Dispersal polymorphism in stable habitats

Eva Kisdi
Department of Mathematics and Statistics
University of Helsinki
PO Box 68, FIN-00014 Finland
email: eva.kisdi@helsinki.fi

Abstract. In fragmented but temporally stable landscapes, kin competition selects for dispersal when habitat patches are small, whereas the loss of dispersal is favoured when dispersal is costly and local populations are large enough for kin interactions to be negligible. In heterogeneous landscapes with both small and large patches, contrasting levels of kin competition facilitate the coexistence of low-dispersal and high-dispersal strategies. In this paper, I use both adaptive dynamics and inclusive fitness to analyse the evolution of dispersal in a simple model assuming that each patch supports either a single individual or a large population. With this assumption, many results can be obtained analytically. If the fraction of individuals living in small patches is below a threshold, then evolutionary branching yields two coexisting dispersal strategies. An attracting and evolutionarily stable dimorphism always exists (also when the monomorphic population does not have a branching point), and contains a strategy with zero dispersal and a strategy with dispersal probability between one half and the ESS of the classic Hamilton-May model. The present model features surprisingly rich population dynamics with multiple equilibria and unprotected dimorphisms, but the evolutionarily stable dimorphism is always protected.

1 Introduction

In a landmark paper, Hamilton and May (1977) showed that kin competition can maintain dispersal in stable habitats. The simplest version of their model assumes that the habitat consists of many small sites, each supporting just one adult individual. All offspring born in a site are therefore siblings, and the offspring who remain in their natal site kill their own siblings when they compete for establishment. Dispersal can be seen as an altruistic act; the dispersed offspring accept the cost of dispersal, such as the associated mortality risk, in order to save their siblings (Taylor 1988; Ronce 2007). If the habitat consists of larger patches that can support more than one adult, then the offspring are not all relatives and therefore kin competition weakens, resulting in less dispersal (Comins et al. 1980; Taylor 1988; Ajar 2003).

In natural habitats, the size of habitat patches is likely variable. Since differently sized patches select for different levels of dispersal, variable patch size offers an attractively economical explanation for the evolution of dispersal polymorphisms. The expectation that variable patch size can select for the diversification of dispersal strategies was shown to be correct by Massol et al. (2011), who found evolutionary branching of dispersal when patch size is distributed with positive skew. The model of Massol et al. (2011) excels at combining analytical results with biological realism (and indeed they demonstrate that their condition for evolutionary branching is satisfied in several natural systems), but is relatively technical.

In this paper, I propose a simpler model to demonstrate evolutionary branching and to explore the joint evolution of coexisting dispersal strategies under contrasting levels of kin competition. The advantage of the present model is threefold. First, it is a direct extension of the Hamilton-May model, and hence easier comparable to the classic results than the continuous-time model of Massol et al. (2011; see section 6 for details). Second, the analysis is carried beyond evolutionary branching, i.e., I investigate how two dispersal strategies coevolve. For some dimorphisms, I find multiple population dynamic equilibria such that the dimorphism is not protected and can be lost without warning, through a catastrophic bifurcation, when an environmental change induces changes in the model parameters. Third, the simplicity of the present model makes it analytically highly tractable.

In the first part of the paper, I obtain simple formulas for the invasion boundaries of pairwise invasibility plots (Geritz et al. 1998), for the evolutionarily singular strategies and their bifurcations, and also for the evolutionarily stable dimorphisms. The adaptive dynamics of this model can be analysed with simple mathematics to an almost unprecedented degree. In the second part, I use the inclusive fitness approach to obtain the monomorphic and dimorphic singularities and their convergence stability in a highly intuitive way (section 7 can be read directly after the model description). I also extend the inclusive fitness approach to derive the condition for evolutionary stability (section 8). Although the latter is technically more cumbersome, it shows explicitly how contrasting habitat sizes facilitate evolutionary branching via spatial sorting of dispersal strategies.

2 The model

I consider a population of an annual organism where juveniles compete for breeding sites. A fraction p of the breeding sites are solitary sites, whereas the remaining fraction $1-p$ of the sites form one large, well mixed patch, where all juveniles compete equally for all sites. The life cycle consists of reproduction, dispersal, and competition. At reproduction, each breeding site is occupied by one adult, who produces C offspring and dies. I assume $C \to \infty$ so that the dynamics is deterministic within each site. A fraction d of the offspring is dispersed and the remaining fraction $1-d$ remains in the natal site. Dispersed offspring survive dispersal with probability s and land either in a solitary site with probability p or in the large patch with probability $1-p$. Hence I assume that immigration is simply proportional to the area covered by the breeding sites, so that the model has no inherent source-sink structure (this is different from the model of Massol et al. (2011), who assumed that each location receives the same number of immigrants irrespectively of its carrying capacity). After dispersal, juveniles compete for a breeding site according to a fair lottery. Kin competition within solitary sites selects for dispersal (Hamilton and May 1977), whereas kin competition vanishes in the large patch and dispersal is selected against due to the survival cost associated with it.

Suppose that L dispersal strategies coexist in the population (later I shall focus on $L = 1$ and $L = 2$). Let d_k denote the fraction of offspring dispersed by strategy k $(k = 1, ..., L)$, let $n_{1k}(t)$ be the fraction of all breeding sites that is solitary and is occupied by an adult of strategy k in generation t $(\sum_{k=1}^{L} n_{1k}(t) = p)$, and let $n_{2k}(t)$ be the fraction of all sites that is in the large patch and is occupied by strategy k $(\sum_{k=1}^{L} n_{2k}(t) = 1 - p)$.

In order to describe the population dynamics concisely, I start with defining three quantities that will play the role of the environmental feedback variables (Metz et al. 1992; Diekmann et al. 2001; see below). The first environmental feedback variable is

$$E_1 = s \sum_k (n_{1k} + n_{2k}) d_k \tag{1a}$$

so that $E_1 C$ is the number of immigrants per site. If a solitary site was occupied by an adult of strategy i, then the site is won by one of the adult's non-dispersed offspring with probability

$$\frac{1 - d_i}{1 - d_i + E_1}.$$

The second environmental feedback variable is

$$E_2 = \frac{1 - p}{\sum_k n_{2k}(1 - d_k) + (1 - p)E_1} \tag{1b}$$

so that E_2/C is the probability that a specific juvenile secures a breeding site if it is competing in the large patch. To interpret this quantity, note that all juveniles in the large patch compete for $(1 - p)N$ sites (where N is the total number of breeding sites), $\sum_k n_{2k}NC(1 - d_k)$ is the number of non-dispersed offspring produced in the large patch and $(1 - p)NE_1C$ is the number of immigrants into the large patch (this is $(1 - p)N$ times what a solitary site receives because the large patch covers $(1 - p)N$ times the area of a single site). The probability of success is

the number of sites divided with the number of competitors. Finally, the third environmental feedback variable is defined as

$$E_3 = \sum_k \frac{n_{1k}}{p} \cdot \frac{1}{1 - d_k + E_1} \tag{1c}$$

so that E_3/C is the probability that a specific immigrant offspring who landed in a solitary site wins the site. Here $\frac{n_{1k}}{p}$ is the probability of landing in a site occupied by an adult of strategy k conditioned on landing in a solitary site, and $(1 - d_k + E_1)C$ is the number of competitors within the site.

Let D denote an arbitrary dispersal strategy (either one of the coexisting L strategies or a new mutant) and let the vector $\mathbf{m}(t) = (m_1(t), m_2(t))^T$ be the population vector of this strategy, containing the fraction of all sites that are occupied with an adult of this strategy and are solitary (m_1) or in the large patch (m_2). Hence if D stands for the kth strategy present in the resident population, then $D = d_k$ and $\mathbf{m} = (n_{1k}, n_{2k})^T$. The projection matrix $\mathbf{A}$, which determines the population vector in the next generation according to $\mathbf{m}(t+1) = \mathbf{A}\mathbf{m}(t)$, is given by

$$\mathbf{A} = \begin{bmatrix} \frac{1-D}{1-D+E_1} + DspE_3 & DspE_3 \\ \\ Ds(1-p)E_2 & (1 - D + Ds(1-p))E_2 \end{bmatrix} \tag{2}$$

Here the element A_{11} is the number of solitary sites won by the offspring of a parent currently in a solitary site, given by the sum of the probability that a solitary site is won by a non-dispersed offspring (first term of A_{11}) and the number of offspring dispersed from a solitary site who survive dispersal, land in a solitary site, and win that site (second term of A_{11}). A_{12} is the number of offspring dispersed from a site in the large patch who survive dispersal, land in a solitary site, and win that site. Similarly, A_{21} is the number of offspring dispersed from a solitary site who survive dispersal, land in the large patch, and secure a site in the large patch. Finally, A_{22} is the number of offspring born in a site of the large patch who either did not disperse or survived dispersal and landed back in the large patch, times the probability of securing a site in the large patch.

Notice that the elements of $\mathbf{A}$ are not constants because the the environmental feedback variables E_1, E_2, E_3 depend on the variables $n_{ik}(t)$. However, all nonlinearities in the population dynamics act through the environmental feedback variables, i.e., the three environmental feedback variables fully describe all biotic interactions between an individual and the resident population (Metz et al. 1992; Diekmann et al. 2001). When the resident population has equilibrated so that the densities n_{ik} are constants, then the dynamics of a mutant is linear as long as its own density is negligible.

$\mathbf{A}$ is irreducible if its off-diagonal elements are positive, which also immediately implies that all elements are positive and therefore $\mathbf{A}$ is primitive. To ensure irreducibility, I assume $0 < s \le 1$ and $0 < p < 1$ throughout. $D > 0$ is also necessary for irreducibility, and therefore zero dispersal requires special attention (see below).

4

3 Adaptive dynamics in monomorphic resident populations

With a single resident strategy ($L = 1$), the environmental feedback variables in equations (1) simplify to

$$E_1 = sd \text{ and } E_2 = E_3 = 1/(1 - d + sd) \tag{3}$$

where $d = d_1$ denotes the dispersal strategy of the only resident. E_2 and E_3 are equal because with a single resident, competition is equally strong in solitary sites and in the large patch (there is no source-sink structure), and therefore a focal individual has equal chance of success when competing in a solitary site occupied by the resident or in the large patch. (This will not be the case in polymorphic populations, where different resident strategies are present at different frequencies in the solitary sites versus in the large patch.)

3.1 Invasion boundaries

Substituting the environmental feedbacks in (3) into (2), the projection matrix of a rare mutant strategy d_{mut} simplifies to

$$\mathbf{A}_{mut} = \begin{bmatrix} \frac{1-d_{mut}}{1-d_{mut}+sd} + \frac{d_{mut}sp}{1-d+sd} & \frac{d_{mut}sp}{1-d+sd} \\[2mm] \frac{d_{mut}s(1-p)}{1-d+sd} & \frac{1-d_{mut}+d_{mut}s(1-p)}{1-d+sd} \end{bmatrix} \tag{4}$$

The invasion fitness, λ_{mut}, is the leading eigenvalue of this primitive matrix, and the mutant invades (subject to demographic stochasticity while present in small numbers) when $\lambda_{mut} > 1$. Instead of using the invasion fitness directly, I use the fitness proxy

$$F(d_{mut}, d) = \text{tr}(\mathbf{A}_{mut}) - \det(\mathbf{A}_{mut}) \tag{5}$$

where tr and det denote the trace and the determinant, respectively. For 2×2 non-negative projection matrices, this fitness proxy has the following property:

$$\text{if } \text{tr}(\mathbf{A}_{mut}) \leq 2 \text{ then } F(d_{mut}, d) \gtreqqless 1 \iff \lambda_{mut} \gtreqqless 1$$

$$\text{if } \text{tr}(\mathbf{A}_{mut}) > 2 \text{ then } \lambda_{mut} > 1 \text{ irrespectively of } F(d_{mut}, d) \tag{6}$$

(Metz and Leimar 2011; Appendix A). The invasion boundary, which is the $\lambda_{mut} = 1$ contour line of the invasion fitness, therefore consists of the points (d, d_{mut}) such that $F(d_{mut}, d) = 1$ and $\text{tr}(\mathbf{A}_{mut}) \leq 2$.

