Evolution of dispersal in a spatially heterogeneous population with finite patch sizes and catastrophes

Kalle Parvinen, Hisashi Ohtsuki, Joe Yuichiro Wakano

PII: S0022-5193(26)00153-0
DOI: https://doi.org/10.1016/j.jtbi.2026.112528
Reference: YJTBI 112528

To appear in: *Journal of Theoretical Biology*

Received date: 14 November 2025
Revised date: 1 June 2026
Accepted date: 10 June 2026

Please cite this article as: Kalle Parvinen, Hisashi Ohtsuki, Joe Yuichiro Wakano, Evolution of dispersal in a spatially heterogeneous population with finite patch sizes and catastrophes, *Journal of Theoretical Biology* (2025), doi: https://doi.org/10.1016/j.jtbi.2026.112528

Highlights:

- We study the evolution of dispersal in a patch-structured population.
- Habitat patches may encounter natural disasters such as flooding or wildfire
- Such catastrophes, even rarely occurring, have major impacts on dispersal evolution
- Increasing the probability to survive dispersal may select against dispersal
- This counterintuitive phenomenon occurs due to interplay with kin selection

Evolution of dispersal in a spatially heterogeneous population with finite patch sizes and catastrophes

Kalle Parvinen[1,2,3], Hisashi Ohtsuki[4] and Joe Yuichiro Wakano[5,6]

June 17, 2026

1. Department of Mathematics and Statistics, FI-20014 University of Turku, Finland. E-mail: kalle.parvinen@utu.fi
2. Advancing Systems Analysis Program, International Institute for Applied Systems Analysis (IIASA), A-2361 Laxenburg, Austria
3. Okinawa Institute of Science and Technology, Onna-son, Kunigami-gun, Okinawa, 904-0495, Japan
4. Research Center for Integrative Evolutionary Science, SOKENDAI (The Graduate University for Advanced Studies), Shonan Village, Hayama, Kanagawa 240-0193, Japan
5. School of Interdisciplinary Mathematical Sciences, Meiji University, Tokyo 164-8525, Japan
6. Meiji Institute for Advanced Study of Mathematical Sciences, Tokyo 164-8525, Japan

Abstract

Evolution of costly dispersal has been one of the main topics in evolutionary biology. Theoreticians have tried to study realistic models including properties such as 1) finite patch sizes, 2) heterogeneity in patch qualities, 3) conditional dispersal strategy based on the patch quality, and 4) catastrophes (i.e, local population extinctions in habitat patches). Finite patch sizes lead to kin competition within each patch, promoting dispersal. Heterogeneity is expected to promote conditional dispersal from low-quality patches. Catastrophes cause temporal heterogeneity, which often promotes dispersal. Furthermore, catastrophes will cause non-dispersing organisms to go extinct. Here we study a model including all the four properties listed above. In parallel with previous studies, we found that the introduction of catastrophes generally promotes dispersal. We found, however, unexpected interactions between catastrophes and other demographic factors. Without catastrophes, increasing the dispersal survival probability increases dispersal, as expected.

In the presence of catastrophes, dispersal can be non-monotonic with respect to the dispersal survival probability. Furthermore, when the dispersal survival probability is very small, including catastrophes in the model will have an abrupt positive effect on dispersal. Our findings strongly suggest that even a tiny chance of catastrophic events happening can qualitatively alter the course of evolution of dispersal. Our results might be applicable when small patches are well-isolated and local natural disasters happen rarely but do occur.

1 Introduction

Dispersal is one of the key aspects of life-history strategies of organisms (Van Valen, 1971; Hamilton and May, 1977; Comins et al., 1980; Dieckmann et al., 1999; Clobert et al., 2001; Ronce, 2007). A metapopulation (Levins, 1969, 1970; Hanski, 1999) is a collection of local populations living in habitat patches connected by dispersal. When considering potential fitness advantages that dispersing individuals can enjoy, one typically first thinks of the direct benefit of finding better habitats. For example, a dispersing individual may find a habitat patch with more resources or less competition (Kisdi, 2002), but on the other hand, the new patch may be less advantageous than the original one, and it is possible that a disperser does not find a patch at all. Although spatial heterogeneity could promote dispersal, under equilibrium dynamics it has been found to select against dispersal (Hastings, 1983; Parvinen, 1999; Gyllenberg et al., 2002; Parvinen, 2006; Parvinen et al., 2020). In contrast, temporal heterogeneity has been found to promote dispersal (McPeek and Holt, 1992; Holt and McPeek, 1996; Doebeli and Ruxton, 1997; Gandon and Michalakis, 1999; Johst et al., 1999; Parvinen, 1999; Mathias et al., 2001; Gyllenberg et al., 2002; Parvinen, 2006). Dispersal can be costly also because dispersers may need physical traits, such as wings or propellers, or require extra energy for movement. Furthermore, dispersers may risk their lives due to, e.g., predation.

An important reference point in the literature is the classical paper by Hamilton and May (1977). In their first model, the living conditions in all patches are the same, so that there are no direct benefits of dispersal, but positive dispersal evolves nevertheless. This model consists of infinitely many patches occupied by one adult individual. Consequently, local kin competition, that is the competition over vacant spots in a local patch among kin, is a driving force for dispersal evolution (Motro, 1982a,b, 1983; Frank, 1986; Taylor, 1988; Ajar, 2003; Massol et al., 2011). For this case the evolutionarily stable dispersal probability is $1/(2-p)$, where p is the probability of surviving dispersal. In their model, increasing dispersal cost always disfavors dispersal. The same phenomenon occurs in an extended model, consisting of a mixture of patches with different size and fecundity (Parvinen et al., 2020). Furthermore, such spatial heterogeneity disfavours dispersal.

Among those ecological factors that can impact the evolution of dispersal traits, the effects of catastrophic events have been discussed (Ronce et al., 2000; Gyllenberg et al., 2002; Parvinen et al., 2003; Parvinen, 2006; Jansen and Vitalis, 2007).

Here, a catastrophe refers to a major disturbance in habitats that causes a loss of all (or a large part of) inhabitants there. Examples include fires, floods, storms, earthquakes and appearance of a predator. Once a catastrophic event, even if their occurrence is extremely rare, happens, organisms are subject to a large negative impact, and therefore an evolutionary impact of those events on dispersal evolution is not at all negligible. Catastrophes also result in empty or sparsely populated patches, into which it may be beneficial to disperse. In fact, previous theoretical works have shown that a moderate level of catastrophes favors dispersal (Ronce et al., 2000; Gyllenberg et al., 2002; Parvinen et al., 2003; Parvinen, 2006). This makes sense because dispersal works for hedging the risk of the mass extinction of one's own lineage.

Inspired by those previous works, we herein study the effect of catastrophes on dispersal evolution, and its interplay with dispersal cost and heterogeneity in the population. The suggested model here is a straightforward extension of Parvinen et al. (2020) by incorporating a positive chance μ of a catastrophe. As in Parvinen et al. (2020), we study the evolution of unconditional dispersal, which means that the probability of dispersal is independent of the current patch type. In addition, we study conditional dispersal as well, which means that the dispersal probability can depend on the current patch type (Weigang, 2017; Parvinen et al., 2023). As a result, we find that the introduction of catastrophes generally promotes dispersal, which is consistent with previous findings. However, we have at the same time found unexpected interactions between catastrophes and other demographic factors. In particular, we have found that in the presence of catastrophes an evolved dispersal probability does not necessarily monotonically increase with the probability p of surviving dispersal (Comins et al., 1980). This means that improved survival during dispersal can sometimes hinder, rather than promote, evolution of dispersal. Moreover, we have found that when the probability of catastrophes is small and the probability of surviving dispersal is low, heterogeneity in the population can sometimes favor, rather than disfavor, evolution of dispersal, which is a contrasting result to our previous report (Parvinen et al., 2020). These results strongly suggest that even a tiny chance that catastrophic events occur, can qualitatively alter the conclusions drawn by previous works.

The remainder of the paper is structured as follows. We will first describe our mathematical models. Then we separately analyze the evolution of unconditional and conditional dispersal rates. Comparisons between these two scenarios are also performed. The discussion concludes the paper.

2 Model and methods

2.1 A spatially heterogeneous model with finite patch sizes and catastrophes

The Wright's island model (Wright, 1931) consists of infinitely many ecologically equal habitat patches (demes) (Ajar, 2003). In previous works (Parvinen et al., 2018, 2020) we included the possibility for spatial heterogeneity, so that patches can be of N different types. Here we, in addition, include the possibility of catastrophes. The details and seasonal life-cycle of the model are as follows:

The proportion of patches of type k is π_k, adding up to one, $\sum_{k=1}^{N} \pi_k = 1$. In the beginning of the season each patch of type k contains n_k adult individuals. In each patch, a local catastrophe occurs with probability μ. Such a catastrophe is assumed to kill all the individuals in a patch, but the patch remains habitable, and can be re-colonized by dispersers. Reproduction occurs in non-empty patches, i.e., where a catastrophe has not occurred. Fecundity in patches of type k is γF_k that represents the number of juveniles that each adult produces. We assume γ to be very large ($\gamma \to \infty$). This means that each adult produces a very large amount of offspring, which is typical for plants producing a large amount of seeds, but differences in fecundities between patches can be taken into account.

As we investigate the evolution of (conditional) dispersal, we allow individuals to differ in their dispersal probabilities m_k, $0 < m_k \leqslant 1$, which is the probability of a juvenile to emigrate from a patch of type k. The proportion $0 < p \leqslant 1$ of dispersed juveniles will survive dispersal and land independently in a random patch. The probability to arrive in a patch of type k is assumed to be ϕ_k, and $\sum_{k=1}^{N} \phi_k = 1$. The relation $\lambda_k = \phi_k/\pi_k$ describes how strongly patches of type k attract immigrants. In the standard case $\lambda_k = 1$, so that $\phi_k = \pi_k$ and all patches receive the same amount of immigrants. The present adults are assumed not to survive until the next season. After immigration there are always plenty of juveniles present, even in such patches where a catastrophe happened, because of the assumption $\gamma \to \infty$. For this reason, the patch size can be assumed to be fixed, so that the n_k individuals to become adults in a patch of type k in the next season are randomly chosen among the juveniles in each patch after immigration.

