Pushed to the edge: Spatial sorting can slow down invasions

Allison K. Shaw*, Frithjof Lutscher†, Lea Popovic‡

October 27, 2022

Running title: Spatial sorting can slow down invasions

Authorship Statement: The project originated from discussions among all authors in 2020-2021; AKS derived and analyzed the simulation model; FL derived and analyzed the analytic model; all authors discussed the results; AKS and FL wrote the initial draft; all authors discussed and edited the manuscript.

Data accessibility statement: Model code and simulation data will be made available via Zenodo (zenodo.org/) upon manuscript acceptance.

Keywords: Allee effect, biological invasion, dispersal kernel, eco-evolutionary dynamics, integrodifference equation, invasion speed, range expansion, spread speed

Article type: Letter

- Words: 150 (abstract), 3582 (main text)

- References: 34

- Items: 5 (Figures) + 0 (Tables) + 0 (Textboxes)

*Department of Ecology, Evolution, and Behavior, University of Minnesota, St. Paul, MN, USA, ashaw@umn.edu, phone: 612-301-7734, fax: 612-624-6777; corresponding author

†Department of Mathematics and Statistics and Department of Biology, University of Ottawa, Ottawa, ON, Canada, frithjof.lutscher@uottawa.ca

‡Department of Mathematics and Statistics, Concordia University, Montreal, QC, Canada, lea.popovic@concordia.ca

1   Abstract

Our ability to understand population spread dynamics is complicated by rapid evolution, which renders simple ecological models insufficient. If dispersal ability evolves, more dispersive individuals may arrive at the population edge than less dispersive individuals (spatial sorting), accelerating spread. If individuals at the low-density population edge benefit (escape competition), high dispersers have a selective advantage (spatial selection). These two processes are often described as forming a positive feedback loop; they reinforce each other, leading to faster spread. Although spatial sorting is close to universal, this form of spatial selection is not: low densities can be detrimental for organisms with Allee effects. Here, we present two conceptual models to explore the feedback loops that form between spatial sorting and spatial selection. We show that the presence of an Allee effect can reverse the positive feedback loop between spatial sorting and spatial selection, creating a negative feedback loop that slows population spread.

2 Introduction

Since 'nothing in biology makes sense except in the light of evolution' (Dobzhansky, 1973), it is somewhat surprising that we are still grappling with understanding how evolution plays out in ecological contexts like population spread (Phillips, 2015; Miller *et al.*, 2020). In spreading populations, evolution can take the form of four processes, each with a parallel to the four typical evolutionary processes. First, under gene surfing, akin to genetic drift in a non-spreading population, stochastic events at the typically low-density edge of the population lead to some alleles reaching high frequency by chance, as the population spreads (Edmonds *et al.*, 2004; Klopfstein *et al.*, 2006). Second, under spatial sorting, akin to gene flow by phenotype, individuals sort by dispersal phenotype within a population, because more dispersive individuals are more likely to arrive at the population edge and thus mate with other highly dispersive individuals (Cwynar & MacDonald, 1987; Shine *et al.*, 2011). Third, under spatial selection, akin to natural selection, selection can vary spatially e.g. if individuals at the low-density population edge experience a reproductive benefit compared to individuals at the high-density population core (Phillips *et al.*, 2008). The fourth process, mutation, acts just as it would in a non-spreading context. Currently, we are grappling with how these processes interact: when and how feedbacks occur, under what conditions one process might override another, and when each process acts most strongly (Miller *et al.*, 2020).

A widespread outcome across theoretical and empirical studies is that spatial sorting and spatial selection can interact to promote faster population spread (Phillips *et al.*, 2008; Perkins *et al.*, 2013; Kubisch *et al.*, 2013; Williams *et al.*, 2016; Ochocki & Miller, 2017). Under spatial selection, if individuals experience a selective benefit by escaping competition at low density (and population density varies spatially across the range), then selection varies spatially, which favors individuals that arrive at the low-density population edge (Travis & Dytham, 2002). Under spatial sorting, more dispersive individuals are more likely to arrive at a population's edge (purely by their tendency to travel further) than less dispersive individuals (Cwynar & MacDonald, 1987; Hanski *et al.*, 2002). The most dispersive individuals at the population edge will tend

3

to mate with each other (Olympic Village effect; Phillips *et al.* 2008), which could in turn generate novel

phenotypes that are even more dispersive (Shine *et al.*, 2011). Thus, acting together, spatial sorting and spatial selection can form a positive feedback loop: more dispersive individuals arrive at the population edge where they have a selective advantage. This feedback loop favors increasingly dispersive individuals and leads to faster population spread (Phillips *et al.*, 2008). However, this logic relies on the assumption that organisms benefit from being at the low density edge of the population. What if this is not the case?