After some algebra (see Appendix B), $F(d_{mut}, d) = 1$ can be rearranged into

$$(d_{mut} - d)\left[pd_{mut}(1 - d_{mut}) - (1 - s)d\big(1 - (1 - s)d\big)\right] = 0 \tag{7}$$

This equation is satisfied on the line $d_{mut} = d$ and on the East-West opening hyperbola

$$\left(\frac{d - A}{\alpha}\right)^2 - \left(\frac{d_{mut} - B}{\beta}\right)^2 = 1 \tag{8a}$$

5

with center point

$$(A, B) = \left(\frac{1}{2(1-s)}, \frac{1}{2} \right) \tag{8b}$$

and

$$\alpha = \frac{\sqrt{1-p}}{2(1-s)}, \quad \beta = \frac{1}{2}\sqrt{\frac{1-p}{p}}. \tag{8c}$$

Since $0 \leq d, d_{mut} \leq 1$ by definition, the intersection of this hyperbola with the unit square is of interest. The left branch of the hyperbola always goes through the points $(d, d_{mut}) = (0,0)$ and $(0,1)$. Depending on whether the left vertex (i.e., the rightmost point of the left branch), $(d, d_{mut}) = (A - \alpha, B)$, is or is not in the unit square, the left branch of the hyperbola gives one or two line segments in the unit square that satisfy $F(d_{mut}, d) = 1$. A part of the right branch of the hyperbola is in the unit square whenever the right vertex $(A + \alpha, B)$ is in there.

In order to see which parts of the $F(d_{mut}, d) = 1$ line segments satisfy the requirement $\text{tr}(\mathbf{A}_{mut}) \leq 2$, note first that $\text{tr}(\mathbf{A}_{mut})$ is a continuous function of d_{mut} and of d in $(0,1]^2$. Hence by varying d_{mut} and d continuously, $\text{tr}(\mathbf{A}_{mut}) - 2$ can switch sign only by passing through a point where $\text{tr}(\mathbf{A}_{mut}) = 2$. Suppose that such a point is encountered while d_{mut} and d vary along a line where $F(d_{mut}, d) = \text{tr}(\mathbf{A}_{mut}) - \det(\mathbf{A}_{mut}) = 1$. $\text{tr}(\mathbf{A}_{mut}) = 2$ then immediately implies $\det(\mathbf{A}_{mut}) = 1$. Recall that the characteristic equation of a 2×2 matrix is $\lambda^2 - \text{tr}(\mathbf{A})\lambda + \det(\mathbf{A}) = 0$, which then simplifies to $\lambda^2 - 2\lambda + 1 = 0$, so that the only eigenvalue is a double root at $\lambda = 1$. However, by the Perron-Frobenius theorem, this contradicts the fact that $\mathbf{A}_{mut}$ is primitive for all $(d, d_{mut}) \in (0,1]^2$. It therefore follows that $\text{tr}(\mathbf{A}_{mut}) - 2$ cannot switch sign along a continuous line in $(0,1]^2$ with $F(d_{mut}, d) = 1$. If $\text{tr}(\mathbf{A}_{mut}) < 2$ holds at one point of the line, then it holds at all points; and the same is true for the converse. In Appendix B, I use this fact to show that $\text{tr}(\mathbf{A}_{mut}) < 2$ holds on the left branch of the hyperbola in the interior of the unit square, but not on the right branch (should it appear in the unit square). It is straightforward to see that $\text{tr}(\mathbf{A}_{mut}) < 2$ holds also on the line $d_{mut} = d$. Hence the invasion boundaries are given by the left branch of the hyperbola in (8) and the line $d_{mut} = d$.

3.2 Evolutionary singularities

With the invasion boundaries derived above, pairwise invasibility plots can readily be drawn (Figure 1). Invasion ("+") and non-invasion ("−") areas are separated by the main diagonal $d_{mut} = d$ and by the left branch of the hyperbola in (8). Since $F(\frac{1}{2}, 0) = 1 + \frac{sp}{4} > 1$, the left edge of the pairwise invasibility plot is always in the "+" area, which implies the signs shown in Figure 1.

The pairwise invasibility plot can have only three qualitatively different shapes: either there is no interior singularity (Figure 1a), or there is an evolutionary branching point (Figure 1b), or there is an attracting ESS (Figure 1c). The vertex of the hyperbola may also lay beyond $d = 1$, but this means no qualitative change from Figure 1c.
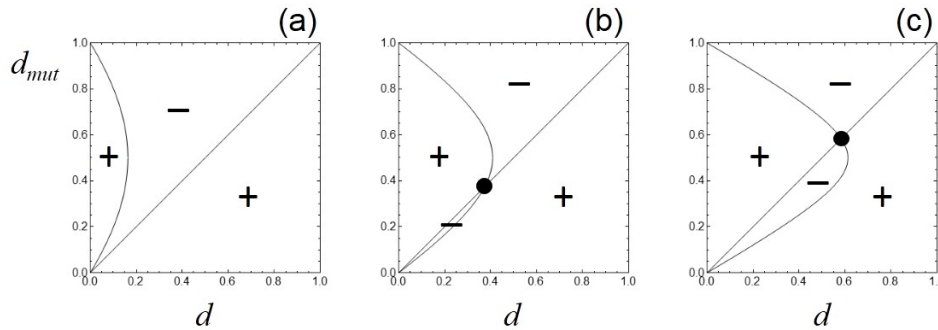
6

Figure 1: Pairwise invasibility plots. The invasion boundaries separate the "+" (invasion) and "−" (non-invasion) areas. The evolutionary singularity is marked with a dot. $s = 0.5$ and (a) $p = 0.3$, dispersal evolves towards zero and undergoes braching in its vicinity; (b) $p = 0.65$, the singularity is an evolutionary branching point; (c) $p = 0.85$, the singularity is a convergence stable ESS.

The evolutionarily singular dispersal strategy d^* is at the interior intersection of the left branch of the hyperbola and the main diagonal. Since it is easily shown that only the left branch of the hyperbola can intersect the diagonal in the unit square, d^* is obtained straightforwardly from (7) as the nonzero root of

$$pd(1 - d) = (1 - s)d\left(1 - (1 - s)d\right)$$

i.e.,

$$d^* = \frac{p - (1 - s)}{p - (1 - s)^2} \tag{9}$$

$0 < d^* \leq 1$ holds and hence an interior singularity exists if and only if $p > 1 - s$. When it exists, the interior singularity is always convergence stable (cf. the sign structure of the pairwise invasibility plots in Figure 1). The hyperbola intersects the main diagonal with a positive slope and therefore the interior singularity is an evolutionary branching point when the intersection is below the vertex, i.e., when $d^* < \frac{1}{2}$; from (9), this is equivalent to $p < 1 - s^2$.

When $p < 1 - s$, there is no interior singularity and dispersal evolves towards zero in monomorphic populations. As Figure 1a illustrates, $d = 0$ is not evolutionarily stable but resembles an evolutionary branching point. With zero dispersal, however, the projection matrix $\mathbf{A}$ is not irreducible, and hence the invasion fitness of the mutant is not well defined; moreover, the leading eigenvalue of the mutant projection matrix is not differentiable at $d_{mut} = d = 0$ so that the selection gradient is not defined. Therefore $d = 0$ is not a regular evolutionarily singular strategy. Yet evolutionary branching happens in the vicinity of $d = 0$. This is because the population becomes dimorphic as soon as d is small enough for a small mutation to produce d_{mut} above the hyperbola. Once the population is dimorphic, evolutionary branching takes off as it does at an interior branching point (Geritz et al. 1998).

In summary, dispersal evolves to an ESS when $p > 1 - s^2$, i.e., when the fraction of solitary sites is high so that the model is sufficiently close to the Hamilton-May (1977) model. When

$p < 1-s$, selection in the large patch dominates and dispersal evolves towards zero and undergoes evolutionary branching near $d = 0$. Inbetween these thresholds, there is an interior evolutionary branching point (Figure 2).
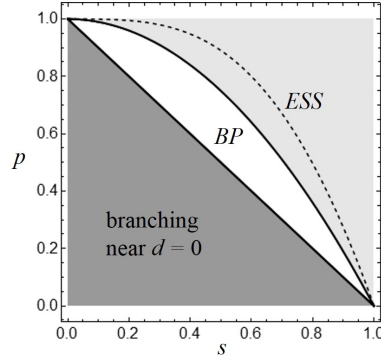

Figure 2: Bifurcations of the monomorphic singularity. Dark grey: no interior singularity, dispersal evolves towards zero and undergoes branching in its vicinity; white: the singularity is an evolutionary branching point (BP); light grey: the singularity is a convergence stable ESS. M (defined in section 4.2) is negative below the dashed line and positive above.

4   Dimorphic resident populations: Analytical results

Due to the relative simplicity of the model, important results can be obtained analytically also in dimorphic resident populations. In particular, here I prove that (i) there is a convergence stable and evolutionarily stable dimorphism where one of the two residents has zero dispersal, and (ii) the model exhibits unprotected dimorphisms with multiple population dynamical attractors.

4.1   Evolutionarily stable dimorphism

Consider a resident population with two dispersal strategies $(d_1, d_2) = (0, d_2)$, i.e., where one of the residents does not disperse but the other resident's strategy $d_2 > 0$ is arbitrary. From (2), the population projection matrix of the nondispersing resident $D = d_1 = 0$ is

$$
\mathbf{A} = \begin{bmatrix} \frac{1}{1+E_1} & 0 \\ 0 & E_2 \end{bmatrix}
\tag{10}
$$

When the resident population is at equilibrium, the leading eigenvalue of this matrix must be 1. The two eigenvalues are $\frac{1}{1+\hat{E}_1}$ and $\hat{E}_2$ (where the hat denotes the equilibrium values), and since $\hat{E}_1 > 0$ (cf. (1a)), the former is necessarily less than 1. It thus follows that $\hat{E}_2 = 1$ must hold. The corresponding eigenvector $\begin{pmatrix} \hat{n}_{11} \\ \hat{n}_{21} \end{pmatrix} = \begin{pmatrix} 0 \\ \hat{n}_{21} \end{pmatrix}$ indicates that in equilibrium, the nondispersing

8

resident is absent from the solitary sites. As expected, the nondispersing resident is confined to the large patch, within which it is favoured by natural selection due to the mortality cost of dispersal. Note that the matrix in (10) is reducible, so that this equilibrium is reached only if the nondispersing resident is initially present in the large patch; in the opposite case, it would go extinct.

Since $\hat{n}_{11} + \hat{n}_{12} = p$ and $\hat{n}_{21} + \hat{n}_{22} = 1 - p$, $\hat{n}_{11} = 0$ implies $\hat{n}_{12} = p$ and $\hat{E}_2$ reduces to $\hat{E}_2 = \frac{1-p}{1-p-\hat{n}_{22}d_2+(1-p)\hat{E}_1}$ with $\hat{E}_1 = s(p+\hat{n}_{22})d_2$. The equilibrium condition $\hat{E}_2 = 1$ is therefore a linear equation for the unknown $\hat{n}_{22}$, which is readily solved to obtain the equilibrium densities

$$\hat{n}_{11} = 0, \quad \hat{n}_{12} = p, \quad \hat{n}_{21} = \frac{(1-s)(1-p)}{1-s(1-p)}, \quad \hat{n}_{22} = \frac{sp(1-p)}{1-s(1-p)}. \tag{11}$$

Interestingly, the equilibrium densities do not depend on d_2 even in the large patch. The higher d_2 is, the more strongly this strategy is selected against in the large patch, but at the same time the more influx it gets from the solitary sites, and these two effects cancel each other. The population dynamic equilibrium is linearly stable provided $s < 1$. At the equilibrium, the environmental feedback variables reduce to

$$\hat{E}_1 = \frac{spd_2}{1-s(1-p)}, \quad \hat{E}_2 = 1, \quad \hat{E}_3 = \frac{1}{1-d_2+\hat{E}_1}. \tag{12}$$

The projection matrix $\mathbf{A}_{mut}$ of a mutant strategy d_{mut} in the resident population of strategies $(0, d_2)$ is the matrix in (2) with $D = d_{mut}$ and with the environmental feedback variables as given in (12). Let $F_2(d_{mut}, 0, d_2) = \text{tr}(\mathbf{A}_{mut}) - \det(\mathbf{A}_{mut})$ denote the fitness proxy of a mutant in a resident population of strategies $d_1 = 0$ and d_2. F_2 relates to the invasion fitness of the mutant as shown in (6). However, below I shall consider only mutants that have similar dispersal strategies to one of the residents, and this simplifies the use of F_2. Since for the residents $0 < \text{tr}(\mathbf{A}) < 2$ always holds at equilibrium, by continuity the same is true for $\text{tr}(\mathbf{A}_{mut})$ when the mutant strategy is sufficiently close to a resident; hence F_2 can be used in place of the invasion fitness for mutations of small effect without separately checking $\text{tr}(\mathbf{A}_{mut})$ (Metz and Leimar 2011).