2.2 Metapopulation fitness and adaptive dynamics

Consider a situation, in which (almost) all individuals have the same dispersal probability m (for conditional dispersal, it is a vector of dispersal probabilities, $\boldsymbol{m}$). Then we investigate, whether a rare mutant with dispersal probability m^{mut} (vector $\boldsymbol{m}^{\mathrm{mut}}$ for the conditional dispersal case) is able to grow in population size. For this purpose, we calculate the metapopulation reproduction ratio (metapopulation fitness) $R_{\mathrm{m}}(m^{\mathrm{mut}}, m)$ (Metz and Gyllenberg, 2001; Gyllenberg and Metz, 2001). It is the expected number of dispersers produced by a colony founded by an immigrating mutant. If $R_{\mathrm{m}}(m^{\mathrm{mut}}, m) > 1$, the mutant population may grow and potentially invade the resident population. Details of how to calculate the metapopulation fitness $R_{\mathrm{m}}(m^{\mathrm{mut}}, m)$ for this model are explained in SI section A. With the help of the implicit function theorem, we are able to obtain an explicit expression for the fitness gradient (selection gradient) (See SI section B.1), generalizing the result of Parvinen et al. (2020) to cover the model with catastrophes. The fitness gradient $D_1(m)$ is the first-order derivative of (metapopulation) fitness with respect to the strategy m^{mut} of the mutant, evaluated when $m^{\mathrm{mut}} = m$. It provides the direction into which the strategy of the monomorphic population is

First-hand model features		
variable	meaning	note
N	number of different patch types	
k	type of a patch	$k = 1, 2, ..., N$
n_k	number of adults in a patch of type k	
F_k	(relative) fecundity in a patch of type k	
π_k	proportion of patches of type k	
λ_k	attractiveness of a patch of type k	
ϕ_k	probability to arrive in a patch of type k	$\phi_k = \lambda_k \pi_k$
m_k	dispersal probability from a patch of type k	evolving trait
p	survival probability of dispersing juveniles	
μ	probability of catastrophe	
Compound expressions		
variable	meaning	note
f	$f = F_1/F_2$	used when $N = 2$
d_k^{cond}	conditional backward migration probability	$\frac{mp(1-\mu)}{(1-m)V_k+mp(1-\mu)}$
d_k	backward migration probability	$d_k = (1-\mu)d_k^{\text{cond}} + \mu$
$R_{2,k}^{\text{R}}$	relatedness with replacement at a patch of type k	$R_{2,k}^{\text{R}} = \frac{1}{n_k-(n_k-1)(1-\mu)(1-d_k^{\text{cond}})^2}$
V_k	relative reproductive potential of a patch of type k	Eq. 3.2
$H_k(\mu)$	ratio of backward migration probabilities	$\frac{d_k^{\text{cond}}}{d_k} = \frac{mp(1-\mu)}{\mu(1-m)V_k+mp(1-\mu)}$
$\langle \ldots \rangle$	average of k-indexed variables with weight π_k	
R_{m}	metapopulation fitness	SI section A

Table 1: List of selected variables

expected to evolve with small mutational steps. A singular strategy m^* is that which satisfies $D_1(m^*) = 0$.

When $R_{\text{m}}(m^{\text{mut}}, m^*) < 1$ for all $m^{\text{mut}} \neq m^*$, the singular strategy m^* is uninvadable (evolutionarily stable strategy, ESS; (Maynard Smith and Price, 1973; Maynard Smith, 1976)). Furthermore, it is a fitness maximum, and thus the second-order derivative of the metapopulation fitness with respect to the strategy m^{mut} of the mutant that is then evaluated at $m^{\text{mut}} = m = m^*$, denoted by $D_2(m^*)$, is negative. An uninvadable singular strategy that is also evolutionarily attracting (convergence stable; $D_1'(m^*) < 0$) is an evolutionary endpoint. In contrast, an evolutionarily attracting ($D_1'(m^*) < 0$) singular strategy that is not uninvadable $D_2(m^*) > 0$, is an evolutionary branching point. Such a singular strategy can be invaded by nearby mutants, and the population will become dimorphic, i.e., to consist of two parts having different strategies, and disruptive selection will take the strategies farther away from each other. This process is called evolutionary branching (Geritz et al., 1997, 1998). Evolutionary branching of dispersal has been observed in several models before, including the present model without catastrophes (Parvinen et al., 2020).

3 Evolution of unconditional dispersal

3.1 Metapopulation fitness and fitness gradient

When dispersal is unconditional, the dispersal (emigration) probability does not depend on the patch type, $m_k = m$. Direction of its evolution is predicted by the sign of the following fitness gradient (see SI section B.1.1)

$$\begin{aligned} D_1(m) &= \frac{\partial}{\partial m^{\mathrm{mut}}} R_{\mathrm{m}} \bigg|_{m^{\mathrm{mut}}=m} \\ &= \frac{1}{mp} \sum_{k=1}^{N} \phi_k V_k \left[p - V_k H_k(\mu) + (1 - d_k^{\mathrm{cond}}) V_k H_k(\mu) R_{2,k}^{\mathrm{R}} \right]. \end{aligned} \tag{3.1}$$

As in the model without catastrophes (Parvinen et al., 2020), our expression for the fitness gradient (selection gradient) includes the relative reproductive potential of a patch of type k,

$$V_k = \frac{n_k F_k / \lambda_k}{\sum_{l=1}^{N} \phi_l n_l F_l / \lambda_l}, \tag{3.2}$$

a measure of relatedness,

$$R_{2,k}^{\mathrm{R}} = \frac{1}{n_k - (n_k - 1)(1-\mu)(1 - d_k^{\mathrm{cond}})^2} \tag{3.3}$$

and the conditional backward migration probability,

$$d_k^{\mathrm{cond}} = \frac{mp(1-\mu)}{(1-m)V_k + mp(1-\mu)}, \tag{3.4}$$

which is the proportion of adults that are immigrants given that the patch has not experienced a catastrophe last year. In contrast, the overall backward migration probability is $d_k = (1-\mu) d_k^{\mathrm{cond}} + \mu \cdot 1$, because in a patch that has experienced a catastrophe last year all individuals are immigrants. Finally, the new ingredient in the fitness gradient (3.1),

$$H_k(\mu) = \frac{d_k^{\mathrm{cond}}}{d_k} = \frac{mp(1-\mu)}{\mu(1-m)V_k + mp(1-\mu)} \tag{3.5}$$

is the ratio of backward migration probabilities, which is by definition smaller than or equal to 1. The equivalence of (3.1) with $\mu = 0$ and the expression derived in our earlier work (Parvinen et al., 2020) is most easily observed by looking at equation [2.8] of the SI of (Parvinen et al., 2020). It matches exactly with our expression (3.1) for $\mu = 0$, because we have $H_k(0) = 1$.

Analogous to Parvinen et al. (2020), the fitness gradient in form (3.1) consists of three terms, in which the third one is positive, and corresponds to kin selection. As noted by Frank (1998, p. 120), dispersal depends on the coefficient of relatedness, $R_{2,k}^{\mathrm{R}}$, and $R_{2,k}^{\mathrm{R}}$ depends on dispersal probability, m. Also Jansen and Vitalis (2007) presented a decomposition of the fitness gradient of their model into different terms, involving kin selection. The first two components of (3.1) can be written as (see details in SI section B.1.2)

$$\begin{aligned}\sum_{k=1}^{N}\phi_k V_k\left(p - V_k H_k(\mu)\right) = \quad & \underbrace{-(1-p)}_{\substack{\leq 0 \\ \text{direct cost} \\ \text{of dispersal}}} \quad \underbrace{-\operatorname{Cov}\left[V_k H_k(\mu), V_k\right]}_{\substack{\leq 0 \\ \text{negative effect of spatial} \\ \text{heterogeneity}}} \\ & \underbrace{-\operatorname{Cov}\left[V_k, H_k(\mu)\right]}_{\substack{\geq 0 \\ \text{positive effect of catastrophes} \\ \text{and spatial heterogeneity}}} \quad + \quad \underbrace{\left(1-\operatorname{E}\left[H_k(\mu)\right]\right)}_{\substack{\geq 0 \\ \text{positive effect of} \\ \text{catastrophes}}}\end{aligned} \tag{3.6}$$

In (3.6), the first term, $-(1-p)$ is non-positive, and can, as in Parvinen et al. (2020), be interpreted as the direct cost of dispersal. The second term of (3.6) is also non-positive. In the absence of catastrophes $H_k(0) = 1$ holds, and thus $-\operatorname{Cov}\left[V_k H_k(\mu), V_k\right] = -\operatorname{Var}\left[\mathbf{V}\right]$, matching exactly the term presented by Parvinen et al. (2020) in the absence of catastrophes. This term can be interpreted as the negative effect of spatial heterogeneity. The remaining terms in (3.6) are non-negative, and become zero when $\mu = 0$. They can be interpreted as the positive effect of catastrophes and spatial heterogeneity.

Analogous to fitness gradient $D_1(m)$, we were able to derive the second-order derivative of metapopulation fitness at singular strategy, denoted by $D_2(m^*)$, in SI section B.2, Eq. (B.21). It tells us if "evolutionary branching" occurs and if the population becomes dimorphic in the dispersal trait. Drawing general conclusions from that expression is, however, not straightforward. As in Parvinen et al. (2020), the format suggests that evolutionary branching may happen in the presence of spatial heterogeneity. This expectation is confirmed by our numerical explorations below.

