within the population. However, reduced variation implies lower average of phenotype-environment mismatch in the population, and hence less average phenotypic potential to disperse optimally. Furthermore, reduced phenotypic variation increases the level of competition between the individuals. This imposes a higher competition (settlement) cost on the evolution of optimal dispersal, and depresses the effects of competitive release caused by other ecological factors.

In metapopulations living in fragmented habitats, the reduced within-population trait variation and enhanced local adaptation resulted from optimal dispersal increase the level of trait variation between the local populations. Although not directly predicted by our model, the increased phenotypic divergence can drive reproductive isolation, reduce dispersal between the patches, promote assortative mating, and contribute to sympatric or parapatric speciation (Benkman, 2017; Berner and Thibert-Plante, 2015; Bolnick and Otto, 2013; Cote et al., 2017; Edelaar and Bolnick, 2012, 2019; Garant et al., 2007; Lenormand, 2002; Nicolaus and Edelaar, 2018). For example, phenotype-dependent native habitat preference has been identified as the factor promoting the sharp genetic divergence observed between parapatric populations of three-spined stickleback, which live in adjoining lake and stream habitats with significantly reduced dispersal between the habitats (Bolnick

et al., 2009; Edelaar and Bolnick, 2012).

Infrequency of Phenotype-optimal Dispersal

Currently available evidence for matching habitat choice strategies is fairly limited. The reason is partly because this mode of adaptation has been largely overlooked due to the major focus on natural selection (Edelaar and Bolnick, 2019). However, it can be argued that matching habitat choice is often insufficiently consequential to be detected, or is insufficiently beneficial to evolve—considering its potentially high cost of evolution. In fact, the evolution of matching habitat choice in a population can be inhibited by several factors, such as, high cost of the evolution of the dispersal trait, individuals' difficulty in obtaining accurate information about their performance and matching habitats, dependence of adaptation on multiple uncorrelated traits, and movement restrictions imposed by strong territoriality (Edelaar et al., 2017; Nicolaus and Edelaar, 2018). Moreover, studies on joint evolution of different adaptation modes using individual-based models, although (inevitably) not very realistic, have suggested that although matching habitat choice can have a greater adaptive potential than plasticity or natural selection (Nicolaus and Edelaar, 2018), it evolves less commonly than phenotypic plasticity (Edelaar et al., 2017), and is usually less favored than habitat construction (Scheiner et al., 2022).

Our results predict that a steep environmental gradient in trait optimum is necessary for phenotype-optimal dispersal to be sufficiently consequential on adaptation and range expansion of a species. Specifically, we observed strong effects for gradients greater than 1 $\mathtt{Q/X}$ in magnitude, that is, when trait optimum changes by more than one standard deviation per ($1/\sqrt{2}$ times) root mean square of (undirected) dispersal distance in one generation time. Although such a steep gradient can be plausible, it is estimated to be much steeper than the gradient that the majority of species in nature typically experience (Shirani and Miller, 2022). We should also note that we did not directly include any dispersal costs (except for indirect cost caused by increased competition) in our study. Inclusion of the dispersal costs is therefore expected to make optimal dispersal even less beneficial to evolve in non-steep environments. Additionally, our discussion above predict that phenotype-optimal dispersal should be more prevalent among slowly growing species which are exposed to sharp, strong, and frequent changes in their environment. Nomadic birds (Benkman, 2017) and long-range dispersing mammals such as walruses, seals, sea lions, cougars, jaguars, elephants, buffaloes, moose, rhinoceros, hippopotamuses, giraffes, and wolves can be examples of species that may benefit from evolving phenotype-optimal dispersal. Microorganisms that live in high chemical or temporal gradients may also optimally move to better matching environment through taxis.

Detection of the Phenotype-optimal Dispersal

Dispersal with matching habitat choice strategies is often hard to detect in nature. Local adaptation and genetic differentiation between populations are evolutionary effects of optimal dispersal that can also be caused by—and are often attributed to—strong natural selection. Local adaptation can also be effectively facilitated by phenotypic plasticity. It is usually not very straightforward

to evaluate which mode of adaptation has resulted in an observed pattern of local adaptation. Therefore, observational and experimental approaches for detecting phenotype-optimal dispersal should look for specific factors that can distinguish this mode of adaptation from the other modes. The key feature of phenotype-optimal dispersal, that is directed movement in the direction of the environmental gradient, is one of such important factors. The directed (non-random) gene flow created by such movements can be tested by several complementary techniques (Edelaar and Bolnick, 2012: Box 2). Our results particularly suggest two key considerations: (i) the impacts of gene flow on the fitness of individuals should specifically be measured at population's range margins, and (ii) the maximum level of trait variation should be measured at the core of the population. As we discussed before, phenotype-optimal dispersal results in adaptive gene flow to range margins, which is otherwise expected to be maladpative when dispersal is predominantly random. Moreover, phenotype-optimal dispersal can reduce trait variation to strikingly low levels that are unlikely to be maintained by realistically strong selection. We should also note that high range expansion speed and persistence capability under strongly changing environments, especially in slowly growing species with high mobility, can serve as preliminary indications for possible existence of optimal dispersal strategies.