Straightforward calculations (Appendix C) show that the selection gradients of the two residents are given by

$$\partial_1 F_2(0,0,d_2) = -(1-s(1-p))\frac{\hat{E}_1}{1+\hat{E}_1} \tag{13a}$$

$$\partial_1 F_2(d_2,0,d_2) = \frac{spd_2}{(1-d_2+\hat{E}_1)^2}h(d_2) \tag{13b}$$

where the notation $\partial_i f(x_1, x_2, x_3)$ is used for the partial derivative of f with respect to x_i evaluated at (x_1, x_2, x_3) and

$$h(d_2) = 1 - d_2\frac{2(1-s)+sp}{1-s(1-p)}.$$

Since $\partial_1 F_2(0,0,d_2)$ in (13a) is always negative, a mutant of the nondispersing resident can never invade the dimorphism $(0, d_2)$ and the boundary $d_1 = 0$ is attracting for any d_2. A resident

9

strategy with low dispersal is favoured in the large patch (when $s < 1$), but is present also in the solitary sites, where kin competition selects against it. A mutant with even lower dispersal is increasingly confined to the large patch, where it is even more favoured. Hence the low dispersal resident of the dimorphism evolves towards full loss of dispersal.

The selection gradient of the second strategy vanishes when $h(d_2) = 0$, i.e., at

$$d_2^* = \frac{1 - s + sp}{2(1 - s) + sp} \tag{14}$$

Since $h(d_2)$ is a strictly decreasing function of d_2, d_2^* is attracting on the boundary $(d_1, d_2) = (0, d_2)$ and therefore the dimorphism $(0, d_2^*)$ is absolutely convergence stable (Matessi and di Pascuale 1996, Leimar 2009). Finally, the dimorphism $(0, d_2^*)$ is also evolutionarily stable because the second derivative

$$\partial_{11} F_2(d_2^*, 0, d_2^*) = -\frac{2sp}{d_2^*} \tag{15}$$

is always negative (see Appendix C).

The present model thus always has an evolutionarily stable dimorphism where one of the residents is not dispersing at all and is confined to the large patch, whereas the other resident disperses with probability d_2^* given in (14) and thereby partly specializes on the solitary sites. It is straightforward to show that $d_2^* > d^*$ (unless $s = 1$, in which case both d_2^* and d^* equal 1), i.e., the dispersing resident of the dimorphism has a higher dispersal probability than the generalist monomorphic singularity. On the other hand, d_2^* is less than $\frac{1}{2-s}$, the ESS of the Hamilton-May model (again unless $s = 1$, in which case both equal 1). The presence of the large patch and of the nondispersing resident has two consequences for the dispersing resident: first, a solitary site receives less immigrants, which implies stronger kin competition and selects for more dispersal; second, dispersal into the large patch, where the dispersing resident is at a disadvantage, selects against dispersal. As evidenced by $d_2^* < \frac{1}{2-s}$, the second mechanism is stronger. Note that as $p \to 1$ (almost all sites are solitary), $d_2^* \to \frac{1}{2-s}$. If $p \to 0$ while $s < 1$, then $d_2^* \to \frac{1}{2}$, which is the ESS of the Hamilton-May model in the limit when all dispersed offspring perish. In the present model, $p \to 0$ implies that an offspring dispersed away from a solitary site will almost surely land in the large patch, and its descendants have a vanishing chance of returning to a solitary site. Since with $s < 1$ the dispersing resident is selected against in the large patch, an offspring dispersed away from a solitary site will have no descendants in the distant future. At the same time, $\hat{E}_1$ goes to 0, meaning that a solitary site receives no immigrants. The situation of no descendants from dispersed offspring and no immigrants in a solitary site is analogous to the limit $s \to 0$ in the Hamilton-May model.

The evolutionarily stable dimorphism $(0, d_2^*)$ exists and is locally attracting whether or not the monomorphic population has an evolutionary branching point. This implies an evolutionary bistability: when the monomorphic singularity is an ESS, some initially dimorphic populations evolve to the dimorphic ESS whereas others lose the dimorphism and eventually evolve to the monomorphic ESS (Geritz et al. 1999; see an example in Section 5).

4.2 Unprotected dimorphisms

Despite its simplicity, the present model exhibits alternative stable equilibria of population dynamics in dimorphic populations where neither resident has zero dispersal. These alternative stable equilibria correspond to unprotected dimorphisms: at one stable equilibrium all four population densities $\hat{n}_{11}, \hat{n}_{12}, \hat{n}_{21}, \hat{n}_{22}$ are positive, whereas at the other stable equilibrium one of the two resident strategies is extinct. The next section will illustrate this with examples. To prove analytically that alternative stable equilibria of n_{ij} exist, I use a result of Priklopil (2012). Let

$$M = \left. \frac{\partial^2 F(d_{mut}, d)}{\partial d_{mut} \partial d} \right|_{d_{mut}=d=d^*} \tag{16}$$

where F is the monomorphic fitness proxy and d^* is the monomorphic singularity (see Section 3). M characterises the existence of protected dimorphisms in the neighbourhood of a monomorphic singularity: if $M < 0$, then the neighbourhood of (d^*, d^*) contains strategy pairs that mutually invade each other's monomorphic population, whereas if $M > 0$, the neighbourhood of the monomorphic singularity contains pairs that mutually exclude each other (Geritz et al. 1998). Priklopil (2012) showed that if M changes sign in the interior of the parameter space, then there are parameter values in the neighbourhood of this sign change for which the model exhibits unprotected dimorphisms.

In the present model, M simplifies to

$$M = \frac{1}{p} \left(\frac{p - (1-s)^2}{sp} \right)^2 \left[p^2 - 3s(1-s)p - (1-s)^3 \right]$$

Since $p > 1 - s$ is necessary for an interior singularity to exist (see Section 3), the factor in front of the brackets is positive. M vanishes when

$$p = \frac{1-s}{2} \left(3s + \sqrt{9s^2 - 4s + 4} \right)$$

(the other root of the brackets is negative for $s < 1$). This curve is shown by the dotted line in Figure 2. The non-degeneracy condition of Priklopil (2012) is satisfied by the hyperbolic invasion boundary derived in Section 3. Since singularities with $M = 0$ exist in the interior of the parameter space and M changes sign there, the theorems of Priklopil (2012) ensure the existence of unprotected dimorphisms.

5 Numerical results

In dimorphic resident populations of arbitrary dispersal strategies (d_1, d_2), the equilibrium densities are hard to obtain analytically (equation (11) is valid only for the special case when $d_1 = 0$). I have determined the resident densities by numerical continuation of equilibria (see e.g. Kuznetsov 1995). Once the resident densities $\hat{n}_{ij}$ $(i, j = 1, 2)$ are known, the selection gradients $G_1(d_1, d_2) = \partial_1 F_2(d_1, d_1, d_2)$ and $G_2(d_1, d_2) = \partial_1 F_2(d_2, d_1, d_2)$ are easily obtained from the fitness proxy $F_2(d_{mut}, d_1, d_2) = \text{tr}(\mathbf{A}_{mut}) - \det(\mathbf{A}_{mut})$, where $\mathbf{A}_{mut}$ is the projection matrix in

(2) with $D = d_{mut}$ and the environmental feedbacks assuming their equilibrium values obtained by substituting $\hat{n}_{ij}$ into (1).

Figure 3 shows the isoclines of the selection gradients, i.e., the lines along which $G_1(d_1, d_2) = 0$ and $G_2(d_1, d_2) = 0$ (green and blue dashed lines, respectively). Since the isoclines do not intersect, there is no interior dimorphic singularity in these examples. The filled dots at the points where the isoclines connect to the boundary mark the boundary singularity $(0, d_2^*)$ found analytically in section 4.1 and its symmetric pair $(d_2^*, 0)$ (since the labelling of the resident strategies is arbitrary, the plots of Figure 3 are symmetric about the main diagonal). An exhaustive numerical search (Appendix D) found no interior singularity for any parameter combination, hence the boundary singularity appears to be the only dimorphic singularity in this model.
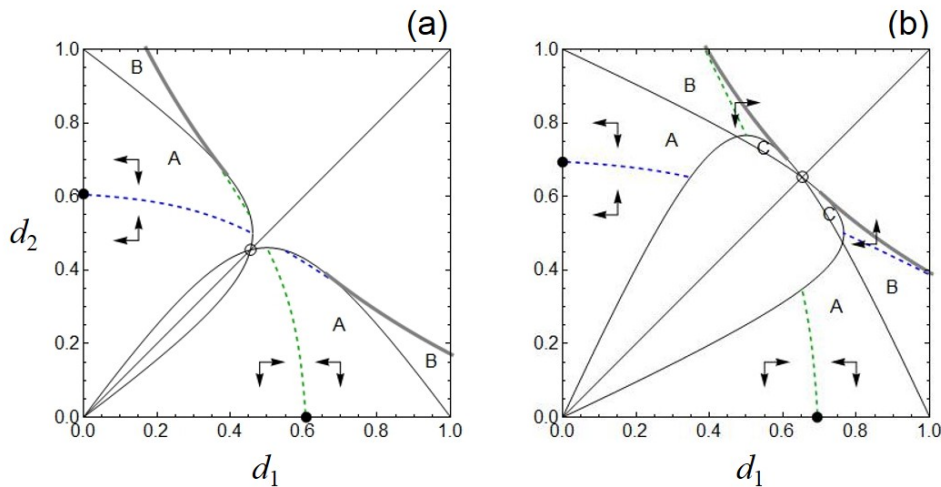

Figure 3: Examples for evolution in dimorphic populations. Black lines: invasion boundaries and their mirror images along the diagonal $d_2 = d_1$; thick grey lines: saddle-node bifurcation of population dynamic equilibria. Area marked with "A": protected dimorphisms; "B": unprotected dimorphisms; "C": mutual exclusion (panel (b) only). The monomorphic and dimorphic singularities are marked with empty and filled dots, respectively. The green dashed line is the d_1-isocline ($G_1(d_1, d_2) = 0$), the blue dashed line is the d_2-isocline ($G_2(d_1, d_2) = 0$). The arrows show the direction of evolution as determined by the sign of the selection gradients. The symmetry on the diagonal is due to the arbitrary labelling of the two resident strategies as d_1 and d_2. Parameter values: (a) $s = 0.4, p = 0.8$; (b) $s = 0.6, p = 0.85$.

Figure 3 also illustrates the unprotected dimorphisms the existence of which was proven in section 4.2. The black lines in Figure 3 are the invasion boundaries with their mirror images along the diagonal $d_2 = d_1$, so that in the area marked with "A", the two resident strategies can mutually invade each other (protected dimorphisms; cf. Geritz et al. 1998). The invasion boundary between areas "A" and "B" is a line of backward transcritical bifurcation of equilibrium densities (Boldin 2006), whereby an unstable interior equilibrium appears and an extinction equilibrium becomes stable. In area "B", the two residents coexist at a locally stable interior equilibrium but this dimorphism is unprotected. The thick grey line is a saddle-node

12

bifurcation line, where the two interior equilibria of population dynamics collide and disappear.

In Figure 3a, the parameters are chosen such that the monomorphic evolutionary singularity is a branching point. After evolutionary branching, the two residents evolve to the boundary singularity $(0, d_2^*)$ with d_2^* given in (14). The same happens also when the model has no interior monomorphic singularity and branching occurs near $d = 0$ (not shown). In Figure 3b, the monomorphic singularity is an ESS. The parameters are chosen such that in the area adjacent to the ESS, marked with "C", M is positive, which means that neither strategy can invade the monomorphic population of the other (mutual exclusion; Geritz et al. 1998). Note that even though there is a monomorphic ESS, the dimorphic singularity on the boundary still exists and is locally attracting and evolutionarily stable.

When the model parameters s and p are varied, the invasion boundaries and the saddle-node bifurcation line in Figure 3 shift, such that a given dimorphism (d_1, d_2) exhibits different population dynamics. Two examples are shown in Figure 4, which plots the bifurcation diagram of population dynamics varying s but keeping (d_1, d_2) fixed. In both cases, decreasing s leads to a saddle-node bifurcation (i.e., in Figure 3, decreasing s causes the saddle-node bifurcation line shift past the point (d_1, d_2)). This means a catastrophic collapse of the dimorphism, where strategy d_1 goes extinct abruptly, without warning from its equilibrium population density declining towards zero. In Figure 4a, increasing survival first turns the unprotected dimorphism into a protected one (in Figure 3, the invasion boundary shifts such that the point (d_1, d_2) gets from area B into area A), and eventually leads to the loss of dimorphism via a transcritical (non-catastrophic) bifurcation. In Figure 4b, increasing survival leads to mutual exclusion (in Figure 3, the invasion boundary shifts such that the point (d_1, d_2) gets from area B into area C).