3.2 Intuitions obtained from an inclusive-fitness based derivation

To obtain intuitions from a different perspective, we have also performed an inclusive-fitness based analysis, as performed in previous papers that studied evolution of dispersal Taylor (1988); Jansen and Vitalis (2007). The detailed calculations are described in SI section F. The analysis performed there shows that, given the resident population adopts dispersal probability m, the inclusive fitness $W_{\mathrm{IF}} = W_{\mathrm{IF}}(m)$ of a mutant with a slightly deviant dispersal probability $m + \Delta m$ is given, up to the first order of Δm, by

$$W_{\mathrm{IF}} = 1 + (\Delta m) \cdot W_{\mathrm{IF}}^{(1)}, \tag{3.7a}$$

where

$$W_{\mathrm{IF}}^{(1)} = \sum_{k=1}^{N} \underbrace{c_k}_{\substack{\text{impact of}\\ \text{each patch type}}} \left[\underbrace{\frac{d_k}{m}}_{\substack{\text{gain of}\\ \text{dispersing juveniles}}} - \underbrace{\frac{1-d_k}{1-m}}_{\substack{\text{loss of}\\ \text{natal juveniles}}} + \underbrace{\frac{1}{1-m} R_{2,k}}_{\substack{\text{gain from}\\ \text{weakened kin-competition}}} \right]. \tag{3.7b}$$

(see Eq. (F.49) in SI section F; we put $m_k = m$ there). Note that the inclusive fitness effect, $W_{\mathrm{IF}}^{(1)} = W_{\mathrm{IF}}^{(1)}(m)$, is the same as $D_1(m)$ in (3.1) up to a positive multiplicative constant, so the predictions based on one calculation shall be exactly the same as those based on the other. There is also term-wise correspondence between D_1 in (3.1) and $W_{\mathrm{IF}}^{(1)}$ in (3.7b); see Eq. (F.56) for details. Despite this mathematical equivalence, deriving $W_{\mathrm{IF}}^{(1)}$ gives additional intuitions from a different perspective.

Since patches of different types contribute differently to a future gene pool, we rely on the theory of class-structured populations (Taylor, 1990) Specifically, c_k in (3.7b) represents the class reproductive value (Taylor, 1990; Rodrigues and Gardner, 2022) of the patches of type k, which intuitively means the relative genetic contribution of all the patches of type k (see Eq. (F.39) in SI for its explicit expression). Each term in the squared brackets in (3.7b) represents the fitness gain of mutants in type k patches, after a proper adjustment of reproductive values.

For the sake of explanation, suppose that mutants adopt a slightly higher dispersal probability than that of residents. Then, the first term in the square brackets in (3.7b) accounts for the direct benefit through the increased number of dispersing juveniles. This benefit is proportional to the backward migration probability d_k, intuitively because larger d_k means more intense competition at type k patches, which makes dispersal from those patches more adaptive. At the same time, this benefit is inversely proportional to the dispersal probability m, intuitively because smaller dispersal probability by others make dispersal more adaptive due to frequency-dependent selection.

Similarly, the second term in the square brackets in (3.7b) accounts for the direct cost through the decreased number of natal juveniles. This cost is proportional to the backward non-migration probability $1 - d_k$ (i.e., the probability with which a random individual is natal to its patch), intuitively because larger $1 - d_k$ means milder competition at type k patches, which makes staying in those patches more adaptive, which in turn means that dispersal from those patches is more costly. At the same time, this cost is inversely proportional to the non-dispersal probability $1 - m$, intuitively because smaller non-dispersal probability by others make non-dispersal more adaptive due to frequency-dependent selection, which means that dispersal from those patches is more costly.

Moreover, the third term in the square brackets in (3.7b) accounts for the kin-selected indirect benefit through the increased number of dispersing juveniles. Because relatedness $R_{2,k}$ includes m in it, discussing this term's dependence on m is not straightforward, but $1/(1-m)$ represents the benefit of mitigated competition due to increased dispersal, which is enjoyed by the neighboring individuals in the same patch. An adult individual in a patch of type k is on average related to those neighbors by the coefficient of relatedness $R_{2,k}$, which is formally defined as the average relatedness between two individuals sampled without replacement from the same patch of type k (see Eq. (F.47)). The focal adult individual, therefore, enjoys the kin-selected benefit of this amount.

The net inclusive fitness effect of dispersal from type k patches, which is the sum of the three terms in the square brackets in (3.7b), is then summed over all patch types, while each being weighted by its class reproductive value, c_k. The resulting weighted sum is the inclusive fitness effect $W_{\mathrm{IF}}^{(1)}$ in (3.7b), which is sign-equivalent to the fitness gradient $D_1(m)$ in (3.1) that was calculated based on the meta-population fitness approach. It is now clear that those benefit- and cost- terms (3 terms) summed over different patch types (N types) determines selection. In other words, the balance of these $3N$ terms shapes the evolution of unconditional dispersal.

3.3 Singular strategies

Singular strategies are such strategies for which the fitness gradient vanishes, $D_1(m) = 0$. In SI sections C.1, C.2, and C.4.2, we show that, unless some special conditions are met, there exists a unique singular strategy m^* with $0 < m^* < 1$, and it is evolutionarily attracting.

The general expression for the fitness gradient (3.1) is so complicated that we do not obtain a general explicit expression for singular strategies. In case of just one patch type, $N = 1$, the expression of the fitness gradient is simpler, but for $\mu > 0$ obtaining singular strategies requires solving a cubic equation. Explicit formulas for its solutions exist, but they involve complex numbers, and are of little practical use for our purpose. However, in a few special cases, we are able to obtain an explicit expression for the singular strategy, as will be explained in the following section 3.3.1. After that we consider the effect of model parameters on the singular strategies, for which some analytical results can be obtained (sections 3.3.2-3.3.5), and on evolutionary branching (section 3.3.6).

3.3.1 Singular strategies in the absence of spatial heterogeneity

When we have only one patch type ($N = 1$), we can for simplicity denote the number of adults in a patch as $n = n_1$. First, in the absence of catastrophes, $\mu = 0$, we again obtain (as we did in Parvinen et al. (2020), therein Eq. 8)

$$m^* = \frac{1 + 2n(1-p) - \sqrt{1 + 4n(n-1)(1-p)^2}}{2n(1-p)(2-p)}, \tag{3.8}$$

which with $c = 1 - p$ is familiar from previous work (Ajar, 2003, therein Eq. 14), see also Frank (1986). For $n = 1$ (3.8) becomes the classical Hamilton and May (1977) result, $m^* = 1/(2 - p)$.

Second, with catastrophes, for $n \to \infty$ we obtain $m^* = \mu/(1 - (1 - \mu)p)$, which was obtained already by Comins et al. (1980, therein Eq. 4).

Third, with catastrophes, for $n = 1$ we obtain

$$m^* = \frac{2\mu}{2\mu - p(1-\mu^2) + (1-\mu)\sqrt{p(p(1-\mu)^2 + 4\mu)}} \tag{3.9}$$

For $\mu > 0$, the denominator of (3.9) is positive and m^* given by (3.9) satisfies $0 < m^* \leqslant 1$, and it is a non-monotonic function of the dispersal survival probability p. We will investigate this phenomenon for the general model below (section 3.3.4).

3.3.2 Catastrophes promote the evolution of dispersal

Consider the ingredients of the fitness gradient, (3.1). Based on SI Eq. (C.44), $H_k(\mu)$ is a decreasing function of μ. Consequently, increasing the catastrophe probability μ will make the second component, $-V_k H_k(\mu)$, weaker. If this was the only effect of μ on the fitness gradient, increasing μ would clearly promote dispersal. However, the catastrophe probability also affects the third term of the fitness gradient, the positive term corresponding to kin selection. Based on SI Eq. (C.45), d_k^{cond} decreases with μ, and thus the multiplier $(1 - d_k^{\text{cond}})$ increases with μ, but as noted above, $H_k(\mu)$ is a decreasing function of μ. Finally, increasing μ can affect relatedness $R_{2,k}^{\text{R}}$ in a non-monotonic way. Therefore, it is not so straightforward to see what the effect of increasing the catastrophe probability μ is on the fitness gradient, and further, on the singular strategies. We were, however, able to prove that the singular strategies increase with respect to μ in our model (details in SI section C.4.5). This is illustrated in Fig. 1, and Figs. S2, S3 and S5.

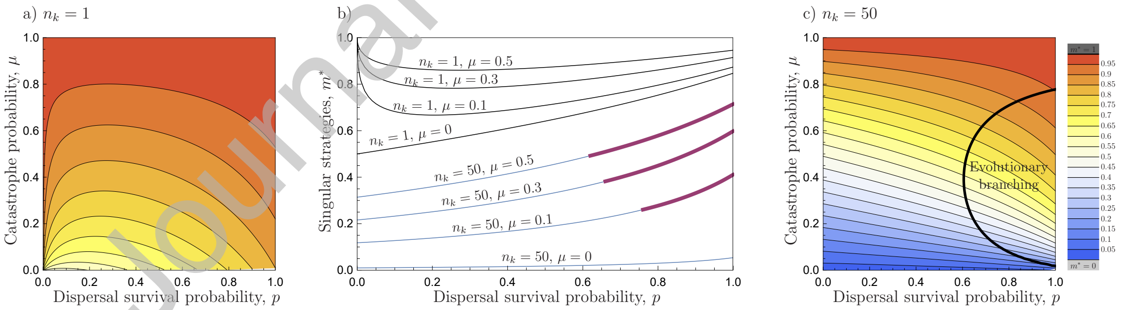

Figure 1: Catastrophes promote unconditional dispersal, and the dispersal survival probability can have a non-monotonic effect. Dependence of the singular strategy on the dispersal survival probability p and catastrophe probability μ in a model with two patch types ($N = 2$) with equal size ($n_1 = n_2 = n_k$) and attractivity $\lambda_1 = \lambda_2 = 1$, but different fecundity ($F_1 \neq F_2$). In panel b, thin curves correspond to uninvadable strategies, thick purple curves to branching points. Parameters: $f = F_1/F_2 = 3$, $\pi_1 = \pi_2 = 0.5$.