Experimental approaches to test for the non-random gene flow directed by phenotype-optimal dispersal can be particularly insightful. In such approaches, the individuals' trait (performance) or the environmental characteristics that affect the optimum trait can be manipulated, and then possible directed dispersal can be tested as a response. Or, the individuals can be artificially displaced to known habitat conditions, and then their subsequent phenotype-dependent movement can be observed. (Edelaar and Bolnick, 2012). Nevertheless, we note that microcosms (Jacob et al., 2017) or microclimatic mosaic arenas (Karpestam et al., 2012) that are usually constructed in experiments to observe directed dispersal behaviors can artificially induce steep environmental gradients that are unlikely to be present in the natural habitat of the model species. Since our results identifies steep environmental gradient as a key factor in evolution of matching habitat choice strategies, a direct comparison of the environmental gradient between the experimental and natural habitats should be performed before making predictions on natural dispersal behaviors based on the experimentally observed behaviors.

Conclusion

The fact that an environmental gradient is necessary for habitat choice strategies to take place is trivial: without a gradient all habitat locations look the same to the individuals. However, our results particularly predict that the gradient must indeed be very steep to make phenotype-optimal dispersal sufficiently beneficial. This can serve as an explanation for why optimal dispersal is not very prevalent in nature. Yet, we should note that our estimates of typical values for steepness of environmental gradients are based on limited data. This is because environmental gradients are hard to measure, and the available measurements are often based on different choices of units which make steepness comparisons rather impossible. Our particular choices of units can provide

sufficient generality for future measurements, based on which better estimates for typical steepness of gradients can be obtained.

Phenotypic adaptation to a heterogeneous environment can occur through several modes, including habitat choice, selection, and phenotypic plasticity. How strong are each of these modes relative to the others, how they interact, and under what conditions they are more favored are questions that are not yet fully addressed (Edelaar and Bolnick, 2019; Edelaar et al., 2017; Scheiner et al., 2022). We included phenotype-optimal dispersal and selection in our model. We observed that strong optimal dispersal tends to reduce trait variation relatively fast, leaving less "fuel" for selection to operate. When a population experiences a sharp change in the environment, it suffers from a strong selective load. We observed that optimal dispersal is quite effective in mitigating the effects of this load and facilitating rapid recovery. This implies that, in rapidly changing environments, selection can indeed increase the potential for evolution of optimal dispersal. Detailed analysis of the interactions between these modes, possibly including other modes such as plasticity, can be an important direction of future research.

Our results confirm that phenotype-optimal dispersal can significantly contribute to genetic differentiation and speciation. Therefore—when sufficiently steep environmental gradient is present—we suggest that conservation efforts for protecting or promoting biodiversity plan for assisting the evolution of phenotype-optimal dispersal, possibly by constructing corridors or stepping stones aligned with the direction of the gradient.

Finally, our results emphasize the importance of intraspecific trait variation in adaptation, individuals interactions, and dispersal (Bolnick et al., 2011), and hence its necessity to be incorporated in models of range evolution. By doing so in our mode, we could specifically make observations on the interactions between the level of trait variation and the strength of phenotype-optimal dispersal, and confirm important consequences of optimal dispersal on population differentiation.

Appendix A: Perceived Environmental Gradient

The perceived gradient partially defined in (8) over a smaller habitat Ω_δ requires certain technical considerations near the boundary of the habitat so that it can be extended to the whole habitat Ω. To include such extension, we assume that the individuals of the species are able to perceive the boundary of the habitat once they become sufficiently close to it (closer than the constant δ), and they avoid crossing the boundary. That means, we assume that the normal component of $\widetilde{\nabla_x \mathrm{Q}}$ to the boundary of Ω is zero. Over a neighborhood of with δ around the boundary, we smoothly extend the $\widetilde{\nabla_x \mathrm{Q}}$ as defined by (8) so that its normal component to the boundary gradually vanishes to zero. As a result, the extended $\widetilde{\nabla_x \mathrm{Q}}$ will be tangential to the boundary of Ω. For this, we first define the following cut-off function

$$\chi_\delta(y) = \begin{cases} \exp\left(\frac{y^2}{y^2-\delta^2}\right), & |y| < \delta \\ 0, & |y| \geq \delta \end{cases} \tag{15}$$

which smoothly declines from 1 at $y = 0$ to 0 at $|y| = \delta$.