6 Limit to the Hamilton-May model

If there are only solitary sites and no large patch, as in the model of Hamilton and May (1977), then the projection matrix in (2) reduces to the element A_{11} with $p = 1$. The invasion fitness of the mutant strategy $D = d_{mut}$ is thus given by

$$\lambda_{mut} = \frac{1 - d_{mut}}{1 - d_{mut} + E_1} + d_{mut} s E_3 \tag{17}$$

and in the monomorphic resident population of strategy d, the environmental feedbacks in (1) take the values

$$E_1 = sd \quad \text{and} \quad E_3 = \frac{1}{1 - d + E_1}$$

(note that E_2 does not appear when there is no large patch). The invasion boundary $\lambda_{mut} = 1$ is given by two straight lines, $d_{mut} = d$ and

$$d_{mut} = 1 - (1 - s)d, \tag{18}$$

the intersection of which yields the well-known ESS of the Hamilton-May model, $d_{HM}^* = \frac{1}{2-s}$. In the present model, the monomorphic singularity d^* given in (9) goes to d_{HM}^* as $p \to 1$, and
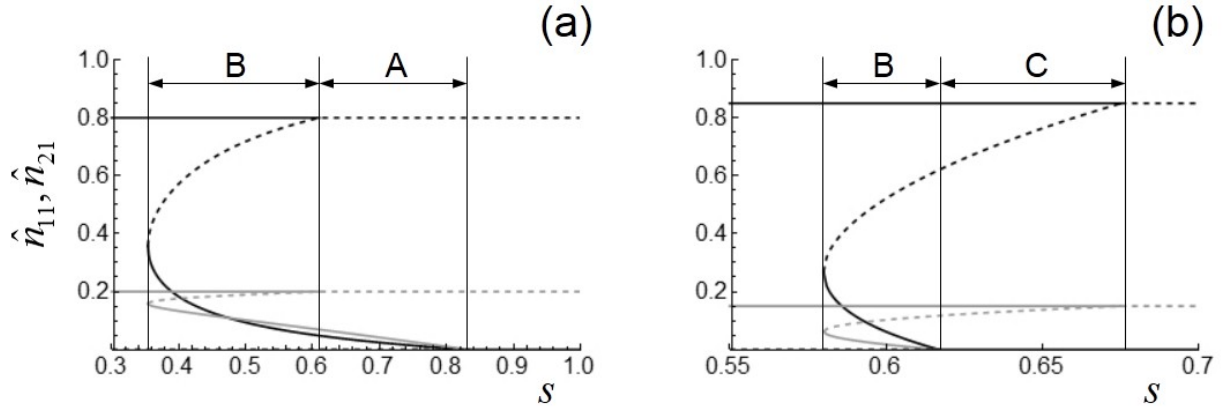
Figure 4: Bifurcations of population dynamic equilibria in dimorphic resident populations for (a) $p = 0.8$ and $(d_1, d_2) = (0.2, 0.9)$; (b) $p = 0.85$ and $(d_1, d_2) = (0.5, 0.8)$. The equilibrium densities $\hat{n}_{11}$ and $\hat{n}_{21}$ are shown with black and grey lines, respectively; the remaining densities are $\hat{n}_{12} = p - \hat{n}_{11}$ and $\hat{n}_{22} = 1 - p - \hat{n}_{21}$. Continuous and dashed lines denote stable and unstable equilibria, respectively. In the leftmost part of both panels there is one stable equilibrium (one pair of continuous black and grey lines) with equilibrium densities $\hat{n}_{11} = p$ and $\hat{n}_{21} = 1 - p$; this represents a stable boundary equilibrium where strategy d_2 is extinct ($\hat{n}_{12} = \hat{n}_{22} = 0$). The boundary equilibrium where strategy d_1 is extinct (both densities shown are zero) is unstable. In part B, a new pair of equilibria appears via a saddle-node bifurcation; the pair of black and grey dashed lines shows the saddle and the new pair of black and grey continuous lines shows the node born in the bifurcation. In panel (a), the right edge of part B marks a transcritical bifurcation where the saddle collides with the boundary equilibrium where d_2 is extinct, and the right edge of part A corresponds to another transcritical bifurcation where the interior node collides with the boundary equilibrium where d_1 is extinct. In panel (b), the two transcritical bifurcations occur in the opposite order. In the rightmost part of both panels, the boundary equilibrium with d_2 extinct is unstable and the boundary equilibrium with d_1 extinct is stable. The letters A, B, C mark protected dimorphism, unprotected dimorphism, and mutual exclusion, respectively (cf. Figure 3).

it is indeed an ESS in the limit for any $s > 0$ (cf. section 3.2).

The hyperbola-shaped invasion boundary in (7), however, does not converge to the straight line in (18), and therefore the pairwise invasibility plot of the present model remains different from the pairwise invasibility plot of the Hamilton-May model as $p \to 1$ (Figure 5). The reason for this is that the present model assumes infinitely many sites in the large patch even as the relative size of the large patch, $1 - p$, goes to zero. A mutant with sufficiently small dispersal can invade because it is largely confined to the large patch, where low dispersal is advantageous. Even though the relative frequency of the invading mutant remains infinitesimal (since the large patch itself contains only an infinitesimal fraction of the sites when $p \to 1$), the absolute number of mutants increases exponentially. In other words, the present model assumes that N, the total number of sites, goes to infinity faster than p goes to 1, so that $(1 - p)N$ remains infinite;

the Hamilton-May model however assumes that $(1-p)N$ is zero. Since the limits $N \to \infty$ and $p \to 1$ do not commute, the present model is different from the Hamilton-May model when $p \to 1$.

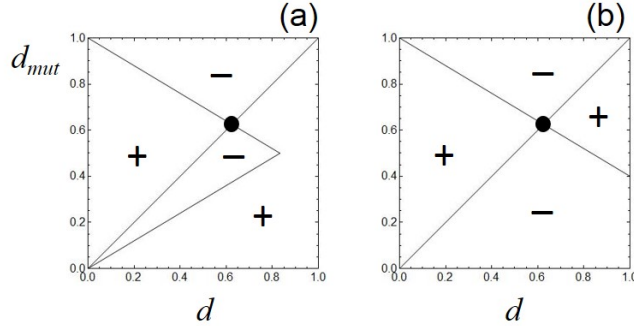
Figure 5: (a) Pairwise invasibility plot of the present model with $p \to 1$; (b) pairwise invasibility plot of the Hamilton-May model ($s = 0.4$ in both panels). Notations as in Figure 1.

In the present model, the evolutionarily stable dimorphism found in section 4.1 exists also in the limit $p \to 1$, and the dispersing strategy of this dimorphism, d_2^* given in (14), goes to d_{HM}^* as $p \to 1$. When almost all sites are solitary, the dispersing strategy evolves as in the Hamilton-May model, but the nondispersing strategy is still present at an infinitesimal frequency (yet in a large number). The evolutionarily stable dimorphism is in the area of protected dimorphisms. Using Theorem 1 of Priklopil (2012), it is straightforward to show that also unprotected dimorphisms exist in the limit $p \to 1$. To the contrary, in the Hamilton-May model there is neither protected nor unprotected dimorphism, only an area of mutual exclusion (Motro 1982; Appendix E).

7   Inclusive fitness approach: Singularities and convergence stability

In the remainder of this paper, I use an alternative approach based on inclusive fitness to derive the monomorphic and dimorphic singularities and their stability properties (see also Ajar (2003) and Parvinen et al. (in prep.) on the comparison of adaptive dynamics and inclusive fitness). In this section, I follow the classical inclusive fitness argument (most closely I follow Taylor 1988), which is suitable to determine evolutionary singularities ("candidate ESSs") and their convergence stability. I consider evolutionary stability using inclusive fitness in the next section.

By Hamilton's (1963) rule, kin selection favours dispersal if

$$rb > c \tag{19}$$

where b is the benefit others receive from the focal individual leaving her natal site, c is the cost suffered by the focal individual, and r is the relatedness between the focal individual and others competing for her natal site. Focus attention first on a solitary site in a monomorphic

population with dispersal d. If a focal individual stays in this site rather than disperses, she wins the site, and thereby kills all others in the site, with probability

$$P = \frac{1}{[1 - d + sd]C}$$

If the focal individual leaves the site, then she surrenders her chance to win this site, and the benefit to all others competing for the site is the chance forfeited by the focal individual, $b = P$. The cost to the focal individual is P minus the chance that she survives dispersal and wins another site. If she arrives at a solitary site, then the chance of winning is P; if she arrives at a site in the large patch, then the chance of winning, from (1b), is $\frac{(1-p)}{[(1-p)(1-d)+(1-p)sd]C} = P$. The probabilities of winning a solitary site and winning a site in the large patch are the same as long as the population is monomorphic because in the present model, immigration is proportional to the number of sites so that there is no source-sink structure. The cost to the focal dispersing individual is therefore $c = (1-s)P$. Finally, in the present clonal model, the focal individual is an identical twin to all non-dispersing offspring born in her natal site, but no relative to the immigrants, and hence $r = (1-d)/(1-d+sd)$. In a solitary site, therefore, the inclusive fitness increment from dispersal is

$$\Delta w_1 = rb - c = \left[\frac{1-d}{1-d+sd} - (1-s)\right] P$$

Next, consider a site in the large patch. Since all individuals of the large patch compete for this site, the relatedness of the focal individual to the competitors is zero, and the inclusive fitness increment contains only the cost term,

$$\Delta w_2 = -c = -(1-s)P$$

The expected inclusive fitness increment from dispersal is the average of Δw_1 and Δw_2 weighted by the probabilities that the focal individual is in a solitary site or in the large patch,

$$\Delta w = p \Delta w_1 + (1-p)\Delta w_2 = \left[p \cdot \frac{1-d}{1-d+sd} - (1-s)\right] P \tag{20}$$

Dispersal is favoured and therefore d increases by long-term evolution if the inclusive fitness increment Δw is positive, i.e., if

$$d < d^* = \frac{p - (1-s)}{p - (1-s)^2} \tag{21}$$

This result recovers the singular dispersal strategy given in (9) and also proves its convergence stability. d^* is outside the interval $[0, 1]$ if $p < 1 - s$, and in this case, dispersal evolves towards zero since the expression in the brackets of (20) is negative for all $0 \leq d \leq 1$.

The inclusive fitness approach can be extended to dimorphic resident populations and can be used to obtain the dimorphic singularity $(0, d_2^*)$ as well as to prove its convergence stability (see Appendix F).

16

8 Inclusive fitness approach: Evolutionary stability

In order to establish the evolutionary stability of the singular strategy d^*, a second-order expansion of the inclusive fitness increment Δw is needed. Instead of a single focal individual, suppose now that m additional siblings leave the site (in the adaptive dynamics notation, $d_{mut} = d + m/C$). Assume that m is such that $\varepsilon = m/C$ is small, yet $1/m = O(\varepsilon)$; this implies $1/C = O(\varepsilon^2)$, i.e., C goes fast to infinity (at the end of this section, I consider an alternative limit). Assume also $1/N = O(\varepsilon)$ where N is the total number of sites.

In a solitary site, the benefit from the dispersal of the m focal offspring,

$$b = \frac{m}{(1 - d + sd)C} = \frac{\varepsilon}{1 - d + sd} \tag{22}$$

is linear in ε. To calculate the cost, assume that each offspring disperses independently, such that a solitary site receives at most one of the m offspring whereas the large patch receives $(1 - p)m$ (see Gandon and Michalakis (1999) on siblings dispersing together as a group). The number of solitary sites won by the m focal dispersers is

$$spm\frac{1}{(1 - d + sd)C + 1} = \frac{sp\varepsilon}{1 - d + sd + \varepsilon/m} = \frac{sp\varepsilon}{1 - d + sd} + O(\varepsilon^3) \tag{23a}$$

whereas the number of sites won in the large patch is

$$s(1 - p)m\frac{(1 - p)N}{[(1 - p)(1 - d) + (1 - p)sd]CN + s(1 - p)m} =$$

$$= \frac{s(1 - p)\varepsilon}{1 - d + sd + s\varepsilon/N} = \frac{s(1 - p)\varepsilon}{1 - d + sd} + O(\varepsilon^3) \tag{23b}$$

The cost incurred by the m focal dispersers is the forfeited chance to win the natal site (in (22)) minus the number of sites won elsewhere (the sum of (23a) and (23b), i.e.,

$$c = \frac{(1 - s)\varepsilon}{1 - d + sd} + O(\varepsilon^3) \tag{24}$$

Analogously to the previous section, dispersal from the large patch incurs the same cost but the benefit is received by nonrelated individuals.