One mechanism how catastrophes promote dispersal is that they result in empty patches, dispersing into which is profitable. This mechanism is present in our model. One should note, however, that non-monotonic effects of catastrophes on dispersal evolution have been observed in other models (Ronce et al., 2000; Gyllenberg et al., 2002; Parvinen et al., 2003; Parvinen, 2006). In these models, when the catastrophe rate is high, most individuals reside in sparsely populated patches anyway, so the benefit of dispersal is diminished. This mechanism resulting in non-monotonic effects of catastrophes is not present in our model. This is because in our model, patches are either empty due to catastrophes, or fully occupied (saturation). The same feature holds for the model by Jansen and Vitalis (2007), and catastrophes promote dispersal therein. Another mechanism, which could cause non-monotonic effects of catastrophes is kin selection (relatedness). That mechanism is not present in Gyllenberg et al. (2002) and Parvinen (2006), because local populations therein are large. In our model dispersal is monotonically increasing with respect to the catastrophe probability μ, because the direct effect of empty patches is stronger than the indirect effects on relatedness (Theorem 11 in SI section C.4.5).

3.3.3 High sensitivity on parameters p and μ when they are small

If p is exactly zero, the model is not meaningful, because no dispersers will survive dispersal. However, it is of interest to study what happens when p is small, as the metapopulation is viable and positive dispersal will always evolve in our model, even with p very close to zero. SI section C.4.3 shows the limit of the fitness gradient, $\lim_{p\to 0} D_1(m)$. Interestingly, it is not continuous with respect to μ at $\mu = 0$. The singular strategy in the model with an infinitesimally small catastrophe probability is significantly larger than the singular strategy in the model without catastrophes, when we let $p \to 0$. Such a discontinuity between cases of $\mu = 0$ and $\mu > 0$ was earlier suggested by Comins et al. (1980) (described as "*abrupt change*"; pp. 214). This effect is clearly visible in Fig. 1b for $n_k = 1$. For $\mu > 0$, the singular dispersal rate approaches $m^* = 1$ as $p \to 0$, whereas in the absence of catastrophes (i.e., $\mu = 0$) the singular dispersal rate approaches $m^* = 1/2$ in the same limit, $p \to 0$. In case all patches have the same population size, $n_k = n$, the limit of the singular strategy is

$$\begin{cases} \lim_{p\to 0} \left(m^* \big|_{\mu>0} \right) & = \frac{1+\mu(n-1)}{n}, \\ \lim_{p\to 0} \left(m^* \big|_{\mu=0} \right) & = \frac{1}{2n}. \end{cases} \tag{3.10}$$

It turns out, that the singular strategy m^* does not have a well-defined limit when both p and μ approach zero. The limit value namely depends on the route taken. This is further illustrated in Fig. 2: Fig. 2a is a magnification of Fig. 1a near $(p, \mu) \approx (0, 0)$. Generically, contour curves should not intersect, because at such points the illustrated function is not well-defined. However, each contour curve in Fig. 2a approaches the origin. The solid arrow illustrates the route of taking the limit of the singular strategy m^* for $(p, \mu) \approx (0, 0)$, while keeping the ratio $\mu/p = \tan\alpha$ constant. Fig. 2b shows the value of this limit as a function of the angle of approach, α, for a metapopulation with one patch type ($N = 1$), for different values of the number of adults in a patch, n. In the most illuminating example we have

$$\begin{cases} \lim\limits_{\mu\to 0}\left(\lim\limits_{p\to 0} m^*\right) = \dfrac{1}{n} & (\text{case } \alpha = 90^\circ), \\ \lim\limits_{p\to 0}\left(\lim\limits_{\mu\to 0} m^*\right) = \dfrac{1}{2n} & (\text{case } \alpha = 0^\circ). \end{cases} \tag{3.11}$$

The former double-limit corresponds to the case of $p \ll \mu$, and the latter double-limit corresponds to the case of $\mu \ll p$. In fact, we can prove that the limit value of m^* for $p, \mu \to 0$ is larger when they approach zero with keeping a larger (μ/p)-ratio (Figure 2b; see also SI section F.9 for the detail).

a) Zoomed part of Fig. 1a

b) Limit when $(p, \mu) \to (0, 0)$

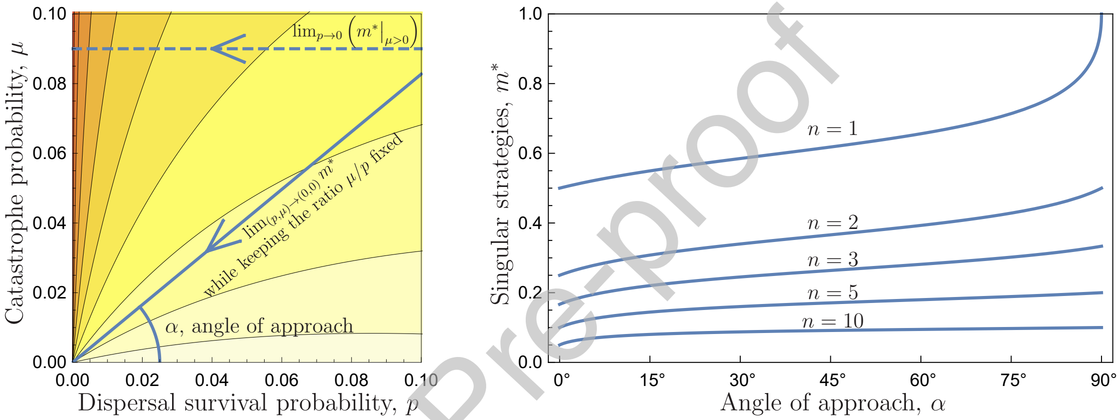

Figure 2: a) A magnification of Fig. 1a near $(p, \mu) \approx (0, 0)$ together with arrows illustrating different limits. b) The limit value of m^* when $p \to 0$ and $\mu \to 0$, while keeping the ratio $(\mu/p) = \tan\alpha$, plotted with respect to α, for a metapopulation with one patch type ($N = 1$), for different numbers of adults in the patch (n). The effect of the route choice is strongest for small n.

An intuition behind this strong dependence on the (μ/p)-ratio is as follows. When p is close to zero, there are few survivors of dispersal. When μ is close to zero, on the other hand, there are few empty patches created by catastrophes. The balance between the availability of those two is crucial, because if there are relatively much more empty patches than surviving dispersers, then those dispersers who have successfully arrived in an empty patch can enjoy a big leap of benefits of colonizing it because the competition there is weak due to relatively small p, which makes dispersal more adaptive. On the other hand, if there are relatively much more surviving dispersers than empty patches, the competition over empty patches is very intense due to relatively large p, and there is little merit in dispersal. Therefore, the balance between p and μ matters however small they are; large μ/p-ratio means less competition at empty patches for dispersers, and hence a relatively larger dispersal rate is selected for by natural selection.

In our model, the metapopulation is viable for all $p > 0$, because fecundities in different patch types are assumed to be infinite ($\lambda \to \infty$). The situation is different, if fecundities are finite. An example of such a model is the model of Comins et al. (1980), which is analysed in more detail in SI section G. As illustrated by the gray regions in Fig. S15b-e, their metapopulation is not viable, if p is too small. Consequently, even though there are "abrupt changes", when p and μ are small (Fig. S15), investigating the limit $p \to 0$ is not meaningful with finite fecundity.

3.3.4 Non-monotonic effect of the dispersal survival probability

The dispersal survival probability p appears in the first component of the fitness gradient (3.1) in our model. This component increases with p, as increasing the dispersal survival probability will increase the direct benefit of dispersal. For this reason, one could easily expect the singular dispersal strategy m^* to be an increasing function of p. Such a phenomenon has indeed been observed in many models, including the present model in the absence of catastrophes ($\mu = 0$) (Parvinen et al., 2020). However, in our model, other components of (3.1) are also affected by p, and the singular dispersal strategy m^* can be non-monotonic with respect to p, when $\mu > 0$. In SI section C.4.4 we show that the derivative of the fitness gradient with respect to the dispersal survival probability p is given by

$$\frac{\partial}{\partial p} D_1(m,p) = \frac{1}{mp^2} \sum_{k=1}^{N} \phi_k \left(V_k\right)^2 H_k(\mu) A(d_k^{\text{cond}}, \mu, n_k), \tag{3.12}$$

in which

$$\begin{aligned} A(d_k^{\text{cond}}, \mu, n_k) =& 1 - (1 - d_k^{\text{cond}}) \Bigg(\frac{\mu}{\mu + (1-\mu) d_k^{\text{cond}}} \\ &+ \frac{d_k^{\text{cond}}}{(\mu + (1-\mu) d_k^{\text{cond}})[n_k - B]} + \frac{d_k^{\text{cond}}(n_k + B)}{[n_k - B]^2} \Bigg). \end{aligned} \tag{3.13}$$

and

$$B = (n_k - 1)(1-\mu)(1 - d_k^{\text{cond}})^2 \geqslant 0. \tag{3.14}$$

In the absence of catastrophes we have $A(d_k^{\text{cond}}, 0, n_k) > 0$ for $d_k^{\text{cond}} > 0$, which means that increasing p promotes dispersal, and the singular strategy m^* increases with p, as already shown before (Parvinen et al., 2020). In contrast, $A(d_k^{\text{cond}}, \mu, n_k)$ can be negative when $\mu > 0$. Such parameter regions are illustrated in Figs. 3 and S1 with different shades of blue for different n_k. If

$$\begin{cases} 0 < \mu < 1, & \text{when } n_k = 1 \\ 0 < \mu < \frac{1}{n_k - 1}, & \text{when } n_k \geqslant 2 \end{cases} \tag{3.15}$$

holds, the expression $A(d_k^{\text{cond}}, \mu, n_k)$ is negative for $d_k^{\text{cond}} > 0$ below some threshold. Consequently, if (3.15) holds, the singular strategy is a non-monotonic function of p (Figs. 1, S2 and S3). This effect is very clearly visible for $n_k = 1$ in Fig. 1, but not so clearly for $n_k = 50$. However, the effect is present even for large n, when μ is small enough, as illustrated by Fig. S4. Similar non-monotonic effects of