Now, for a rectangular habitat $\Omega = (a_1, b_1) \times \cdots \times (a_{\mathrm{m}}, b_{\mathrm{m}})$, we extend (8) as

$$\widetilde{\nabla_x \mathrm{Q}(x)} := \frac{\Pi}{\Pi + \|\nabla_x \mathrm{Q}(x)\|_{\mathbb{R}^{\mathrm{m}}}} \Bigg(\nabla_x \mathrm{Q}(x) - \sum_{i=1}^{\mathrm{m}} \chi_\delta(x_i - a_i) \partial^+_{x_i} \mathrm{Q}(x|_{a_i}) \hat{e}_i - \sum_{i=1}^{\mathrm{m}} \chi_\delta(x_i - b_i) \partial^-_{x_i} \mathrm{Q}(x|_{b_i}) \hat{e}_i \Bigg), \qquad x \in \Omega. \tag{16}$$

where $\hat{e}_i$ denotes the ith standard unit vector in $\mathbb{R}^{\mathrm{m}}$, and $x|_{c_i} := (x_1, \ldots, c_i, \ldots, x_{\mathrm{m}}) = x - (x_i - c_i)\hat{e}_i$. Moreover, $\partial^+_{x_i} \mathrm{Q}(z)$ and $\partial^-_{x_i} \mathrm{Q}(z)$ denote, respectively, the right-hand and left-hand partial derivatives of Q with respect to x_i evaluated at a point z. The summation terms in (16) gradually remove the normal component of $\nabla_x \mathrm{Q}$ to the boundaries $x_i = a_i$ and $x_i = b_i$, over a δ-neighborhood of the boundaries, so that $\widetilde{\nabla_x \mathrm{Q}(x)}$ eventually becomes completely tangent to the entire boundary of Ω. The removal of the normal components of the perceived gradient $\widetilde{\nabla_x \mathrm{Q}(x)}$ to habitat boundary automatically results in no phenotypic flux through the boundary due to the directed dispersal term (1b). Therefore, the no-flux (reflecting) boundary conditions that we discussed in Remark 1 and Appendix A.5 of our previous work (Shirani and Miller, 2022) can also be applied to the model (10)–(12) we present in this work. If a periodic boundary condition is considered across the jth spatial direction, then the terms associated with $i = j$ are excluded from the summation terms in (16). This is because imposing periodic boundary condition in a spatial direction is equivalent to considering a habitat that is periodically extended in that direction.

For the one-dimensional habitats $\Omega = (a, b)$ used in the results presented in Figures 1–6, the perceived environmental gradient (16) can be written as

$$\widetilde{\nabla_x \mathrm{Q}(x)} := \frac{\Pi}{\Pi + \|\nabla_x \mathrm{Q}(x)\|_{\mathbb{R}}} \Big(\nabla_x \mathrm{Q}(x) - \chi_\delta(x - a) \mathrm{d}^+_x \mathrm{Q}(a) - \chi_\delta(x - b) \mathrm{d}^-_x \mathrm{Q}(b) \Big), \qquad x \in \Omega. \tag{17}$$

where $\mathrm{d}^+_x \mathrm{Q}(z)$ and $\mathrm{d}^-_x \mathrm{Q}(z)$ denote, respectively, the right-hand and left-hand derivatives of Q with respect to x evaluated at a point z. Note that $\|\nabla_x \mathrm{Q}(x)\|_{\mathbb{R}} = |\mathrm{d}_x \mathrm{Q}(x)|$.

For the two-dimensional habitat $\Omega = (a_1, b_1) \times (a_2, b_2)$ used in the results presented in Figure 7, we considered reflecting boundary conditions at $x_1 = a_1$ and $x_1 = b_1$, and periodic boundary condition across the x_2-axis. Therefore, the perceived environmental gradient for this problem can be written as

$$\widetilde{\nabla_x \mathrm{Q}(x)} := \frac{\Pi}{\Pi + \|\nabla_x \mathrm{Q}(x)\|_{\mathbb{R}^2}} \Bigg(\nabla_x \mathrm{Q}(x) - \chi_\delta(x_1 - a_1) \begin{bmatrix} \partial^+_{x_1} \mathrm{Q}((a_1, x_2)) \\ 0 \end{bmatrix} - \chi_\delta(x_1 - b_1) \begin{bmatrix} \partial^-_{x_1} \mathrm{Q}((b_1, x_2)) \\ 0 \end{bmatrix} \Bigg), \qquad x \in \Omega. \tag{18}$$

In all of the simulations presented in this work, we set the maximum perceived gradient to be $\Pi = 1\,\mathtt{Q}/\mathtt{X}$, and we assume the individuals can sense the habitat boundary at a distance smaller than or equal to $\delta = 2\,\mathtt{X}$.

Appendix B: Model Derivation

We use the basic equation (1) to derive the equations of our model (10)–(12). As constructing components of the basic equation (1), we use the intrinsic growth rate (5), the perceived dispersal force (9), and the rate of mutational changes $\partial_t^{(\mathrm{M})}\phi$ given by Equation (20) in our previous work (Shirani and Miller, 2022). Our derivation relies on the major assumptions (i)–(viii) provided in the Model Assumptions section.

The basic equation (1) that we use to derive the equations of our model differs from the the Equation (25) in our previous work (Shirani and Miller, 2022) only in the inclusion of the optimal dispersal term (1b). As a result, applying the following derivation steps to the rest of the terms, (1a), (1c), and (1d), will give the terms (10a), (10c), (11a), (11b), (11e), (11f), and (12a)–(12c), whose derivation can be obtained as a single-species version of our general multi-species model presented in our previous work; see, for example, the Equation (12)–(14) given in that work for a one-dimensional habitat. Therefore, we exclude the derivation of these terms from this work and refer the reader to our previous work. In the following, we show the derivation of the new terms (10b), (11c)–(11e), (12d) and (12e) that model the effects of optimal dispersal.