The inclusive fitness increment is given by $\Delta w = \pi(rb - c) + (1 - \pi)(-c) = \pi rb - c$, where π is the probability that the focal family is in a solitary site and r is relatedness in a solitary site. Since b is of order ε, r and π have to be expanded to first order in ε to approximate Δw to second order.

To expand r, note that when a fraction ε of the offspring leaves a solitary site, their relatedness to those who compete for their natal site is

$$r = \frac{1 - d - \varepsilon}{1 - d - \varepsilon + sd} = r_0 + r_1\varepsilon + O(\varepsilon^2) \text{ with } r_0 = \frac{1 - d}{1 - d + sd}, \quad r_1 = \frac{-sd}{(1 - d + sd)^2} \tag{25}$$

If all sites are solitary as in the Hamilton-May (1977) model, then $\Delta w = rb - c$ and the coefficient of the ε^2 term is the negative because r_1 is negative. This implies that at the singular strategy, where the linear term of Δw vanishes, both increasing and decreasing dispersal (positive and negative ε) cause a loss in inclusive fitness, i.e., the singular strategy is evolutionarily stable.

To expand π, consider an infinitesimally rare mutant subpopulation with the dispersal strategy $d_{mut} = d + \varepsilon$. The population dynamics of this subpopulation are determined by the projection matrix $\mathbf{A}_{mut}$ in equation (4), and its eigenvector $(\pi, (1-\pi))^T$ gives the distribution of focal dispersers over the states of being in a solitary site or in the large patch. Taking $d = d^*$, after some cumbersome algebra (*Wolfram Mathematica*® notebook available upon request) one obtains

$$\pi = p + \pi_1 \varepsilon + O(\varepsilon^2) \text{ with } \pi_1 = \frac{(1-s)(1-p)}{sd^*} \tag{26}$$

Substituting the expansions (25) and (26) and d^* from (21) into $\Delta w = \pi rb - c$, the inclusive fitness increment in the population of the singular dispersal strategy d^* is

$$\Delta w = \frac{p - (1-s)^2}{sp}[pr_1 + \pi_1 r_0]\varepsilon^2 + O(\varepsilon^3) \tag{27}$$

(at $d = d^*$, the terms linear in ε vanish). Recall that the model has an interior singularity if $p > 1 - s$, and this implies that the factor in front of the brackets in (27) is positive. In the brackets, the first term pr_1 is negative because r_1 is negative; this term describes the same nonlinearity effect what makes the singularity of the Hamilton-May model evolutionarily stable. The second term in the brackets, $\pi_1 r_0$, is however positive. This term comes from the spatial sorting of dispersers. According to (26), the focal dispersers are more often in solitary sites relative to the distribution of a neutral mutant (with $\varepsilon > 0$, $\pi > p$). This spatial sorting implies that mutants with higher dispersal accumulate in the solitary sites, where high dispersal is favourable; and mutants with lower dispersal accumulate in the large patch, where low dispersal is favourable. Spatial sorting thus gives an advantage to mutants with dispersal either higher or lower than the singularity, and thereby leads to disruptive selection. Evolutionary branching occurs if the effect of spatial sorting is stronger than the effect of nonlinearity in the Hamilton-May model, i.e., if $\pi_1 r_0 > -pr_1$ in (27). Substituting the definitions of π_1, r_0 and r_1, one readily recovers $p < 1 - s^2$, the condition for evolutionary branching (given an interior singularity exists; cf. section 3.2).

Note that the above derivation of the evolutionary stability condition assumes a relatively large mutation (m rather than one extra individual dispersing) and takes limits such that $1/C = O(\varepsilon^2)$, which implies that the probability of winning a site after dispersal and hence c is proportional to ε at least up to second order (as in (24)). This is the limit considered by adaptive dynamics. An alternative is to assume only one extra disperser and $\varepsilon = 1/C$. In the latter case, c has a term quadratic in ε, which contributes a negative term in the brackets of (27) and thereby hinders evolutionary branching (see Appendix G for details).

9 Discussion

In this paper, I analysed a minimal model for the coexistence and evolutionary branching of dispersal strategies due to contrasting levels of kin competition. Kin competition is well known to drive the evolution of dispersal despite of the costs associated with dispersal (Hamilton and May 1977; Comins et al. 1980; Motro 1982; Frank 1986; Taylor 1988; Gandon and Michalakis 1999; Ajar 2003; Berdahl et al. 2015). Most models predicting polymorphism or evolutionary branching of dispersal, however, assume large populations with negligible competition among kin. Instead, in these models environmental stochasticity, catastrophes, or non-equilibrium population dynamics select for dispersal (Doebeli and Ruxton 1997; Johst et al. 1999; Parvinen 1999; Mathias et al. 2001; Kisdi 2002; Parvinen 2002; Nurmi and Parvinen 2011; see also Cohen and Levin 1991, whose "evolutionarily compatible strategy" is an evolutionary branching point; and Holt and McPeek 1996, whose Fig. 2 exhibits an evolutionary branching point).

Kin competition is likely to be an important factor also in the diversification of dispersal strategies. This is because the strength of kin competition, and therefore the level of dispersal selected for, differ in differently sized habitat patches. Variable patch size therefore facilitates the evolution of diverse dispersal strategies (cf. Massol et al. 2011). The present model is the direct extension of the Hamilton-May (1977) model to variable patch size, with the simplifying assumption that patches are either small (solitary sites supporting a single adult) or large (with negligible kin competition). For simplicity, I assumed a single large patch, but the model could easily be extended to include several large patches. The only difference is that the frequencies of dispersal strategies need to equilibrate in the large patches before one can pool them and define n_{2k}, the fraction of sites that are in a large patch and is occupied by strategy k.

Since competition is different in solitary sites and in the large patch, the population is structured and the population dynamics of a mutant strategy is described by the 2×2 projection matrix $\mathbf{A}$ given in (2). The invasion fitness of the mutant is the leading eigenvalue of this matrix, which is cumbersome to use. The adaptive dynamic analysis of the present model was greatly facilitated by the fitness proxy introduced by Metz and Leimar (2011; see Appendix A): For non-negative 2×2 matrices, one can use $\mathrm{tr}(\mathbf{A}) - \det(\mathbf{A})$ instead of the invasion fitness, as long as $\mathrm{tr}(\mathbf{A}) \leq 2$ (which holds for mutants similar to a resident strategy). An alternative possibility, particularly for models with many patch types, is to calculate the basic reproduction number (R_0) over dispersal generations, i.e., the number of dispersers produced by the descendants of an offspring entering the dispersal pool (Gyllenberg and Metz 2001; Metz and Gyllenberg 2001; Massol et al. 2011).

I could obtain the pairwise invasibility plots analytically (see equation (8)), from which the full characterisation of monomorphic evolution follows. For the monomorphic evolutionary singularity (equation (9)), I recover the formula of Frank (1986) with $F_{ST} = p$ (offspring taken from one patch are identical by descent if and only if the patch is a solitary site); note that Frank (1986) assumed patches of equal size, yet his result generalizes to the present model. Contrary to the case of equal patches, in the present model the singularity can be an evolutionary branching point (Figure 2), and its bifurcation lines are given by very simple formulas (transcritical bifurcation at $p = 1 - s$; bifurcation between an ESS and an evolutionary branching point at

$p = 1 - s^2$, see section 3.2).

The population dynamics of dimorphic populations exhibit multiple attractors and unprotected dimorphisms (Figures 3 and 4; analytic proof in section 4.2). If model parameters change due to a change in the environment (e.g. if the probability of survival during dispersal, s, decreases), then an unprotected dimorphism can collapse through a catastrophic saddle-node bifurcation (Figure 4). The possibility of such catastrophic extinctions is an important concern for management and conservation. The same catastrophic extinction may happen also due to evolution: If the initial dimorphic population is above the isocline in area B of Figure 4b, then the adaptive dynamics can push the system through the saddle-node bifurcation line, causing the collapse of the dimorphism. Starting from the evolutionary branching point of Figure 4a, however, the evolving dimorphism stays in the area of protected dimorphisms.

For arbitrary dimorphic resident populations, obtaining the population dynamic equilibrium was not possible analytically; as so often, the hard step of the analysis is finding the attractor of the ecological dynamics. However, if one of the resident strategies has zero dispersal, then in equilibrium it is confined to the large patch, and in this simpler case one can obtain the dimorphic equilibrium population densities in equation (11). It is therefore possible to find the evolutionarily stable dimorphism $(0, d_2^*)$ analytically (see (14)).

The model has three environmental feedback variables, and therefore no more than three strategies can coexist at equilibrium (Levin 1970; Geritz et al. 1997; Meszéna et al. 2006). I have not explored the trimorphic equilibria, but an analogy suggests that these are unstable. The Hamilton-May model has two environmental feedbacks (E_1 and E_3 in the present notation) and all its dimorphic equilibria are unstable (Motro 1982; Appendix E). The presence of the large patch adds one more environmental feedback (E_2) so that the trimorphic equilibra may now be unstable in the same way as the dimorphic equilibra of the Hamilton-May model. In any case, trimorphisms are not reachable by small mutational steps from initially monomorphic or dimorphic populations, because dimorphic populations have no singularities where evolutionary branching could happen.

In the second part of the paper, I re-derived some of the results using an inclusive fitness approach. This intuitive approach provides an easy way to find the singular strategy and its convergence stability (section 7), but establishing evolutionary stability using inclusive fitness in a structured population involves cumbersome algebra (section 8). The reason is that for the singularity and its convergence stability one needs only the selection gradient, and the selection gradient is essentially an eigenvalue sensitivity problem. The leading eigenvalue of the mutant projection matrix $\mathbf{A}_{mut}$ in (4) is 1 when $d_{mut} = d$, and the selection gradient is the eigenvalue sensitivity $[\partial\lambda/\partial d_{mut}]_{dmut=d} = \mathbf{v}^T\mathbf{B}\mathbf{u}/\mathbf{v}^T\mathbf{u}$, where $B_{ij} = [\partial A_{ij}/\partial d_{mut}]_{dmut=d}$ and $\mathbf{u}$ and $\mathbf{v}$ are respectively the right and left eigenvectors of $\mathbf{A}_{mut}$ with $d_{mut} = d$. In the present model, it is easy to see that $\mathbf{u} = [p, (1-p)]^T$ and $\mathbf{v} = [1, 1]^T$ (the latter is because there is no source-sink structure). For the eigenvalue sensitivity, therefore, one only needs to combine the elements of $[1, 1]^T\mathbf{B}$, i.e., the derivatives of $A_{11} + A_{21} = \frac{1-d_{mut}}{1-d_{mut}+sd} + \frac{sd_{mut}}{1-d+sd}$ (which is fitness in the Hamilton-May model) and $A_{12} + A_{22} = \frac{1-d_{mut}+sd_{mut}}{1-d+sd}$ (which is fitness in a large patch with no

20

kin competition). These derivatives are easy to intuit (see section 7).

To obtain the condition for evolutionary stability with the inclusive fitness approach, one needs also the eigenvector sensitivity (cf. equation (26)), which is technically more cumbersome. Yet it is the eigenvector sensitivity that explains how the spatial structure can turn selection to be disruptive at the singular strategy. As equation (26) shows, a mutant with higher dispersal accumulates in the solitary sites, where it is at an advantage, whereas a mutant with lower dispersal accumulates in the large patch, where it has an advantage. This spatial sorting makes it possible that at the singularity, mutants with both higher and lower dispersal are at an advantage, i.e., selection is disruptive.

A somewhat surprising conclusion of the inclusive fitness analysis is that to recover the condition for evolutionary stability as obtained from adaptive dynamics, one has to make specific assumptions about the order of limits (see section 8). Specifically, in this model the adaptive dynamics approach assumes that fecundity (C) goes to infinity faster than the inverse of the mutation stepsize $|d_{mut} - d|$. If instead $|d_{mut} - d| = 1/C$ such that one individual changes her dispersal behaviour as commonly assumed when using the inclusive fitness approach, then the condition for evolutionary branching becomes somewhat more restrictive (Figure 6; derivation in Appendix G).