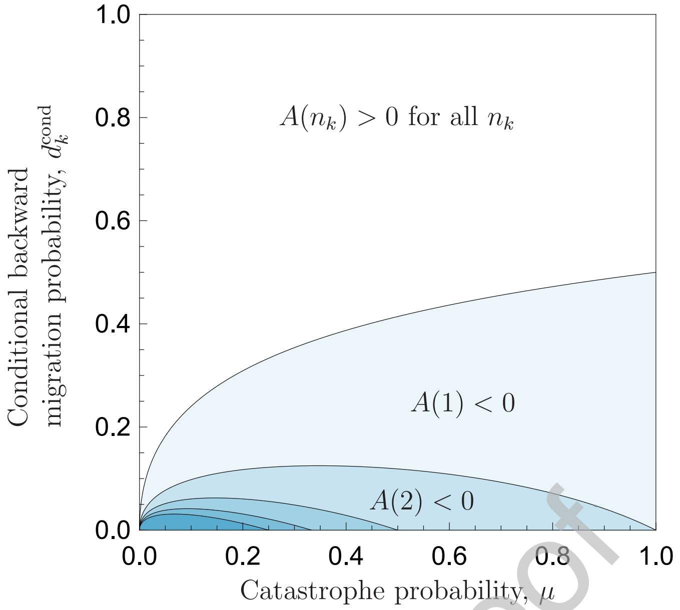

Figure 3: Parameter regions of the sign of $A(d_k^{\mathrm{cond}}, \mu, n_k)$ given by (3.13), with respect to d_k^{cond} and μ for different n. See also Fig S1.

the dispersal survival probability p has been earlier found numerically by Comins et al. (1980), their Figs. 4 and 5. Their model has one patch type, and fecundity is assumed to be finite. Their and our results together illustrate that the non-monotonic effect of p is not an artifact of either type of fecundity alone. We investigate their model in SI section G, in case the number of adults in a patch is (at most) one, $n = 1$. As illustrated by Fig. S15b-e, non-monotonic effect of p is observed when fecundity is large enough.

Altogether, when p is very small, there are only few immigrants arriving into patches. This has twofold consequences. On one hand, a quite small amount of natal juveniles is enough to colonize the natal patch with high probability. Since also relatedness is high, avoiding kin competition also promotes dispersal. On the other hand, because of catastrophes, there are empty patches, and those dispersers who arrive in such patches, have a very high probability of establishment. This effect is significant even when μ is small and the amount of empty patches is low.

3.3.5 Spatial heterogeneity typically selects against dispersal

Based on (3.6), analogous to the model without catastrophes (Parvinen et al., 2020), spatial heterogeneity, in the form of variability in V_k, is expected to select against dispersal. This is observed in Figs. 4, S5 and S6. However, in the presence of catastrophes, spatial heterogeneity may promote dispersal when p is small, as illustrated by Fig. S8.

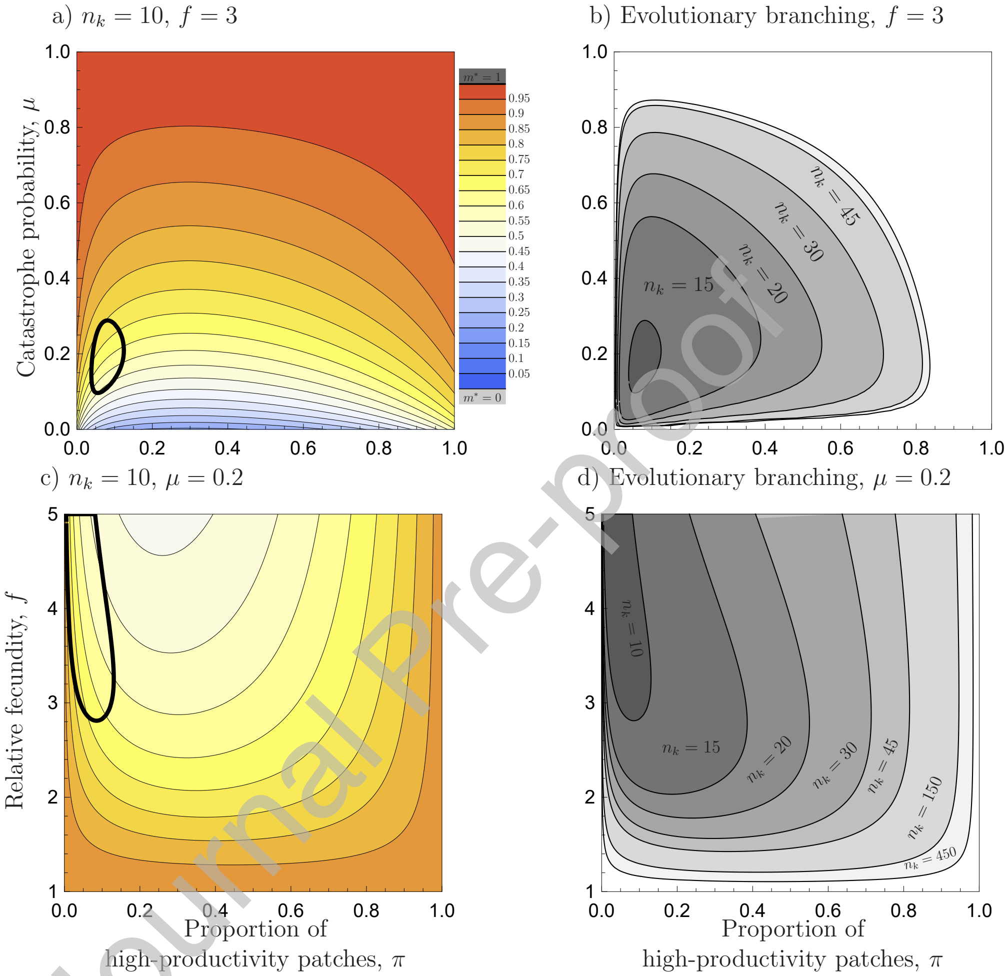

Figure 4: Spatial heterogeneity typically selects against unconditional dispersal. Singular strategies (a,c) and parameter regions with evolutionary branching (b,d) with respect to the proportion of high-productivity patches $\pi = \pi_1 = 1 - \pi_2$ and (a,b) the catastrophe probability, or (c,d) the relative fecundity $f = F_1/F_2$ in a model with two patch types ($N = 2$) with equal size ($n_1 = n_2 = n_k$) and attractivity $\lambda_1 = \lambda_2 = 1$, but different fecundity. Common parameters: $p = 0.95$, $\lambda_1 = \lambda_2 = 1$.

3.3.6 Evolutionary branching is promoted by spatial heterogeneity and large patch sizes

In the absence of catastrophes, Parvinen et al. (2020) observed that evolutionary branching is possible under spatial heterogeneity, and is promoted by differences in patch size. More specifically, variability in V_k generated by differences in patch sizes resulted in evolutionary branching under much wider parameter ranges than variability in V_k generated by differences in fecundity or attractivity only. In contrast, in the presence of catastrophes, we observe that spatial heterogeneity generically promotes evolutionary branching. In specific, the size of the parameter range in which evolutionary branching occurs seems to be more or less of the same size, independently of the origin of the spatial heterogeneity (compare Fig. S2 with S3, and the rows of Fig. S5), but note that they differ qualitatively, when $\mu \approx 0$. Those figures also clearly demonstrate, that increasing patch sizes, while keeping V_k unchanged, promotes evolutionary branching.

4 Evolution of conditional dispersal

Let us next investigate the evolution of conditional dispersal, i.e., we assume that the dispersal probability m_k may depend on the patch type k.

4.1 The fitness gradient

We were able to derive an explicit expression also for the components of the fitness gradient of conditional dispersal (Theorem 3 in the SI section B.1):

$$\begin{aligned} D_{1,k}((m_1,\dots,m_N)) &= \frac{\partial}{\partial m_k^{\mathrm{mut}}} R_{\mathrm{m}} \bigg|_{m_1^{\mathrm{mut}}=m_1,\dots,m_N^{\mathrm{mut}}=m_N} \\ &= \frac{\phi_k W_k}{m_k p} \left(p - W_k H_k(\mu) + (1-d_k^{\mathrm{cond}}) W_k H_k(\mu) R_{2,k}^{\mathrm{R}} \right). \end{aligned} \tag{4.1}$$

When compared with the fitness gradient of unconditional dispersal ((3.1)), the relative reproductive potential V_k is replaced by

$$W_k = \frac{m_k n_k F_k/\lambda_k}{\langle mnF/\lambda \rangle} = \frac{m_k n_k F_k/\lambda_k}{\sum_{l=1}^N \phi_l m_l n_l F_l/\lambda_l}, \tag{4.2}$$

which measures the relative production of dispersers. For $m_k = m$ we have $W_k = V_k$. Furthermore, the conditional backward migration probability involves the strategy component m_k,

$$d_k^{\mathrm{cond}} = \frac{m_k p(1-\mu)}{(1-m_k)W_k + m_k p(1-\mu)}, \tag{4.3}$$

using which relatedness $R_{2,k}^{\mathrm{R}}$ is given by (3.3), and and the relative conditional backward migration probability $H_k(\mu) = d_k^{\mathrm{cond}}/d_k = m_k p(1-\mu)/(\mu(1-m_k)W_k + m_k p(1-\mu))$.