We first substitute (9) into the basic equation (1) to write

$$n(t+\tau)\phi(t+\tau,p) - n(t)\phi(t,p) = -\tau \operatorname{div}\left(\mathrm{A}\, n(t)\phi(t,p)\frac{p-\mathrm{Q}}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right) + \ldots, \tag{19}$$

where "…" denotes the terms we have excluded from our derivation, as stated above. Note that, for simplicity of exposition, in writing (19) and the rest of the derivations that follow, we do not explicitly show the dependence of the variables and parameters on x.

Now, we integrate both sides of (19) with respect to p over $\mathbb{R}$ to obtain

$$n(t+\tau) - n(t) = -\tau \operatorname{div}\left(\mathrm{A}\, n(t)\frac{q(t)-\mathrm{Q}}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right) + \ldots. \tag{20}$$

Dividing both sides of (20) by τ and taking the limit as $\tau \to 0$ yields (10).

To derive (11), we multiply both sides of (19) by p and integrate the result with respect to p over $\mathbb{R}$. Noting that $\int_{\mathbb{R}} p^2\phi(t,p)\mathrm{d}p = v(t) + q^2(t)$, we obtain

$$n(t+\tau)q(t+\tau) - n(t)q(t) = -\tau \operatorname{div}\left(\mathrm{A}\, n(t)\frac{v(t)+q^2(t)-q(t)\mathrm{Q}}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right) + \ldots.$$

which, after dividing by τ and taking the limit as $\tau \to 0$, gives

$$\partial_t(n(t)q(t)) = -\operatorname{div}\left(\mathrm{A}\, n(t)\frac{v(t)+q^2(t)-q(t)\mathrm{Q}}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right) + \ldots. \tag{21}$$

Now, we use the chain rule on the left-hand side of (21) and substitute (10) into the result to obtain

$$\begin{aligned}\partial_t q(t) = \frac{1}{n(t)}\Bigg[&-\operatorname{div}\left(\mathrm{A}\, n(t)\frac{v(t)+q^2(t)-q(t)\mathrm{Q}}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right)\\ &+ q(t)\operatorname{div}\left(\mathrm{A}\, n(t)\frac{q(t)-\mathrm{Q}}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right)\Bigg] + \ldots.\end{aligned} \tag{22}$$

Denoting the first term within the brackets in (22) by [(22).1st], we can write

$$\begin{aligned}
[(22).1\text{st}] :=& -\operatorname{div}\left(\mathrm{A}\, n(t)\frac{v(t)+q^2(t)-q(t)\mathrm{Q}}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right)\\
=& -\operatorname{div}\left(\mathrm{A}\, n(t)\frac{q(t)(q(t)-\mathrm{Q})}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right)-\operatorname{div}\left(\mathrm{A}\, n(t)\frac{v(t)}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right)\\
=& -q(t)\operatorname{div}\left(\mathrm{A}\, n(t)\frac{q(t)-\mathrm{Q}}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right)-\left\langle \nabla_x q(t)\,,\, \mathrm{A}\, n(t)\frac{q(t)-\mathrm{Q}}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right\rangle_{\mathbb{R}^{\mathrm{m}}}\\
& -n(t)\operatorname{div}\left(\mathrm{A}\,\frac{v(t)}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right)-\left\langle \nabla_x n(t)\,,\, \mathrm{A}\,\frac{v(t)}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right\rangle_{\mathbb{R}^{\mathrm{m}}}.
\end{aligned}$$

Now, we substitute the result into (22) for the first term within the brackets and obtain (11).

Finally, to derive (12), we multiply both sides of (19) by $(p-q(t+\tau))^2$ and integrate the result with respect to p over $\mathbb{R}$. It yields

$$\begin{aligned}
n(t+\tau)v(t+\tau) =& \int_{\mathbb{R}}(p-q(t+\tau))^2\left[n(t)\phi(t,p)-\tau\operatorname{div}\left(\mathrm{A}\, n(t)\phi(t,p)\frac{p-\mathrm{Q}}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right)\right]\mathrm{d}p+\ldots\\
=& -\tau\int_{\mathbb{R}}p^2\operatorname{div}\left(\mathrm{A}\, n(t)\phi(t,p)\frac{p-\mathrm{Q}}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right)\mathrm{d}p+n(t)\int_{\mathbb{R}}p^2\phi(t,p)\mathrm{d}p\\
& -n(t+\tau)q^2(t+\tau)+\ldots\\
=& -\tau\operatorname{div}\left(\mathrm{A}\, n(t)\frac{3v(t)q(t)+q^3(t)-v(t)\mathrm{Q}-q^2(t)\mathrm{Q}}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right)+n(t)v(t)\\
& -\left(n(t+\tau)q^2(t+\tau)-n(t)q^2(t)\right)+\ldots.
\end{aligned}$$