Since the inclusive fitness approach is based on local expansions, this method is not suitable to obtain the full pairwise invasibility plot or the sets of protected and unprotected dimorphisms. I have also refrained from using the inclusive fitness approach near $d = 0$, the point where the selection gradient is not defined (see section 3.2).
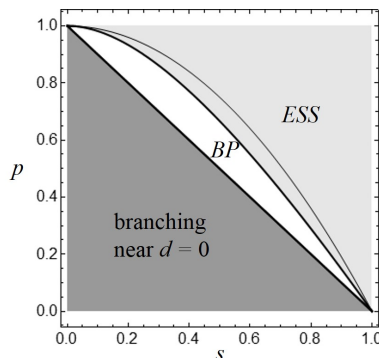

Figure 6: Bifurcations of the monomorphic singularity with the alternative assumption about limits used in Appendix G. Notation as in Figure 2. The thin line shows the boundary between BP and ESS from Figure 2 for comparison.

The main simplifying assumption of the present model is the specific distribution of patch sizes (solitary sites plus a large patch). This is of course not realistic, but makes the model highly tractable and the results simple ("proof-of-concept" model, Servedio et al. 2014). Massol et al. (2011) investigated a more realistic model with arbitrary distribution of patch sizes, and found

evolutionary branching of dispersal when the patch size distribution is positively skewed. The positive skew means that the habitat consists of many small and few large patches, such that the numbers of individuals experiencing selection in small and in large patches are comparable, a situation similar to the present model.

In Appendix H, I evaluate the results of Massol et al. (2011) for the extreme patch size distribution used in this paper, and show that their model predicts evolution to $d = 0$ for all combinations of my parameters and predicts evolutionary branching at $d = 0$ when $p > 1/2$. The discrepancy between this prediction and Figure 2 is likely to be the consequence of the source-sink structure inherent in the model of Massol et al. (2011) but absent in my model. Massol et al. (2011) assumed that immigration is independent of patch size. This creates a source-sink structure with a net flow of individuals from patches of high carrying capacity to patches of low carrying capacity. Such a source-sink structure becomes stronger when the variance of patch size increases, and it selects against dispersal (Hastings 1983). The source-sink structure thus explains why, in the model of Massol et al. (2011), the singular dispersal strategy decreases with the variance of the patch size distribution, and why dispersal evolves to $d = 0$ with the extreme patch size distribution used in my model, which has infinite variance (Appendix H). Since my model assumes that immigration is proportional to patch size, there is no source-sink structure and there can be an interior singularity (Figure 2).

Further, my model predicts that when dispersal evolves to zero, it always undergoes evolutionary branching, whereas the model of Massol et al. (2011) predicts branching for the extreme patch size distribution of my model only if $p > 1/2$. When the resident dispersal is close zero, a second strategy with higher dispersal is favoured in the solitary sites. In the model of Massol et al. (2011), such a strategy can invade only if solitary sites are sufficiently common, otherwise the net flow from the large patch is strong enough to swamp the solitary sites even if the large patch is dominated by a near-zero dispersal strategy. Since there is no source-sink structure in my model, such swamping does not occur and therefore evolutionary branchig always happens when dispersal evolves to $d = 0$.

Whether immigration is independent of patch size (as in Massol et al. 2011) or is proportional to patch size (as in the present model) or relates to patch size in some other way depends on exactly how the patches differ from each other. If all patches have the same area but large patches have high resource density so that they can support many individuals, then, assuming that immigration depends on area as in case of passive dispersal, immigration is independent of patch size and the population has a source-sink structure. If however patches differ in area rather than in resource density, then immigration is proportional to patch size and there is no source-sink structure. With the extreme patch size distribution of solitary sites and a very large patch, the former scenario would assume a very big difference in resource densities and therefore the latter possibility seems somewhat more realistic, especially if one considers an organism with limited active movement and passive long-distance dispersal, such that individuals are well mixed within the large patch but can move to solitary sites only via costly dispersal. With the general patch size distribution of Massol et al. (2011), both scenarios are possible.

Massol et al. (2011) considered a continuous-time model similar to the Moran process, which hinders a quantitative comparison with the discrete-time Hamilton-May model (see the Appendix of Massol et al. 2011). In the present model, $p \to 1$ (almost all sites are solitary) represents the limit to the simplest model version of Hamilton and May (1977). I recover the Hamilton-May ESS in this limit, but, interestingly, the pairwise invasibility plots and the dynamics of dimorphic populations do not converge to those of the Hamilton-May model (Figure 5). This is because the limit $p \to 1$ does not commute with the system size going to infinity (see section 6).

A common assumption of Massol et al. (2011) and the present model is that the individuals have no knowledge of the strength of local kin competition. In both models, disruptive selection emerges because of the contrasting levels of kin competition in small versus large patches. If the individuals can assess how strong kin competition they experience, then a conditional dispersal strategy can replace the dispersal polymorphism (Ezoe and Iwasa 1997; Kisdi 2004; Parvinen et al. 2012).

I followed the traditional assumption that the probability of survival during dispersal (s) is constant, which yields a linear trade-off between the number of nondispersed and successfully dispersed offspring (i.e., between $(1 - d)C$ and sdC). An alternative mechanism leading to evolutionary branching of dispersal in stable environments operates when dispersal is part of a nonlinear trade-off, for example dispersal is traded off against fecundity (Weigang and Kisdi 2015). A nonlinear trade-off may exist also if the probability of survival during dispersal depends on the amount of extra resources invested into dispersers (e.g. in the form of specific dispersal structures or an energy reserve that help survival during dispersal), so that if a higher fraction of the offspring disperse, then each offspring is less well equipped and hence has a lower probability to survive. While a nonlinear trade-off may either facilitate or hinder evolutionary branching depending on its convexity, the contrasting levels of kin competition explored in this paper arise from the very natural assumption of variable patch sizes and always facilitate diversity relative to the case of equally sized patches.

Acknowledgements

References

Ajar E. 2003. Analysis of disruptive selection in subdivided populations. BMC Evolutionary Biology 3: 22.

Berdahl A., C. J. Torney, E. Schertzer & S. A. Levin. 2015. On the evolutionary interplay between dispersal and local adaptation in heterogeneous environments. Evolution 69:13901405.

Boldin B. 2006. Introducing a population into a steady community: The critical case, the center manifold, and the direction of bifurcation. SIAM J. Appl. Math. 66: 1424-1453.

Cohen D. & S. A. Levin. 1991. Dispersal in patchy environments: The effects of temporal and spatial structure. Theor. Pop. Biol. 39: 63-99.

Comins H. N., W. D. Hamilton & R. M. May. 1980. Evolutionarily stable dispersal strategies. J. theor. Biol. 32: 205-230.

Diekmann O., M. Gyllenberg, H. Huang, M. Kirkilionis, J. A. J. Metz & H. R. Thieme. 2001. On the formulation and analysis of general deterministic structured population models. II. Nonlinear theory. J. Math. Biol. 43: 157-189.

Doebeli M. & G. D. Ruxton. 1997. Evolution of dispersal rates in metapopulation models: Branching and cyclic dynamics in phenotype space. Evolution 51: 1730-1741.

Ezoe H. & Y. Iwasa. 1997. Evolution of condition-dependent dispersal: A genetic-algorithm search for the ESS reaction norm. Res. Pop. Ecol. 39: 127-138.

Frank S. A. 1986. Dispersal polymorphisms in subdivided populations. J. theor. Biol. 122: 303-309.

Gandon S. & 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., E. Kisdi, G. Meszéna & 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, E. Kisdi & G. Meszéna. 1997. Dynamics of adaptation and evolutionary branching. Phys. Rev. Letters 78: 2024-2027.

Geritz S. A. H., E. van der Meijden & J. A. J. Metz. 1999. Evolutionary dynamics of seed size and seedling competitive ability. Theor. Pop. Biol. 55: 324-343.

Gyllenberg M. & J. A. J. Metz. 2001. On fitness in structured metapopulations. J. Math. Biol. 43: 545-560.

Hamilton W. D. 1963. The evolution of altruistic behavior. Am. Nat. 97: 354-356.

Hamilton W. D. & R. M. May. 1977. Dispersal in stable habitats. Nature 269: 578-581.

Hastings A. 1983. Can spatial variation alone lead to selection for dispersal? Theor. Pop. Biol. 24: 244-251.

Holt R.D. & M. A. McPeek. 1996. Chaotic population dynamics favors the evolution of dispersal. Am. Nat. 148: 709-718.

Johst K., M. Doebeli & R. Brandl. 1999. Evolution of complex dynamics in spatially structured populations. Proc. R. Soc. Lond. B 266: 1147-1154.

Kisdi E. 2002. Dispersal: Risk spreading versus local adaptation. Am. Nat. 159: 579-596.

Kisdi E. 2004. Conditional dispersal under kin competition: Extension of the Hamilton-May model to brood size-dependent dispersal. Theor. Pop. Biol. 66: 369-380.

Kuznetsov Y. A. 1995. Elements of applied bifurcation theory. Springer-Verlag, Berlin.

Leimar O. 2009. Multidimensional convergence stability. Evol. Ecol. Res. 11: 191-208.

Levin S. A. 1970. Community equilibria and stability, and an extension of the competitive exclusion principle. Am. Nat. 104: 413-423.

Massol F., A. Duputié, P. David & P. Jarne. 2011. Asymmetric patch size distribution leads to disruptive selection on dispersal. Evolution 65: 490-500.

Matessi C. & C. Di Pasquale. 1996. Long-term evolution of multilocus traits. J. Math. Biol. 34: 613-653.

Mathias A., E. Kisdi & I. Olivieri. 2001. Divergent evolution of dispersal in a heterogeneous and variable landscape. Evolution 55: 246-259.

Meszéna G., M. Gyllenberg, L. Pásztor & J. A. J. Metz. 2006. Competitive exclusion and limiting similarity: A unified theory. Theor. Pop. Biol. 69: 68-87.

Metz J. A. J. & M. Gyllenberg. 2001. How should we define fitness in structured metapopulation models? Including an application to the calculation of evolutionarily stable dispersal strategies. Proc. R. Soc. Lond. B 268: 499-508.

Metz J. A. J. & O. Leimar. 2011. A simple fitness proxy for structured populations with con-

tinuous traits, with case studies on the evolution of haplo-diploids and genetic dimorphisms. J. Biol. Dyn. 5: 163-190.

Metz J. A. J., R. M. Nisbet & S. A. H. Geritz, 1992. How should we define fitness for general ecological scenarios? Trends Ecol. Evol. 7:198-202.

Motro U. 1982. Optimal rates of dispersal. I. Haploid populations. Theor. Pop. Biol. 21: 394-411.

Nurmi T. & K. Parvinen. 2011. Joint evolution of specialization and dispersal in structured metapopulations. J. theor. Biol. 275: 78-92.

Parvinen K. 1999. Evolution of migration in a metapopulation. Bull. Math. Biol. 61: 531-550.

Parvinen K. 2002. Evolutionary branching of dispersal strategies in structured metapopulations. J. Math. Biol. 45: 106-124.

Parvinen K., J. Ohtsuki & J. Wakano. The effects of fecundity derivatives on the condition of evolutionary branching in spatial models. in prep.

Parvinen K., A. Seppänen & J. D. Nagy. 2012. Evolution of complex density-dependent dispersal strategies. Bull. Math. Biol. 74: 2622-2649.

Priklopil T. 2012. On invasion boundaries and the unprotected coexistence of two strategies. J. Math. Biol. 64: 1137-1156.

Ronce O. 2007. How does it feel to be like a rolling stone? Ten questions about dispersal evolution. Annu. Rev. Ecol. Evol. Syst. 38: 231-253.

Servedio M. R., Y. Brandvain, S. Dhole, C. L. Fitzpatrick, E. E. Goldberg, C. A. Stern, J. van Cleve & D. J. Yeh. 2014. Not just a theory - The utility of mathematical models in evolutionary biology. PLoS Biology 12: e1002017.

Taylor P. D. 1988. An inclusive fitness model for dispersal of offspring. J. theor. Biol. 130: 363-378.

Weigang H. C. & E. Kisdi. 2015. Evolution of dispersal under a fecundity-dispersal trade-off. J. theor. Biol. 371: 145153.

Appendix A

Let $\mathbf{A}$ be a non-negative 2×2 projection matrix, let λ be its leading eigenvalue (i.e., the eigenvalue with the largest absolute value), and $F = \mathrm{tr}(\mathbf{A}) - \det(\mathbf{A})$. $\mathrm{tr}(\mathbf{A}) \geq 0$ by $\mathbf{A}$ being non-negative and λ is real and non-negative by the Perron-Frobenius theorem. In this Appendix, I show that assuming $\mathrm{tr}(\mathbf{A}) \leq 2$, $F \gtreqless 1$ is equivalent to $\lambda \gtreqless 1$. This result was first obtained by Metz and Leimar (2011), who also discuss the general case of $n \times n$ matrices.