Fig. 5 shows qualitatively different phase-plots of the fitness gradient of conditional dispersal for two patch types, $N = 2$. If the fitness gradient isoclines would be straight lines, all qualitatively different phase-plots would be as illustrated in Fig. 5. In reality the isoclines are not straight lines, but close to such. Nevertheless, these are the qualitatively different phase-plots that we could find in our numerical explorations for two patch types, $N = 2$.

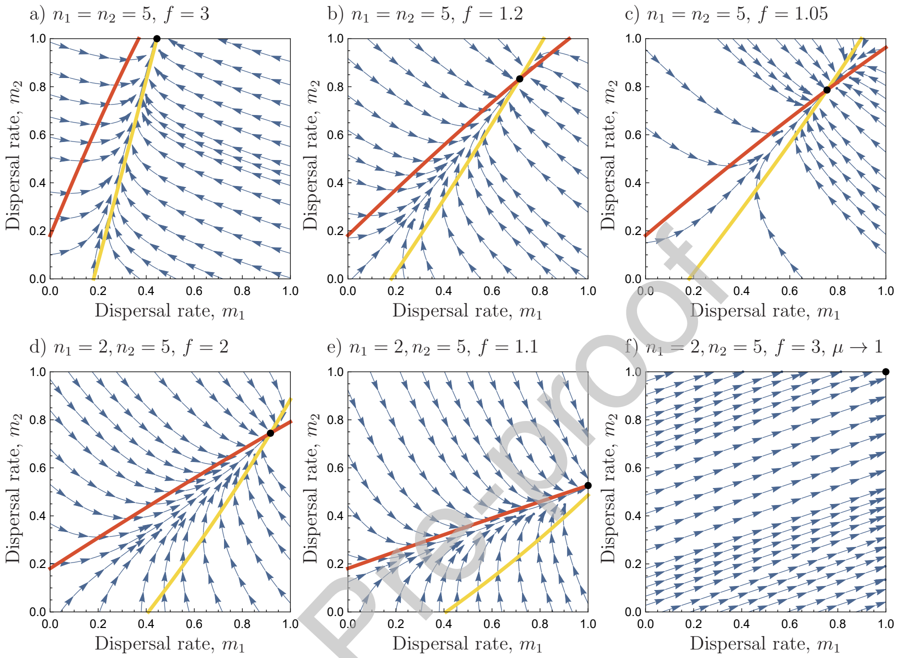

Figure 5: Qualitatively different phase-plots of conditional dispersal, illustrating the direction of the fitness gradient. The isoclines of the fitness gradient are shown with thick yellow ($D_{1,1} = 0$) and red ($D_{1,2} = 0$) curves. Singular (or boundary-singular) strategies are marked with a black dot. Parameters: $p = 0.95$, $\pi = 0.5$, $\mu = 0$, $\lambda = 1$.

As in the previous section, we are able to obtain an equivalent expression to the fitness gradient $D_{1,k}$ in (4.1) by using an inclusive-fitness based calculation. Details are described in SI section F, and we simply describe the result here. The inclusive fitness of a mutant that adopts the dispersal strategy $\boldsymbol{m}^{\text{mut}} = (m_1 + \Delta m_1, \ldots, m_N + \Delta m_N)$ in a resident population of $\boldsymbol{m} = (m_1, \ldots, m_N)$ is given, up to the first order of Δm_k's, by

$$W_{\text{IF}} = 1 + \sum_{k=1}^{N} (\Delta m_k) W_{\text{IF}}^{(1,k)}, \tag{4.4a}$$

where

$$W_{\text{IF}}^{(1,k)} = \underbrace{c_k}_{\substack{\text{impact of}\\\text{each patch type}}} \left[\underbrace{\frac{d_k}{m_k}}_{\substack{\text{gain of}\\\text{dispersing juveniles}}} - \underbrace{\frac{1-d_k}{1-m_k}}_{\substack{\text{loss of}\\\text{natal juveniles}}} + \underbrace{\frac{1}{1-m_k} R_{2,k}}_{\substack{\text{gain from}\\\text{weakened kin-competition}}} \right]. \tag{4.4b}$$

(see Eq. (F.49) in SI section F). Similar to the previous section, the inclusive fitness effect, $W_{\mathrm{IF}}^{(1,k)} = W_{\mathrm{IF}}^{(1,k)}(\boldsymbol{m})$, is the same as $D_{1,k}(\boldsymbol{m})$ in (4.1) up to a positive multiplicative constant. There is also term-wise correspondence between $D_{1,k}$ in (4.1) and $W_{\mathrm{IF}}^{(1,k)}$ in (4.4b); see SI Eq. (F.55) for details. The intuitions we obtain from each term in the square brackets in (4.4b) are parallel to those we obtained from (3.7b), so we do not repeat the arguments here.

4.2 Singular strategies

In SI sections C.1 and C.2 we show that positive dispersal evolves for all dispersal strategy components. Furthermore, unless some special conditions are met, complete dispersal ($m_1^* = \cdots = m_N^* = 1$) does not evolve. Potential evolutionary endpoints are thus singular strategies, satisfying

$$D_{1,k}((m_1, \ldots, m_N)) = 0, \quad 0 < m_k < 1 \tag{4.5}$$

for all k, and boundary-singular strategies, satisfying $D_{1,k}((m_1, \ldots, m_N)) \geqslant 0, m_k = 1$ for some but not all components, (4.5) for other components.

For conditional dispersal, we were not able to prove the uniqueness of the (boundary-)singular strategy. Nevertheless, in our numerical explorations for two patch types, $N = 2$, we did always find a unique (boundary-)singular strategy (m_1^*, m_2^*) satisfying the conditions stated above. We found this strategy to be uninvadable, i.e., we did not observe evolutionary branching of conditional dispersal. Qualitatively different phase-plots are illustrated in Fig. 5. Next we investigate how the singular conditional dispersal strategy is affected by parameters.

4.2.1 Catastrophes promote the evolution of dispersal

Analogous to the case of unconditional dispersal, we proved that the components of the fitness gradient vector of conditional dispersal increase with respect to μ in our model (details in SI section C.4.5). Provided with the numerically observed shape of the fitness gradient isoclines near the singular strategy, the components of the singular strategy also increase with μ. (Figs. 6a, S13)

4.2.2 High sensitivity on parameters p and μ when they are small & Non-monotonic effect of the dispersal survival probability

Analogous to the case of unconditional dispersal, the conditional singular strategy components are sensitive to the parameters p and μ when they are small. This effect can be seen in Fig. 6bc for $n_k = 1$. For $\mu = 0.1$, the singular dispersal strategy components approach $m_1^* = m_2^* = 1$ as $p \to 0$ (Fig. 6c), whereas for $\mu = 0$ they approach $m^* = 1/2$ (Fig. 6b). In fact, the components of the conditional singular dispersal strategy, when the dispersal survival probability p goes to zero, are

$$\begin{cases} \lim_{p\to 0} \left(m_k^* \big|_{\mu>0} \right) & = \frac{1+\mu(n_k-1)}{n_k}, \\ \lim_{p\to 0} \left(m_k^* \big|_{\mu=0} \right) & = \frac{1}{2n_k}. \end{cases} \tag{4.6}$$

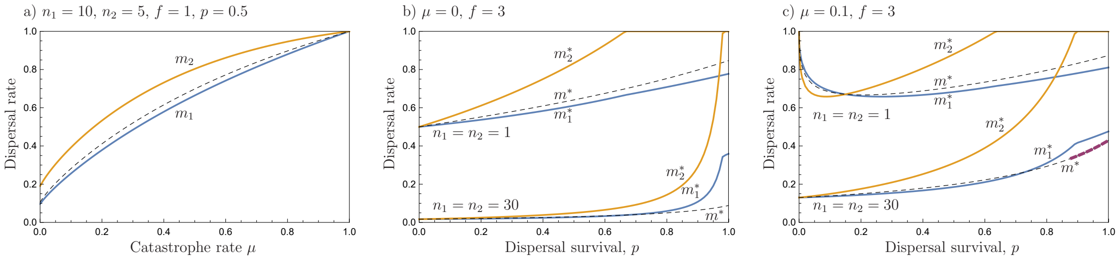

Figure 6: Dependence of singular conditional (m_1^*, m_2^*) and unconditional (m^*) (dashed curve) dispersal strategies on a) the catastrophe probability μ, b and c) dispersal survival probability for different values on n_k and μ in a model with two patch types $(N = 2)$. Parameters: a) $p = 0.5$, b and c) $f = F_1/F_2 = 3$, Common parameters: $\lambda_1 = \lambda_2 = 1$, $\pi_1 = \pi_2 = 0.5$.

We thus observe an analogous discontinuity with respect to μ as for unconditional dispersal (3.10).

The singular strategy components m_k^* do not have a well-defined limit when both p and μ approach zero. As in the unconditional dispersal case, the limit value depends on the route taken (see Fig 7). In specific, analogous to (3.11), we have

$$\begin{aligned} \lim_{\mu \to 0} \left(\lim_{p \to 0} m_k^* \right) &= \frac{1}{n_k} \quad (\text{case } \alpha = 90^\circ) \\ \lim_{p \to 0} \left(\lim_{\mu \to 0} m_k^* \right) &= \frac{1}{2n_k} \quad (\text{case } \alpha = 0^\circ). \end{aligned} \tag{4.7}$$

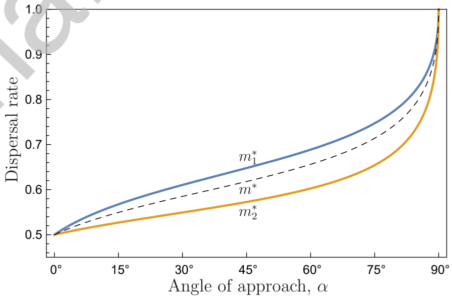

Figure 7: The limit values of the singular conditional dispersal strategy components m_1^* and m_2^* (solid curves) together with the unconditional strategy m^* (dashed curve), when $p \to 0$ and $\mu \to 0$, while keeping the ratio $(\mu/p) = \tan\alpha$, plotted with respect to α, for a metapopulation with two patch types $(N = 2)$. Parameters: $n_1 = n_2 = 1$, $f = F_1/F_2 = 3$, $\lambda_1 = \lambda_2 = 1$, $\pi_1 = \pi_2 = 0.5$.