Next, we divide both sides of the above equation by τ and take the limit as $\tau\to 0$. We obtain

$$\begin{aligned}
\partial_t(n(t)v(t)) =& -\operatorname{div}\left(\mathrm{A}\, n(t)\frac{3v(t)q(t)-v(t)\mathrm{Q}+q^2(t)(q(t)-\mathrm{Q})}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right)\\
& -\partial_t\left(n(t)q^2(t)\right)+\ldots. \qquad (23)
\end{aligned}$$

If we use the chain rule to write $\partial_t(n(t)v(t)) = n(t)\partial_t v(t)+v(t)\partial_t n(t)$, wherein $\partial_t n(t)$ is given by (10), and split the divergence term in (23) into three terms, we can write

$$\begin{aligned}
n(t)\partial_t v(t) =& v(t)\operatorname{div}\left(\mathrm{A}\, n(t)\frac{q(t)-\mathrm{Q}}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right) && (24a)\\
& -\operatorname{div}\left(\mathrm{A}\, n(t)\frac{2v(t)q(t)}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right) && (24b)\\
& -\operatorname{div}\left(\mathrm{A}\, n(t)\frac{v(t)(q(t)-\mathrm{Q})}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right) && (24c)\\
& -\operatorname{div}\left(\mathrm{A}\, n(t)\frac{q^2(t)(q(t)-\mathrm{Q})}{\mathrm{V}}\widetilde{\nabla_x \mathrm{Q}}\right) && (24d)\\
& -\partial_t\left(n(t)q^2(t)\right)+\ldots. && (24e)
\end{aligned}$$

Now, for the terms (24b), (24d), and (24e), we have

$$\begin{aligned}(24b) = &- 2\, n(t)q(t)\,\mathrm{div}\left(\mathrm{A}\,\frac{v(t)}{\mathrm{V}}\,\widetilde{\nabla_x \mathrm{Q}}\right) - 2\left\langle q(t)\nabla_x n(t)\,,\,\mathrm{A}\,\frac{v(t)}{\mathrm{V}}\,\widetilde{\nabla_x \mathrm{Q}}\right\rangle_{\mathbb{R}^{\mathrm{m}}} \\ &- 2\left\langle n(t)\nabla_x q(t)\,,\,\mathrm{A}\,\frac{v(t)}{\mathrm{V}}\,\widetilde{\nabla_x \mathrm{Q}}\right\rangle_{\mathbb{R}^{\mathrm{m}}}, \end{aligned} \tag{25}$$

$$(24d) = - q^2(t)\,\mathrm{div}\left(\mathrm{A}\,n(t)\frac{q(t)-\mathrm{Q}}{\mathrm{V}}\,\widetilde{\nabla_x \mathrm{Q}}\right) - 2\left\langle q(t)\nabla_x q(t)\,,\,\mathrm{A}\,n(t)\frac{q(t)-\mathrm{Q}}{\mathrm{V}}\,\widetilde{\nabla_x \mathrm{Q}}\right\rangle_{\mathbb{R}^{\mathrm{m}}}, \tag{26}$$

$$\begin{aligned}(24e) = &\, q^2(t)\,\mathrm{div}\left(\mathrm{A}\,n(t)\frac{q(t)-\mathrm{Q}}{\mathrm{V}}\,\widetilde{\nabla_x \mathrm{Q}}\right) + 2\, n(t)q(t)\,\mathrm{div}\left(\mathrm{A}\,\frac{v(t)}{\mathrm{V}}\,\widetilde{\nabla_x \mathrm{Q}}\right) \\ &+ 2\left\langle q(t)\nabla_x n(t)\,,\,\mathrm{A}\,\frac{v(t)}{\mathrm{V}}\,\widetilde{\nabla_x \mathrm{Q}}\right\rangle_{\mathbb{R}^{\mathrm{m}}} + 2\left\langle q(t)\nabla_x q(t)\,,\,\mathrm{A}\,n(t)\frac{q(t)-\mathrm{Q}}{\mathrm{V}}\,\widetilde{\nabla_x \mathrm{Q}}\right\rangle_{\mathbb{R}^{\mathrm{m}}}. \end{aligned} \tag{27}$$

Moreover, denoting the right-hand side of (24a) by (24a).RHS, we can write

$$(24a).\mathrm{RHS} = \mathrm{div}\left(\mathrm{A}\,n(t)v(t)\frac{q(t)-\mathrm{Q}}{\mathrm{V}}\,\widetilde{\nabla_x \mathrm{Q}}\right) - \left\langle \nabla_x v(t)\,,\,\mathrm{A}\,n(t)\frac{q(t)-\mathrm{Q}}{\mathrm{V}}\,\widetilde{\nabla_x \mathrm{Q}}\right\rangle_{\mathbb{R}^{\mathrm{m}}}. \tag{28}$$

Substituting (25), (26), (27), and (28) into (24) and simplifying the result gives (12).

References

A. L. Angert, M. G. Bontrager, and J. Ågren. What do we really know about adaptation at range edges? *Annual Review of Ecology, Evolution, and Systematics*, 51(1):341–361, 2020. doi: 10.1146/annurev-ecolsys-012120-091002.

P. R. Armsworth. Conditional dispersal, clines, and the evolution of dispersiveness. *Theoretical Ecology*, 2(2):105–117, 2009. doi: 10.1007/s12080-008-0032-2.