The characteristic equation of $\mathbf{A}$, $\lambda^2 - \mathrm{tr}(\mathbf{A})\lambda + \det(\mathbf{A}) = 0$, yields the eigenvalues $(\mathrm{tr}(\mathbf{A}) \pm \sqrt{\mathrm{tr}(\mathbf{A})^2 - 4\det(\mathbf{A})})/2$. Since $\mathrm{tr}(\mathbf{A}) \geq 0$, the leading eigenvalue is the "+" root. Hence we have

$$\lambda = \frac{\mathrm{tr}(\mathbf{A}) + \sqrt{\mathrm{tr}(\mathbf{A})^2 - 4\det(\mathbf{A})}}{2} \gtreqless 1$$

$$\iff \sqrt{\mathrm{tr}(\mathbf{A})^2 - 4\det(\mathbf{A})} \gtreqless 2 - \mathrm{tr}(\mathbf{A})$$

$$\stackrel{*}{\iff} \mathrm{tr}(\mathbf{A})^2 - 4\det(\mathbf{A}) \gtreqless 4 - 4\,\mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{A})^2$$

$$\iff \mathrm{tr}(\mathbf{A}) - \det(\mathbf{A}) \gtreqless 1$$

where the step marked with $*$ used that $\mathrm{tr}(\mathbf{A}) \leq 2$ and λ is real (the expression under the square root is non-negative). Note that if $\mathrm{tr}(\mathbf{A}) > 2$, then $\lambda > 1$ follows immediately.

Appendix B

In this Appendix, I first outline the proof that $F(d_{mut}, d) = 1$ is equivalent to equation (7) of the main text. Substituting the elements of $\mathbf{A}_{mut}$ from (4) into $F(d_{mut}, d) = \mathrm{tr}(\mathbf{A}_{mut}) - \det(\mathbf{A}_{mut}) = 1$ gives

$$\frac{1 - d_{mut}}{1 - d_{mut} + sd} + \frac{d_{mut}sp}{1 - d + sd} + \frac{1 - d_{mut} + d_{mut}s(1-p)}{1 - d + sd} -$$
$$- \left[\frac{1 - d_{mut}}{1 - d_{mut} + sd} + \frac{d_{mut}sp}{1 - d + sd} \right] \frac{1 - d_{mut} + d_{mut}s(1-p)}{1 - d + sd} + \frac{d_{mut}sp}{1 - d + sd} \frac{d_{mut}s(1-p)}{1 - d + sd} - 1 = 0$$

Multiplying with the denominators yields a cubic polynomial in d_{mut} and d. This cubic polynomial can be divided with $(d_{mut} - d)$, and the result readily simplifies into (7).

Next, I prove that the left branch of the hyperbola in (8) is an invasion boundary. As shown in the main text, it suffices to demonstrate that $\mathrm{tr}(\mathbf{A}_{mut}) < 2$ holds at one point of each segment of the left branch in $(0, 1]^2$. Assume first that $\frac{1 - \sqrt{1-p}}{2(1-s)} \leq 1$, so that the left vertex of the hyperbola is in the unit square and the left branch gives one continuous line in the unit square. Substituting the coordinates of the vertex into

$$\mathrm{tr}(\mathbf{A}_{mut}) = \frac{1 - d_{mut}}{1 - d_{mut} + sd} + \frac{1 - d_{mut} + sd_{mut}}{1 - d + sd} \tag{28}$$

27

yields

$$\text{tr}(\mathbf{A}_{mut}) = \frac{2 + \sqrt{1-p}(1 - s(2+s))}{(1 + \sqrt{1-p})(1 - s\sqrt{1-p})}.$$

The claim $\text{tr}(\mathbf{A}_{mut}) < 2$ is equivalent to

$$
\begin{aligned}
2 + \sqrt{1-p}(1 - s(2+s)) &<& 2 + 2\sqrt{1-p} - 2s\sqrt{1-p}(1 + \sqrt{1-p}) \\
-s(2+s)\sqrt{1-p} &<& \sqrt{1-p} - 2s\sqrt{1-p}(1 + \sqrt{1-p}) \\
-s(2+s) &<& 1 - 2s(1 + \sqrt{1-p}) \\
1 - 2s\sqrt{1-p} + s^2 &>& 0
\end{aligned}
$$

where I have used that $p < 1$ by assumption. The last inequality holds since $1 - 2s + s^2 \geq 0$ and $p > 0$ is assumed with strict inequality. Because $\text{tr}(\mathbf{A}_{mut}) < 2$ holds at the left vertex, it must hold on the entire left branch in the interior of the unit square (see main text).

Assume now that $\frac{1-\sqrt{1-p}}{2(1-s)} > 1$, so that the left vertex of the hyperbola is outside the unit square. The left branch of the hyperbola has then two separate segments in the unit square, which intersect the $d = 1$ edge at the points $(1, d_{mut}^{(1)})$ and $(1, d_{mut}^{(2)})$, where, from equation (7) with $d = 1$, $d_{mut}^{(1)}$ and $d_{mut}^{(2)}$ are the roots of the quadratic equation $p d_{mut}(1 - d_{mut}) = s(1 - s)$. To see that $\text{tr}(\mathbf{A}_{mut}) < 2$ holds at both of these points, evaluate $\text{tr}(\mathbf{A}_{mut})$ in (28) at $d = 1$ to arrive at the claim

$$\frac{1 - d_{mut}}{1 - d_{mut} + s} + \frac{1 - d_{mut} + s d_{mut}}{s} < 2$$

to hold at $d_{mut}^{(1)}$ and $d_{mut}^{(2)}$. This inequality simplifies to

$$(1-s)(1 - d_{mut})(1 - d_{mut} + s) < s^2 \tag{29}$$

where the left hand side is an upward opening parabola with roots $d_{mut} = 1$ and $d_{mut} = 1 + s$. Hence the left hand side of (29) is a decreasing function of d_{mut} for the biologically relevant values $0 \leq d_{mut} \leq 1$, and the inequality holds at both $d_{mut}^{(1)}$ and $d_{mut}^{(2)}$ if it holds at the smaller one. The explicit expressions for $d_{mut}^{(1)}$ and $d_{mut}^{(2)}$ are somewhat unruly and depend on p (whereas (29) does not). However, it is easily seen from the equation $p d_{mut}(1 - d_{mut}) = s(1 - s)$ that $d_{mut}^{(1)}$ and $d_{mut}^{(2)}$ are symmetric about $\frac{1}{2}$ and are the farthest apart, such that the smaller root assumes its minimal value, when $p = 1$. In this case, $d_{mut}^{(1)} = s$ and $d_{mut}^{(2)} = 1 - s$, and since the condition $\frac{1-\sqrt{1-p}}{2(1-s)} > 1$ implies $s > \frac{1}{2}$, $1 - s$ is the smaller root. Substituting $d_{mut} = 1 - s$ into (29) we arrive at the claim $2s^2(1 - s) < s^2$, which holds due to $s > \frac{1}{2}$. Therefore $\text{tr}(\mathbf{A}_{mut}) < 2$ holds at one point of each segment, which implies the same on the entire left branch of the hyperbola in the interior of the unit square (see main text).

The right branch of the hyperbola intersects the unit square if and only if the right vertex is in the unit square, i.e., if $\frac{1+\sqrt{1-p}}{2(1-s)} < 1$. A derivation analogous to the above shows that at the right vertex $\text{tr}(\mathbf{A}_{mut}) > 2$, and therefore, as shown in the main text, the same holds on the entire right branch in the unit square. The right branch is thus not an invasion boundary.

Appendix C

Here I derive the selection gradients of dimorphic populations with strategies $d_1 = 0$ and arbitrary d_2 given in equations (13) and (15) of the main text.

Assume small mutations, i.e., $d_{mut} = d_i + \delta$, where d_i is a resident ($i = 1, 2$) and $|\delta| \ll 1$. Substituting $D = d_{mut}$ and $\hat{E}_2 = 1$ from (12) into (2), the projection matrix of a mutant of strategy i is

$$\mathbf{A}_{mut}^{(i)} = \begin{bmatrix} A_{11}^{(i)} & sp\hat{E}_3(d_i + \delta) \\ s(1-p)(d_i + \delta) & 1 - (d_i + \delta)(1 - s(1-p)) \end{bmatrix} \tag{30}$$

where, by Taylor expanding the nonlinear element A_{11},

$$A_{11}^{(i)} = \frac{1 - d_i}{1 - d_i + \hat{E}_1} - \frac{\hat{E}_1}{(1 - d_i + \hat{E}_1)^2}\delta - \frac{\hat{E}_1}{(1 - d_i + \hat{E}_1)^3}\delta^2 + sp\hat{E}_3(d_i + \delta) + O(\delta^3)$$

$F_2(d_i, 0, d_2) = 1$ for both residents ($i = 1, 2$). Collecting terms linear in δ in the trace and the determinant of $\mathbf{A}_{mut}^{(i)}$, one obtains

$$F_2(d_i + \delta, 0, d_2) = 1 + \left[2sp\hat{E}_3 d_i - (1 - s(1-p))\frac{\hat{E}_1(1 + \hat{E}_1)}{(1 - d_i + \hat{E}_1)^2}\right]\delta + O(\delta^2) \tag{31}$$

Taking $i = 1$ and substituting $d_1 = 0$, (31) simplifies to

$$F_2(\delta, 0, d_2) = 1 - (1 - s(1-p))\frac{\hat{E}_1}{(1 + \hat{E}_1)}\delta + O(\delta^2)$$

which immediately yields the selection gradient in (13a).

For $i = 2$, substitute $\hat{E}_3 = 1/(1 - d_2 + \hat{E}_1)$ from (12) into (31) to arrive at

$$F_2(d_2 + \delta, 0, d_2) = 1 + \frac{1}{1 - d_2 + \hat{E}_1}\left[2spd_2 - (1 - s(1-p))\frac{\hat{E}_1(1 + \hat{E}_1)}{1 - d_2 + \hat{E}_1}\right]\delta + O(\delta^2)$$

Since $\hat{E}_1(1 - s(1-p)) = spd_2$ by (12), this expression simplifies to

$$\begin{aligned} F_2(d_2 + \delta, 0, d_2) &= 1 + \frac{spd_2}{1 - d_2 + \hat{E}_1}\left[2 - \frac{1 + \hat{E}_1}{1 - d_2 + \hat{E}_1}\right]\delta + O(\delta^2) \\ &= 1 + \frac{spd_2}{(1 - d_2 + \hat{E}_1)^2}\left[1 - 2d_2 + \hat{E}_1\right]\delta + O(\delta^2) \end{aligned}$$

which yields (13b), where $h(d_2)$ is $[1 - 2d_2 + \hat{E}_1]$ with $\hat{E}_1$ substituted from (12).

At the singular dimorphism $(0, d_2^*)$, the linear terms of the fitness proxy $F_2(d_2^* + \delta, 0, d_2^*)$ vanish. Collecting terms quadratic in δ in the trace and the determinant of $\mathbf{A}_{mut}^{(2)}$, one arrives at

$$F_2(d_2^* + \delta, 0, d_2^*) = 1 + \left[sp\hat{E}_3 - (1 - s(1 - p)) \left(\frac{\hat{E}_1}{(1 - d_2^* + \hat{E}_1)^2} + \frac{\hat{E}_1 d_2^*}{(1 - d_2^* + \hat{E}_1)^3} \right) \right] \delta^2 + O(\delta^3)$$

$h(d_2^*) = 0$ implies $[1 - 2d_2^* + \hat{E}_1] = 0$ and therefore $1 - d_2^* + \hat{E}_1 = d_2^*$, so that $\hat{E}_3 = 1/d_2^*$ and

$$F_2(d_2^* + \delta, 0, d_2^*) = 1 + \left[\frac{sp}{d_2^*} - (1 - s(1 - p)) \frac{2\hat{E}_1}{(d_2^*)^2} \right] \delta^2 + O(\delta^3)$$

Using $\hat{E}_1(1 - s(1 - p)) = spd_2^*$ from (12), this simplifies to

$$F_2(d_2^* + \delta, 0, d_2^*) = 1 - \frac{sp}{d_2^*} \delta^2 + O(\delta^3)$$

which readily yields (15).