Finally, the sign of $\frac{\partial}{\partial p} D_{1,k}((m_1, \ldots, m_N))$ is given by expression $A(d_k^{\mathrm{cond}}, \mu, n_k)$ (Eq. (C.31) in the SI). Therefore, we observe a similar effect of increasing the dispersal survival probability p also for conditional dispersal (Figs. 6bc, S10): the singular strategy component m_k^* decreases with p when p is small, if (3.15) holds.

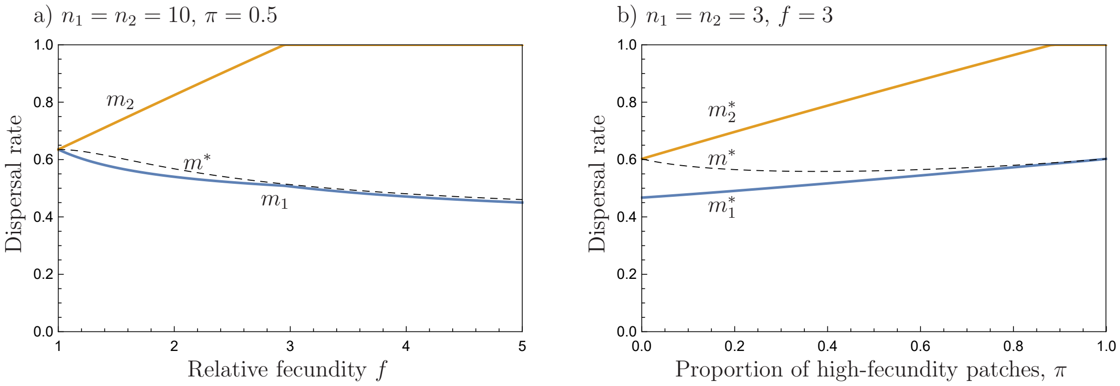

Figure 8: Dependence of singular conditional (m_1^*, m_2^*) and unconditional (m^*) dispersal strategies on a) relative fecundity $f = F_1/F_2$ b) proportion of high-fecundity patches $\pi = \pi_1 = 1 - \pi_2$ in a model with two patch types $(N = 2)$. Parameters: a) $n_1 = n_2 = 10$, $\pi = 0.5$ b) $n_1 = n_2 = 3$, $f = F_1/F_2 = 3$. Common parameters: $\lambda_1 = \lambda_2 = 1$, $\mu = 0.2$, $p = 0.8$.

4.2.3 Difference in evolved dispersal components

Individuals experience different selective pressure in different patch types. Therefore, conditional dispersal strategies usually do not evolve to be equal in different patch types. Which features of different patch types then promote or select against conditional dispersal? Increasing the size n_k or the relative fecundity f_k of a specific patch type increases its relative production of dispersers, W_k. The form of the fitness gradient (4.1) suggests that this would select against dispersal. Furthermore, increasing the patch size will decrease relatedness $R^{\mathrm{R}}_{2,k}$, which is also expected to select against dispersal. In Fig. 6 patch type 1 is more productive either because of larger size ($n_1 > n_2$, $f = F_1/F_2 = 1$, panel a) or larger fecundity ($f = F_1/F_2 > 1$, $n_1 = n_2$, panels bc). We indeed observe in Fig. 6ab that the conditional dispersal strategy $m_2^* > m_1^*$ as expected. Fig. S12c illustrates that the relative production of dispersers alone does not fully explain dispersal evolution: Although $m_2^* > m_1^*$ when f is large and $m_2^* < m_1^*$ when f is small, for equal productivity ($fn_1 = n_2$ when $f = 4$), $m_2^* < m_1^*$ because $n_2 > n_1$, resulting in smaller relatedness in patch type 2 (see also Fig. S12e).

Surprisingly, in Fig. 6c and S10ceg we observe that dispersal is larger in the more productive patch type 1, $m_1^* > m_2^*$, when p is small. Actually, (4.6) shows that when $p \to 0$, the relative fecundities do not affect singular dispersal strategies. Dispersal is then largest in patch types with smallest size even if they would be very productive (S10eg). In contrast, our numerical results suggest that for large p the effect of productivity dominates, and $m_1^* < m_2^*$. For equal patch sizes a similar phenomenon occurs, when (3.15) holds (Fig. 6c, S10c).

When comparing the components of the singular conditional dispersal strategy with the singular unconditional dispersal strategy, one would expect the unconditional one to be some kind of biased average of the conditional ones, i.e., to lie in between m_1^* and m_2^* for two patch types. This is what we usually observe in our numerical explorations, but not always (Fig. 6c with $n_1 = n_2 = 30$, large p). At a singular unconditional dispersal strategy, the sum of the components of the conditional dispersal strategy fitness gradient vector is zero. Therefore, if conditional dispersal evolution is initiated at the unconditional dispersal strategy, initially some conditional strategy components will decrease and others increase, leaving the unconditional singular dispersal strategy to lie in between them. However, this situation does not need to hold at the singular conditional dispersal strategy, as illustrated by Fig. S9.

4.2.4 The effect of spatial heterogeneity

When p is large, variability in V_k is expected to select against unconditional dispersal: The unconditional singular dispersal strategy m^* in Fig. 4ac is lowest for an intermediate proportion of high-productivity patches π, and decreases when the relative fecundity f is increased. In contrast, both conditional dispersal strategy components m_1^* and m_2^* in Figs. 8b and S11ac increase, when π is increased. When the proportion of patches with higher productivity of dispersers is increased, the average productivity $\langle mnF/\lambda \rangle$ increases, and thus the relative productivity W_k (4.2) decreases for both patch types, promoting dispersal. Also the effect of increasing the relative fecundity f on conditional dispersal differs from that on unconditional dispersal: m_1^* decreases and m_2^* increases in Fig. 8a. For $f = 1$ the patch types in Figs. 8a are equal, and $m_1^* = m_2^*$. When f is increased, patch type 1 becomes more productive, increasing the benefits of staying in patch type 1. The qualitative effect of variability in V_k can, again, be different, when p is small, as illustrated by Figures S11e and S12g.

5 Discussion

This paper has explored the effect of catastrophes on the evolution of dispersal, especially the interplay between catastrophes and other demographic factors. To this end, we have constructed a fairly general model and studied the consequences of evolution for both when the dispersal probability is independent of the current patch type (unconditional dispersal) and when the dispersal probability can depend on it (conditional dispersal). We proved that in our model catastrophes promote dispersal. This is because catastrophes create empty habitats and therefore confer advantages to dispersing strategies. In our model, patches are either fully occupied or empty due to catastrophes. In models without such site-saturation, increasing the catastrophe rate may cause the total population density to decrease in such a way that most individuals reside in sparsely populated patches. In such a situation, the benefits of dispersal are diminished, and non-monotonic effects of catastrophes on dispersal evolution can be observed (Ronce et al., 2000; Gyllenberg et al., 2002; Parvinen et al., 2003; Parvinen, 2006).

We have found a surprising interplay between catastrophe probability, μ, and the survival probability in dispersal, p, especially when both μ and p are very small (Comins et al., 1980). In such a situation, the evolved dispersal probability can depend on p in a non-monotonic manner, and most importantly, it is possible that a lower chance of survival in dispersal leads to an increased dispersal probability. Moreover, we have found that the introduction of a tiny positive chance of catastrophe can have a dramatic impact on the consequences of dispersal evolution. When the dispersal survival probability p is nearly zero, introduction of catastrophes ($0 < \mu \ll 1$) has a considerable effect on dispersal evolution: dispersal evolution is highly sensitive on p and μ when they are small. In example cases shown in Figs. 1 and 6, increasing $\mu = 0$ to $\mu = 0.1$ sometimes resulted in more than double the dispersal probability. This unexpected result suggests that catastrophes can potentially have much larger impact on evolutionary outcomes than intuitively expected, and that this effect becomes particularly evident when the survival probability in dispersal, p, is very small. A small value of p corresponds, for example, to the case where each habitat is very much isolated from the others as island biogeography often assumes, or to the case where the environmental conditions are very harsh and dispersers can rarely find habitable places. An implication of our results is that in those fragmented landscapes catastrophes play a major role in determining the evolution of dispersal.

Why do very rare catastrophes matter? The interplay of catastrophes and dispersal survival probability can be investigated by looking at the components of the fitness gradient. Without catastrophes ($\mu = 0$), there should be balance among three factors; the gain by increased dispersing juveniles, the loss by decreased natal juveniles, and the gain by decreased kin-competition. When p is small, the effect of the first factor is of order p. The effects of the second and third factors are not small ($\sim O(1)$) but their leading terms cancel out because relatedness is almost one for small p, regardless of m. As a result, the sum of the three effects, which is the fitness gradient, is small ($\sim O(p)$). When we introduce catastrophes, the fraction μ of dispersing juveniles head to disturbed patches that have no natal offspring. The gain brought by these dispersing juveniles hardly depends on p because all competing dispersers suffer from the equal mortality. The other two factors still cancel out for small p and μ values. For details, see SI section F. Thus, roughly speaking, introducing a small catastrophe probability adds a new small gain ($\sim O(\mu)$) through dispersing offspring. Since the fitness gradient at $\mu = 0$ is already small, this new small component can have significant impact on the evolving dispersal probability.