P. R. Armsworth and J. E. Roughgarden. The impact of directed versus random movement on population dynamics and biodiversity patterns. *The American Naturalist*, 165(4):449–465, 2005. doi: 10.1086/428595.

P. R. Armsworth and J. E. Roughgarden. The structure of clines with fitness-dependent dispersal. *The American Naturalist*, 172(5):648–657, 2008. doi: 10.1086/591685.

M. Baguette, S. Blanchet, D. Legrand, V. M. Stevens, and C. Turlure. Individual dispersal, landscape connectivity and ecological networks. *Biological Reviews*, 88(2):310–326, 2013. doi: doi.org/10.1111/brv.12000.

C. B. Baines, I. M. C. Ferzoco, and S. J. McCauley. Phenotype-by-environment interactions influence dispersal. *Journal of Animal Ecology*, 88(8):1263–1274, 2019. doi: 10.1111/1365-2656.13008.

N. Barton. Adaptation at the edge of a species range. In J. Silvertown and J. Antonovics, editors, *Integrating ecology and evolution in a spatial context*, chapter 17, page 365–392. Blackwell, Oxford, 2001.

C. W. Benkman. Matching habitat choice in nomadic crossbills appears most pronounced when food is most limiting. *Evolution*, 71(3):778–785, 2017. doi: 10.1111/evo.13146.

D. Berner and X. Thibert-Plante. How mechanisms of habitat preference evolve and promote divergence with gene flow. *Journal of Evolutionary Biology*, 28(9):1641–1655, 2015. doi: 10.1111/jeb.12683.

E. Bestion, J. Clobert, and J. Cote. Dispersal response to climate change: scaling down to intraspecific variation. *Ecology Letters*, 18(11):1226–1233, 2015. doi: doi.org/10.1111/ele.12502.

D. I. Bolnick and S. P. Otto. The magnitude of local adaptation under genotype-dependent dispersal. *Ecology and Evolution*, 3(14):4722–4735, 2013. doi: https://doi.org/10.1002/ece3.850.

D. I. Bolnick, L. K. Snowberg, C. Patenia, W. E. Stutz, T. Ingram, and O. L. Lau. Phenotype-dependent native habitat preference facilitates divergence between parapatric lake and stream stickleback. *Evolution*, 63(8):2004–2016, 2009. doi: 10.1111/j.1558-5646.2009.00699.x.

D. I. Bolnick, P. Amarasekare, M. S. Araújo, R. Bürger, J. M. Levine, M. Novak, V. H. Rudolf, S. J. Schreiber, M. C. Urban, and D. A. Vasseur. Why intraspecific trait variation matters in community ecology. *Trends in Ecology & Evolution*, 26(4):183–192, 2011. doi: https://doi.org/10.1016/j.tree.2011.01.009.

D. Bonte, H. Van Dyck, J. M. Bullock, A. Coulon, M. Delgado, M. Gibbs, V. Lehouck, E. Matthysen, K. Mustin, M. Saastamoinen, N. Schtickzelle, V. M. Stevens, S. Vandewoestijne, M. Baguette, K. Barton, T. G. Benton, A. Chaput-Bardy, J. Clobert, C. Dytham, T. Hovestadt, C. M. Meier, S. C. F. Palmer, C. Turlure, and J. M. J. Travis. Costs of dispersal. *Biological Reviews*, 87(2):290–312, 2012. doi: 10.1111/j.1469-185X.2011.00201.x.

D. E. Bowler and T. G. Benton. Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biological Reviews*, 80(2):205–225, 2005. doi: 10.1017/S1464793104006645.

J. R. Bridle and T. H. Vines. Limits to evolution at range margins: when and why does adaptation fail? *Trends in Ecology & Evolution*, 22(3):140–147, 2007. doi: 10.1016/j.tree.2006.11.002.

T. J. Case and M. L. Taper. Interspecific competition, environmental gradients, gene flow, and the coevolution of species' borders. *The American Naturalist*, 155(5):583–605, 2000. doi: 10.1086/303351.

T. J. Case, R. D. Holt, M. A. McPeek, and T. H. Keitt. The community context of species' borders: ecological and evolutionary perspectives. *Oikos*, 108(1):28–46, 2005. doi: 10.1111/j.0030-1299.2005.13148.x.

J. Clobert, J.-F. Le Galliard, J. Cote, S. Meylan, and M. Massot. Informed dispersal, heterogeneity in animal dispersal syndromes and the dynamics of spatially structured populations. *Ecology Letters*, 12(3):197–209, 2009. doi: 10.1111/j.1461-0248.2008.01267.x.

J. Cote, E. Bestion, S. Jacob, J. Travis, D. Legrand, and M. Baguette. Evolution of dispersal strategies and dispersal syndromes in fragmented landscapes. *Ecography*, 40(1):56–73, 2017. doi: https://doi.org/10.1111/ecog.02538.

P. W. Dillingham, J. E. Moore, D. Fletcher, E. Cortés, K. A. Curtis, K. C. James, and R. L. Lewison. Improved estimation of intrinsic growth rmax for long-lived species: integrating matrix models and allometry. *Ecological Applications*, 26(1):322–333, 2016. doi: 10.1890/14-1990.