Appendix D

To search for interior evolutionary singularities of dimorphic resident populations, I obtained the equilibrium densities of the resident strategies and the selection gradients as described in the main text. Assume, without loss of generality, that $d_2 > d_1$, i.e., consider the upper left halves (above the diagonal) of the symmetric isocline plots in Figure 3. A d_2-isocline (where $G_2(d_1, d_2) = 0$) must connect to the boundary singularity $(0, d_2^*)$ with d_2^* given in (14). First, I used numerical continuation to follow this isocline in the interior, and verified that $G_1(d_1, d_2)$ does not change sign along this isocline; this means that there is no dimorphic singularity on the d_2-isocline that connects to the boundary singularity. Second, I searched the interior for any other d_2-isocline checking for other zeros of $G_2(d_1, d_2)$ on a 100×100 mesh. The entire procedure was repeated with every combination of parameter values $s \in \{0.1, 0.2, ..., 0.9\}$ and $p \in \{0.1, 0.2, ..., 0.9\}$. This search did not find any d_2-isocline other than the one connecting to the boundary singularity, and found no interior dimorphic singularity.

Appendix E

In this Appendix, I consider the dimorphic equilibria of the Hamilton-May (1977) model. Since this model has only two environmental feedback variables (E_1 and E_3, see section 6), at most two resident strategies can coexist at equilibrium (Levin 1970; Geritz et al. 1997; Meszéna et al. 2006). Motro (1982) has shown that dimorphic equilibria exist, but they are always unstable. Here I put the proof in the context of the present paper.

Let n denote the fraction of sites occupied by the resident strategy d_1; the remaining fraction $1 - n$ is occupied by the resident d_2 (since all sites are solitary and all sites are occupied, the population dynamics is one-dimensional). The environmental feedback variables are

$$E_1 = s(nd_1 + (1 - n)d_2) \quad \text{and} \quad E_3 = \frac{n}{1 - d_1 + E_1} + \frac{1 - n}{1 - d_2 + E_1} \tag{32}$$

At the population dynamic equilibrium of the two resident strategies, n must be such that λ_{mut} given in equation (17) equals 1 when $d_{mut} = d_1$ (or, equivalently, $d_{mut} = d_2$), i.e., we have

$$\frac{1 - d_1}{1 - d_1 + E_1} + sd_1 \left[\frac{n}{1 - d_1 + E_1} + \frac{1 - n}{1 - d_2 + E_1} \right] = 1 \tag{33}$$

at equilibrium. Since E_1 is linear in n, (33) can be rearranged into a quadratic equation for n. The two roots of this quadratic equation are $n = 1$ and the nontrivial root

$$\hat{n} = \frac{1 - d_1 - d_2 + sd_2}{s(d_2 - d_1)} \tag{34}$$

$\hat{n}$ is between 0 and 1 if (d_1, d_2) is between the lines $d_2 = 1 - (1 - s)d_1$ and $d_2 = (1 - d_1)/(1 - s)$, which correspond to the invasion boundary in (18) and its inverse (i.e., mirror image), respectively. Hence the dimorphic equilibrium exists precisely in the area of mutual exclusion, delineated by the invasion boundary and its mirror image in the pairwise invasibility plot.

With a slight abuse of notation, let $\lambda(n)$ denote the annual growth rate of strategy d_1, i.e., λ_{mut} in (17) with $d_{mut} = d_1$ and the environmental feedbacks as in (32). The equilibrium $\hat{n}$ is unstable if $\frac{\partial \lambda(n)n}{\partial n}\big|_{n=\hat{n}} > 1$, which is equivalent to $\frac{\partial \lambda(n)}{\partial n}\big|_{n=\hat{n}} > 0$. This derivative evaluates to

$$\frac{\partial \lambda(n)}{\partial n}\bigg|_{n=\hat{n}} = \frac{s(d_1 - d_2)}{d_1 d_2}(1 - d_1 - d_2 + sd_1) = \frac{s^2(d_1 - d_2)^2}{d_1 d_2}(1 - \hat{n}) \tag{35}$$

which is indeed positive at every biologically relevant dimorphic equilibrium ($0 < \hat{n} < 1$). Hence the Hamilton-May model has no stable equilibria of dimorphic resident populations.

Appendix F

Here I use the inclusive fitness approach to derive the selection gradients in a dimorphic population and thereby obtain the dimorphic singularity $(0, d_2^*)$ and establish its convergence stability. Consider first a solitary site occupied by strategy d_i ($i = 1, 2$). A focal offspring could win this site with probability $P_1^{(i)} = \frac{1}{[1 - d_i + \hat{E}_1]C}$ (where $\hat{E}_1$ is the number of immigrants as in (1a) at population dynamic equilibrium), hence if the focal individual disperses, she loses this chance to win the natal site and thereby gives a benefit $P_1^{(i)}$ to others in her natal site. To calculate the cost to the focal individual, I first calculate the probability of winning another site upon successful dispersal. A disperser arrives at a solitary site occupied by strategy d_j with probability $s\hat{n}_{1j}$ and wins this site with probability $P_1^{(j)}$; and she arrives at the large patch with probability $s(1 - p)$ and wins a site there with probability $P_2 = \hat{E}_2/C$ with $\hat{E}_2$ given in (1b). The cost

to the dispersing focal individual is therefore $P_1^{(i)} - s[\hat{n}_{11} P_1^{(1)} + \hat{n}_{12} P_1^{(2)} + (1-p) P_2]$. Finally, the relatedness between the focal individual and those benefiting from her dispersal is $\frac{1-d_i}{1-d_i+\hat{E}_1}$. Hence the inclusive fitness increment from leaving a solitary site of strategy d_i is

$$\Delta w_1^{(i)} = \frac{1-d_i}{1-d_i+\hat{E}_1} P_1^{(i)} - [P_1^{(i)} - s(\hat{n}_{11} P_1^{(1)} + \hat{n}_{12} P_1^{(2)} + (1-p) P_2)]$$

If the focal individual disperses from the large patch where almost all others are unrelated to her, then the inclusive fitness increment contains only the cost term, i.e.,

$$\Delta w_2 = -[P_2 - s(\hat{n}_{11} P_1^{(1)} + \hat{n}_{12} P_1^{(2)} + (1-p) P_2)]$$

An individual with strategy d_i is born in a solitary site with probability $\hat{n}_{1i}/(\hat{n}_{1i} + \hat{n}_{2i})$ and in the large patch with probabilty $\hat{n}_{2i}/(\hat{n}_{1i} + \hat{n}_{2i})$. The expected inclusive fitness increment from dispersal for an individual with strategy d_i is therefore

$$\Delta w^{(i)} = \frac{\hat{n}_{1i}}{\hat{n}_{1i} + \hat{n}_{2i}} \Delta w_1^{(i)} + \frac{\hat{n}_{2i}}{\hat{n}_{1i} + \hat{n}_{2i}} \Delta w_2$$

Note that the equilibrium densities $\hat{n}_{ij}$ must be determined from the resident population dynamics as in the main text (section 5), and they are complicated functions of the strategies d_1 and d_2. The sign of $\Delta w^{(i)}$ gives the direction of the selection gradient of strategy d_i, and the nullcline of $\Delta w^{(i)}$ is the d_i-isocline (cf. Figure 3).

Consider now the putative dimorphic singularity $(0, d_2^*)$. With $d_1 = 0$, we have $\hat{n}_{11} = 0$, $\hat{n}_{12} = p$ and $P_2 = \hat{E}_2/C = 1/C$ (cf. section 4.1) so that $\Delta w^{(1)}$ simplifies to

$$\Delta w^{(1)} = \Delta w_2 = -\frac{1}{C} + s\left(p P_1^{(2)} + \frac{1-p}{C}\right)$$

Substituting $P_1^{(2)} = \frac{1}{[1-d_i+\hat{E}_1]C}$ and $\hat{E}_1 = \frac{spd_2}{1-s(1-p)}$ (see (12) in section 4.1), the inequality $\Delta w^{(1)} < 0$ simplifies to $(1-d_2)(1-s) > 0$. Since the latter is always satisfied, the putative singularity $(0, d_2^*)$ cannot be invaded by mutants of $d_1 = 0$. Moreover, $\Delta w^{(1)} < 0$ implies that the selection gradient of strategy d_1 near $d_1 = 0$ is negative, i.e., $(0, d_2^*)$ is transversally stable.

To find the singular value d_2^*, substitute $\hat{n}_{ij}$ from (11) in section 4.1 into $\Delta w^{(2)}$. After some algebra, this yields

$$\Delta w^{(2)} = \frac{sp(1 - d_2 + \hat{E}_1) - \hat{E}_1(1 - s(1-p))}{(1 - d_2 + \hat{E}_1)^2 C}$$

where $\hat{E}_1 = \frac{spd_2}{1-s(1-p)}$ as above. It is straightforward to show that the numerator of $\Delta w^{(2)}$ is a linearly decreasing function of d_2 such that $\Delta w^{(2)}$ is positive when

$$d_2 < d_2^* = \frac{1 - s + sp}{2(1-s) + sp}$$

whereby we have recovered the dimorphic singularity in (14) and have also shown its convergence stability.

Appendix G

Here I derive the condition for evolutionary stability of the monomorphic singularity using the inclusive fitness approach and assuming $m = 1$ and $1/C = \varepsilon$, i.e., assuming that fecundity goes slower to infinity than in section 8. In this case, (23a) becomes

$$sp\frac{1}{(1 - d + sd)C + 1} = \frac{sp\varepsilon}{1 - d + sd + \varepsilon} = sp\left[\frac{\varepsilon}{1 - d + sd} - \frac{\varepsilon^2}{(1 - d + sd)^2}\right] + O(\varepsilon^3) \qquad (36)$$

The benefit in (22) and the probability of winning a site in the large patch in (23b) remain the same (the latter assuming $1/N = O(\varepsilon)$ as before), and also the first order expansions of r and π remain the same as in section 8. Using (36) instead of (23a) in the cost term of $\Delta w = \pi r b - c$, the inclusive fitness increment at the singularity is

$$\Delta w = \frac{p - (1 - s)^2}{sp}\left[pr_1 + \pi_1 r_0 - (p - (1 - s)^2)\right]\varepsilon^2 + O(\varepsilon^3)$$

and since $p - (1 - s)^2 > 0$ when there is an interior singularity, this alternative limit yields an extra negative second order term in Δw compared to (27). Figure 6 shows the effect on the bifurcation plot of the monomorphic singularity.

Appendix H

In this Appendix, I investigate the conditions for evolutionary branching given by Massol et al. (2011) for the specific patch size distribution used in this paper. In the model of Massol et al. (2011), the singular dispersal strategy is given by

$$d^* = \min\left[\frac{1}{(1 - s + \gamma_2)\bar{K}}, 1\right] \qquad (37)$$

and it is evolutionarily stable if

$$\gamma_3 < 2\gamma_2^{1/2} + \frac{(1 - s)\bar{K} - 1}{\bar{K}}\gamma_2^{-1/2} + \frac{(1 - s)(\bar{K} - 1)}{\bar{K}}\gamma_2^{-3/2} \qquad (38)$$

where $\bar{K}$ is the mean patch size (i.e., the mean number of individuals supported by a patch), $\gamma_2 = \mathrm{E}[(K - \bar{K})^2]/\bar{K}^2$ is the squared coefficient of variation of the patch size distribution, and $\gamma_3 = \mathrm{E}[(K - \bar{K})^3]/(\bar{K}^3\gamma_2^{3/2})$ is its standardized skewness.

In the patch size distribution used in the present paper, there are pN solitary sites with $K = 1$ and 1 large patch with $K = (1 - p)N$. Hence the mean patch size is

$$\bar{K} = \frac{pN}{1 + pN} \cdot 1 + \frac{1}{1 + pN} \cdot (1 - p)N = \frac{N}{1 + pN}$$

and therefore $\bar{K} \to 1/p$ as $N \to \infty$. The variance of the patch size distribution is

$$\mathrm{E}[(K - \bar{K})^2] = \frac{pN}{1 + pN}\left(1 - \frac{N}{1 + pN}\right)^2 + \frac{1}{1 + pN}\left((1 - p)N - \frac{N}{1 + pN}\right)^2 = \frac{pN(1 - (1 - p)N)^2}{(1 + pN)^2}$$

which yields
$$\gamma_2 = \frac{p(1 - (1-p)N)^2}{N}$$
and therefore γ_2 is asymptotically equivalent to $p(1-p)^2 N$ as $N \to \infty$. Finally, the standardized skewness is
$$\gamma_3 = \frac{pN - 1}{\sqrt{pN}}$$
and therefore γ_3 is asymptotically equivalent to $\sqrt{pN}$ as $N \to \infty$.

   Since $\gamma_2 \to \infty$, the singular dispersal strategy in (37) is zero. In (38), the second and third terms of the right hand side are negligible for $N \to \infty$, and the inequality simplifies to $\sqrt{pN} < 2(1-p)\sqrt{pN}$ or $p < 1/2$. Hence the condition for evolutionary stability is violated and the singularity is an evolutionary branching point if $p > 1/2$.