Our theoretical model here takes into account rather many demographic factors, such as dispersal survival (p), finiteness of local patch (n), spatial heterogeneity (N, and n_k, f_k), arrival bias (λ_k), and catastrophes (μ). As a result, our results here contain some classical explicit expressions of singular dispersal rates as special cases (see Section 3.3.1). In that sense, we believe our analysis here offers a high degree of generality. One of the virtues of such a modeling approach is that can we spot important interplay between multiple demographics factors that models that study a single demographic factor can never detect. As a result, we have found in the current paper an important interaction between catastrophe probability and survival probability in dispersal.

It has been shown (Parvinen et al., 2023) that, without catastrophe, evolutionary branching can happen for the unconditional case, while it does not happen in the conditional case. We have here the same result when the model includes catastrophes. Spatial heterogeneity typically favors unconditional dispersal (see Fig. 4) while the effect is more complicated in the conditional case (see Fig. 8). These results are qualitatively consistent with our previous result (Parvinen et al., 2023), while dispersal from low-productivity patch can evolve with catastrophes.

The evolution of dispersal in models including catastrophes have been investigated before (Comins et al., 1980; Ronce et al., 2000; Gyllenberg et al., 2002; Parvinen et al., 2003). The fact that the ESS dispersal probability can decrease when the dispersal survival probability p increases has been numerically observed before, see Figs. 4 and 5 of Comins et al. (1980). Their model differs from ours by assuming that the number of offspring per adult is finite, whereas we have assumed infinite fecundity (and made a number of extensions, including spatial heterogeneity). Together, Comins et al. (1980) and our work show that the non-monotonic effect of the dispersal survival probability is not an artifact of the assumption of either finite or infinite fecundity. In this paper, we have explored the phenomenon in detail, and with the help of the derived analytical expressions, we have been able to investigate under which conditions this phenomenon occurs. In SI section G, we have made a corresponding analysis of the finite-fecundity model of Comins et al. (1980), when there can be at most one adult individual in a patch ($n = 1$), and showed that the phenomenon becomes more common, when fecundity is increased.

Certainly, there are some factors that our model does not cover. For example, we did not study the effect of non-equilibrium dynamics (including cyclic behavior in demography), which is occasionally caused by finiteness of fecundity. Moreover, using our model we have studied the effect of spatial heterogeneity, but not temporal variation of patch quality. It remains an open question whether our findings here generalize to those situations; we leave those aspects as future works.

To conclude, this paper have elucidated the effect of catastrophes on dispersal evolution. Through a mathematical analysis, we have revealed an intricate interplay between catastrophes and other demographic factors and found an unexpected synergy between catastrophes and survival probability in dispersal. We believe that our results here enhance our understanding of the evolution of dispersal and contribute to identifying environmental factors that play major roles there.

Declaration of interests

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests

Hisashi Ohtsuki reports financial support was provided by JSPS KAKENHI Grant Number No.20K06812. Joe Yuichiro Wakano reports financial support was

provided by JSPS KAKENHI Grant Number No.21K03357 and No.24H00001. Given his role as Editorial Board Member, H.O. had no involvement in the peer review of this article and had no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to another journal editor. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

HO acknowledges the support by JSPS KAKENHI Grant Number No.20K06812. JYW acknowledges the support by JSPS KAKENHI Grant Number No.21K03357 and No.24H00001.

References

Ajar, É. (2003). Analysis of disruptive selection in subdivided populations. *BMC Evolutionary Biology 3:22*, 1–12.

Clobert, J., E. Danchin, A. A. Dhondt, and J. D. Nichols (Eds.) (2001). *Dispersal.* Oxford University Press.

Comins, H. N., W. D. Hamilton, and R. M. May (1980). Evolutionarily stable dispersal strategies. *J. Theor. Biol. 82*, 205–230.

Dieckmann, U., B. O'Hara, and W. Weisser (1999). The evolutionary ecology of dispersal. *Trends Ecol. Evol. 14*, 88–90.

Doebeli, M. and G. D. Ruxton (1997). Evolution of dispersal rates in metapopulation models: branching and cyclic dynamics in phenotype space. *Evolution 51*, 1730–1741.

Frank, S. A. (1986). Dispersal polymorphisms in subdivided populations. *J Theor. Biol. 122*, 303–309.

Frank, S. A. (1998). *Foundations of Social Evolution.* Princeton University Press, Princeton, NJ.

Gandon, S. and Y. Michalakis (1999). Evolutionarily stable dispersal rate in a metapopulation with extinctions and kin competition. *J. Theor. Biol. 199*, 275–290.

Geritz, S. A. H., É. Kisdi, G. Meszéna, and J. A. J. Metz (1998). Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. *Evol. Ecol. 12*, 35–57.

Geritz, S. A. H., J. A. J. Metz, É. Kisdi, and G. Meszéna (1997). Dynamics of adaptation and evolutionary branching. *Phys. Rev. Lett. 78*, 2024–2027.

Gyllenberg, M. and J. A. J. Metz (2001). On fitness in structured metapopulations. *J. Math. Biol. 43*, 545–560.

Gyllenberg, M., K. Parvinen, and U. Dieckmann (2002). Evolutionary suicide and evolution of dispersal in structured metapopulations. *J. Math. Biol. 45*, 79–105.

Hamilton, W. D. and R. M. May (1977). Dispersal in stable habitats. *Nature 269*, 578–581.

Hanski, I. A. (1999). *Metapopulation Ecology.* Oxford University Press, Oxford.

Hastings, A. (1983). Can spatial variation alone lead to selection for dispersal. *Theor. Popul. Biol. 24*, 244–251.

Holt, R. D. and M. McPeek (1996). Chaotic population dynamics favors the evolution of dispersal. *Am. Nat. 148*, 709–718.

Jansen, V. A. A. and R. Vitalis (2007). The evolution of dispersal in a levins' type metapopulation model. *Evolution 61*, 2386–2397.

Johst, K., M. Doebeli, and R. Brandl (1999). Evolution of complex dynamics in spatially structured populations. *Proc. R. Soc. London B 266*, 1147–1154.

Kisdi, É. (2002). Dispersal: Risk spreading versus local adaptation. *Am. Nat. 159*, 579–596.

Levins, R. (1969). Some demographic and genetic consequenses of environmental heterogeneity for biological control. *Bull. Entomol. Soc. Am. 15*, 237–240.

Levins, R. (1970). Extinction. In M. Gerstenhaber (Ed.), *Some Mathematical Problems in Biology*, pp. 77–107. American Mathematical Society, Providence, RI.

Massol, F., A. Duputié, P. David, and P. Jarne (2011). Asymmetric patch size distribution leads to disruptive selection on dispersal. *Evolution 65*, 490–500.

Mathias, A., É. Kisdi, and I. Olivieri (2001). Divergent evolution of dispersal in a heterogeneous landscape. *Evolution 55*, 246–259.

Maynard Smith, J. (1976). Evolution and the theory of games. *Amer. Sci. 64*, 41–45.

Maynard Smith, J. and G. R. Price (1973). The logic of animal conflict. *Nature 246*, 15–18.

McPeek, M. A. and R. D. Holt (1992). The evolution of dispersal in spatially and temporally varying environments. *Am. Nat. 140*, 1010–1027.

Metz, J. A. J. and M. Gyllenberg (2001). How should we define fitness in structured metapopulation models? Including an application to the calculation of ES dispersal strategies. *Proc. R. Soc. London B 268*, 499–508.

Motro, U. (1982a). Optimal rates of dispersal. I. haploid populations. *Theor. Popul. Biol. 21*, 394–411.

Motro, U. (1982b). Optimal rates of dispersal. II. diploid populations. *Theor. Popul. Biol. 21*, 412–429.

Motro, U. (1983). Optimal rates of dispersal. III. parent offspring conflict. *Theor. Popul. Biol. 23*, 159–168.

Parvinen, K. (1999). Evolution of migration in a metapopulation. *Bull. Math. Biol. 61*, 531–550.

Parvinen, K. (2006). Evolution of dispersal in a structured metapopulation model in discrete time. *Bull. Math. Biol. 68*, 655–678.

Parvinen, K., U. Dieckmann, M. Gyllenberg, and J. A. J. Metz (2003). Evolution of dispersal in metapopulations with local density dependence and demographic stochasticity. *J. Evol. Biol 16*, 143–153.

Parvinen, K., H. Ohtsuki, and J. Y. Wakano (2018). Spatial heterogeneity and evolution of fecundity-affecting traits. *J. Theor. Biol. 454*, 190–204.

Parvinen, K., H. Ohtsuki, and J. Y. Wakano (2020). Evolution of dispersal in a spatially heterogeneous population with finite patch sizes. *Proc. Natl. Acad. Sci. USA 117*, 7290–7295.

Parvinen, K., H. Ohtsuki, and J. Y. Wakano (2023). Evolution of dispersal under spatio-temporal heterogeneity. *J. Theor. Biol. 574*, 111612.

Rodrigues, A. M. and A. Gardner (2022). Reproductive value and the evolution of altruism. *Trends in ecology & evolution 37*(4), 346–358.

Ronce, O. (2007). How does it feel to be like a rolling stone? Ten questions about dispersal evolution. *Annu. Rev. Ecol. Evol. Syst. 38*, 231–257.

Ronce, O., F. Perret, and I. Olivieri (2000). Evolutionarily stable dispersal rates do not always increase with local extinction rates. *Am. Nat. 155*, 485–496.

Taylor, P. D. (1988). An inclusive fitness model for dispersal of offspring. *J. Theor. Biol. 130*, 363–378.

Taylor, P. D. (1990). Allele-frequency change in a class-structured population. *Am. Nat. 135*, 95–106.

Van Valen, L. (1971). Group selection and the evolution of dispersal. *Evolution 25*, 591–598.

Weigang, H. C. (2017). Coevolution of patch-type dependent emigration and patch-type dependent immigration. *Journal of Theoretical Biology 426*, 140–151.

Wright, S. (1931). Evolution in mendelian populations. *Genetics 16*, 97–159.