A. N. Dreiss, S. Antoniazza, R. Burri, L. Fumagalli, C. Sonnay, C. Frey, J. Goudet, and A. Roulin. Local adaptation and matching habitat choice in female barn owls with respect to melanic coloration. *Journal of Evolutionary Biology*, 25(1):103–114, 2012. doi: 10.1111/j.1420-9101.2011.02407.x.

R. A. Duckworth and A. V. Badyaev. Coupling of dispersal and aggression facilitates the rapid range expansion of a passerine bird. *Proceedings of the National Academy of Sciences*, 104(38):15017–15022, 2007. doi: 10.1073/pnas.0706174104.

P. Edelaar and D. I. Bolnick. Non-random gene flow: an underappreciated force in evolution and ecology. *Trends in Ecology & Evolution*, 27(12):659–665, 2012. doi: https://doi.org/10.1016/j.tree.2012.07.009.

P. Edelaar and D. I. Bolnick. Appreciating the multiple processes increasing individual or population fitness. *Trends in Ecology & Evolution*, 34(5):435–446, 2019. doi: 10.1016/j.tree.2019.02.001.

P. Edelaar, A. M. Siepielski, and J. Clobert. Matching habitat choice causes directed gene flow: A neglected dimension in evolution and ecology. *Evolution*, 62(10):2462–2472, 2008. doi: https://doi.org/10.1111/j.1558-5646.2008.00459.x.

P. Edelaar, R. Jovani, and I. Gomez-Mestre. Should I change or should I go? Phenotypic plasticity and matching habitat choice in the adaptation to environmental heterogeneity. *The American Naturalist*, 190(4):506–520, 2017. doi: 10.1086/693345.

E. A. Fronhofer and F. Altermatt. Eco-evolutionary feedbacks during experimental range expansions. *Nature Communications*, 6(1):6844, 2015. doi: 10.1038/ncomms7844.

D. Garant, S. E. Forde, and A. P. Hendry. The multifarious effects of dispersal and gene flow on contemporary adaptation. *Functional Ecology*, 21(3):434–443, 2007. doi: 10.1111/j.1365-2435.2006.01228.x.

K. J. Gaston et al. *The structure and dynamics of geographic ranges*. Oxford University Press, Oxford, 2003.

W. Godsoe, N. J. Holland, C. Cosner, B. E. Kendall, A. Brett, J. Jankowski, and R. D. Holt. Interspecific interactions and range limits: contrasts among interaction types. *Theoretical Ecology*, 10(2):167–179, 2017. doi: 10.1007/s12080-016-0319-7.

N. M. Haddad, L. A. Brudvig, J. Clobert, K. F. Davies, A. Gonzalez, R. D. Holt, T. E. Lovejoy, J. O. Sexton, M. P. Austin, C. D. Collins, W. M. Cook, E. I. Damschen, R. M. Ewers, B. L. Foster, C. N. Jenkins, A. J. King, W. F. Laurance, D. J. Levey, C. R. Margules, B. A. Melbourne, A. O. Nicholls, J. L. Orrock, D.-X. Song, and J. R. Townshend. Habitat fragmentation and its lasting impact on earth's ecosystems. *Science Advances*, 1(2):e1500052, 2015. doi: 10.1126/sciadv.1500052.

J. B. S. Haldane. The relation between density regulation and natural selection. *Proceedings of the Royal Society of London. Series B - Biological Sciences*, 145(920):306–308, 1956. doi: 10.1098/rspb.1956.0039.

R. D. Holt and T. H. Keitt. Species' borders: a unifying theme in ecology. *Oikos*, 108(1):3–6, 2005. doi: 10.1111/j.0030-1299.2005.13145.x.

S. Jacob, D. Legrand, A. S. Chaine, D. Bonte, N. Schtickzelle, M. Huet, and J. Clobert. Gene flow favours local adaptation under habitat choice in ciliate microcosms. *Nature Ecology & Evolution*, 1(9):1407–1410, 2017. doi: 10.1038/s41559-017-0269-5.

E. Karpestam, L. Wennersten, and A. Forsman. Matching habitat choice by experimentally mismatched phenotypes. *Evolutionary Ecology*, 26(4):893–907, 2012. doi: 10.1007/s10682-011-9530-6.

J. C. Kingsolver, H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. E. Hill, A. Hoang, P. Gibert, and P. Beerli. The strength of phenotypic selection in natural populations. *The American Naturalist*, 157(3):245–261, 2001. doi: 10.1086/319193.

M. Kirkpatrick and N. H. Barton. Evolution of a species' range. *The American Naturalist*, 150(1):1–23, 1997. doi: 10.1086/286054.

E. J. Kottler, E. E. Dickman, J. P. Sexton, N. C. Emery, and S. J. Franks. Draining the swamping hypothesis: Little evidence that gene flow reduces fitness at range edges. *Trends in Ecology & Evolution*, 36(6):533–544, 2021. doi: https://doi.org/10.1016/j.tree.2021.02.004.

T. Lenormand. Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, 17(4):183–189, 2002. doi: https://doi.org/10.1016/S0169-5347(02)02497-7.

A. M. Louthan, D. F. Doak, and A. L. Angert. Where and when do species interactions set range limits? *Trends in Ecology & Evolution*, 30(12):780–792, 2015. doi: 10.1016/j.tree.2015.09.011.

N. Lustenhouwer, F. Moerman, F. Altermatt, R. D. Bassar, G. Bocedi, D. Bonte, S. Dey, E. A. Fronhofer, E. G. da Rocha, A. Giometto, L. T. Lancaster, R. B. Prather Jr, M. Saastamoinen, J. M. J. Travis, C. A. Urquhart, C. Weiss-Lehman, J. L. Williams, L. Börger, and D. Berger. Experimental evolution of dispersal: Unifying theory, experiments and natural systems. *Journal of Animal Ecology*, 92(6):1113–1123, 2023. doi: https://doi.org/10.1111/1365-2656.13930.

E. Mayr. *Animal Species and Evolution*. Harvard University Press, Cambridge, MA, 1963.

T. E. X. Miller, A. L. Angert, C. D. Brown, J. A. Lee-Yaw, M. Lewis, F. Lutscher, N. G. Marculis, B. A. Melbourne, A. K. Shaw, M. Szűcs, O. Tabares, T. Usui, C. Weiss-Lehman, and J. L. Williams. Eco-evolutionary dynamics of range expansion. *Ecology*, 101(10):e03139, 2020. doi: 10.1002/ecy.3139.

M. Nicolaus and P. Edelaar. Comparing the consequences of natural selection, adaptive phenotypic plasticity, and matching habitat choice for phenotype-environment matching, population genetic structure, and reproductive isolation in meta-populations. *Ecology and Evolution*, 8(8):3815–3827, 2018. doi: 10.1002/ece3.3816.

C. Niel and J. Lebreton. Using demographic invariants to detect overharvested bird populations from incomplete data. *Conservation Biology*, 19(3):826–835, 2005. doi: 10.1111/j.1523-1739.2005.00310.x.

A. Ponchon and J. M. J. Travis. Informed dispersal based on prospecting impacts the rate and shape of range expansions. *Ecography*, 2022(5):e06190, 2022. doi: https://doi.org/10.1111/ecog.06190.

M. Rafajlović, J. M. Alexander, R. K. Butlin, and K. Johannesson. Introduction to the theme issue 'species' ranges in the face of changing environments'. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1848):20210002, 2022. doi: 10.1098/rstb.2021.0002.

V. Ravigné, U. Dieckmann, and I. Olivieri. Live where you thrive: Joint evolution of habitat choice and local adaptation facilitates specialization and promotes diversity. *The American Naturalist*, 174(4):E141–E169, 2009. doi: 10.1086/605369.

O. Ronce. How does it feel to be like a rolling stone? ten questions about dispersal evolution. *Annual Review of Ecology, Evolution, and Systematics*, 38(1):231–253, 2007. doi: 10.1146/annurev.ecolsys.38.091206.095611.

G. D. Ruxton and P. Rohani. Fitness-dependent dispersal in metapopulations and its consequences for persistence and synchrony. *Journal of Animal Ecology*, 68(3):530–539, 1999. doi: 10.1046/j.1365-2656.1999.00300.x.

M. Saastamoinen, G. Bocedi, J. Cote, D. Legrand, F. Guillaume, C. W. Wheat, E. A. Fronhofer, C. Garcia, R. Henry, A. Husby, M. Baguette, D. Bonte, A. Coulon, H. Kokko, E. Matthysen, K. Niitepõld, E. Nonaka, V. M. Stevens, J. M. J. Travis, K. Donohue, J. M. Bullock, and M. del Mar Delgado. Genetics of dispersal. *Biological Reviews*, 93(1):574–599, 2018. doi: 10.1111/brv.12356.

S. M. Scheiner, M. Barfield, and R. D. Holt. Do i build or do i move? Adaptation by habitat construction versus habitat choice. *Evolution*, 76(3):414–428, 2022. doi: 10.1111/evo.14355.

V. Selonen and I. K. Hanski. Habitat exploration and use in dispersing juvenile flying squirrels. *Journal of Animal Ecology*, 75(6):1440–1449, 2006. doi: 10.1111/j.1365-2656.2006.01168.x.

J. P. Sexton, P. J. McIntyre, A. L. Angert, and K. J. Rice. Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution, and Systematics*, 40(1):415–436, 2009. doi: 10.1146/annurev.ecolsys.110308.120317.

F. Shirani and J. R. Miller. Competition, trait variance dynamics, and the evolution of a species' range. *Bulletin of Mathematical Biology*, 84(3):37, 2022. doi: 10.1007/s11538-022-00990-z.

Y. Takahashi, Y. Suyama, Y. Matsuki, R. Funayama, K. Nakayama, and M. Kawata. Lack of genetic variation prevents adaptation at the geographic range margin in a damselfly. *Molecular Ecology*, 25(18):4450–4460, 2016. doi: 10.1111/mec.13782.