Being at low population density is not always the paradise it seems, in that individuals may not always experience an overall benefit at low density. For example, difficulty finding mates, reduced facilitation, increased inbreeding, loss of heterozygosity, and increased demographic stochasticity are all widespread costs of low density (Courchamp *et al.*, 1999; Gascoigne *et al.*, 2009). Each of these mechanisms can lead to an Allee effect, where per capita growth decreases at low density, which can translate into an increased chance of local extinction at the population level (Stephens & Sutherland, 1999). In the case of a strong Allee effect, there is a threshold population density, the Allee threshold, below which population growth is negative. Models of spatial spread that have included Allee effects have found that the presence of an Allee effect can slow population spread compared to its absence (Travis & Dytham, 2002) or even reverse the effect of de novo evolution on spread, favoring individual dispersal that causes slower population spread over time (Shaw & Kokko, 2015), by disfavoring individuals that are at low density (Korolev, 2015). However, how Allee effects interact with spatial sorting, and under what conditions an Allee effect may be sufficiently strong to override accelerating effects of spatial sorting on spread, remains an open question (Miller *et al.*, 2020).

Here, we show how the presence of an Allee effect can flip the positive feedback loop between spatial sorting and spatial selection into a negative feedback loop. To illustrate our point, we develop a pair of conceptual models and use them to explore scenarios with and without spatial sorting. We find that if the Allee threshold is sufficiently small, spatial sorting and spatial selection interact via a positive feedback loop

to speed up population spread. However, if the Allee threshold is sufficiently large, these two processes create a negative feedback loop: spatial sorting pushes higher dispersers to the population edge, where they are selected against by the Allee effect, thus slowing population spread. In this case, populations spread slower in the presence of spatial sorting than in its absence.

3 Methods

Here we consider a population-based model in continuous one-dimensional space x and discrete time t (years) which tracks the density of individuals with dispersal strategy i as $n_i(x,t)$. The processes of dispersal and growth occur sequentially within each year. For simplicity, we focus on evolution from standing variation, and ignore de novo evolution via mutation or recombination. Below, we provide a description of the general model framework, and then details about the two implementations that we considered: simulations of the full model, and an analytic approximation of the model.

3.1 Framework

Dispersal occurs according to a dispersal strategy i with $i = 1, ..., \tau$. A fraction p_i of the individuals with dispersal strategy i disperse while the remaining fraction $(1 - p_i)$ stay in place. Dispersing individuals follow the same dispersal kernel, k, which gives the probability of traveling from a location y to a location x. We assume that the kernel is symmetric and denote by v its variance. A proportion μ of all dispersing individuals die.

Growth is density-dependent with an Allee effect. The density-dependence part of the growth function is given by

$$g(N) = \begin{cases} 0, & \text{if } N(x,t) < a \\[2ex] \frac{b}{b+N(x,t)}, & \text{otherwise} \end{cases} \tag{1a}$$

5

where

$$N(x,t) = \sum_{i=1}^{\tau} n_i(x,t) \tag{1b}$$

is the total number of individuals in location x in year t, a is the Allee threshold, and b is a density-dependence parameter.

We considered two scenarios for determining newborn dispersal strategy: with and without evolution via spatial sorting. For simulations with spatial sorting, offspring inherit their parent's dispersal strategy exactly. Thus, the number of individuals after growth (offspring are born, parents die) is given by

$$f(n_i, N) = \lambda g(N) n_i(x,t) \tag{2a}$$

where λ is the growth rate. This scenario results in a gradient of dispersal strategies across the population, with more dispersive strategies (high p_i) at the population edge and less dispersive strategies (low p_i) at the core – i.e., spatial sorting. For simulations without spatial sorting, offspring inherit a dispersal strategy from a uniform distribution, so the number of individuals is given by

$$f(N) = \left(\frac{1}{\tau}\right) \lambda g(N) N(x,t) \ . \tag{2b}$$

This scenario results in no gradient of dispersal strategies across the population; all locations have the same even distribution of dispersal strategies – i.e., no spatial sorting.

Concatenating the processes of growth and dispersal, the population density next year $(t+1)$ is given by the integrodifference equation

$$n_i(x, t+1) = (1 - p_i) f\big(n_i(x,t)\big) + p_i (1 - \mu) \int_{-\infty}^{\infty} k(x - y) f\big(n_i(y,t)\big) \, dy \tag{3}$$

for each strategy i.

6

3.2 Simulation approach

First, we ran numerical simulations of the full model. We used a Laplace dispersal kernel

$$k(x - y; v) = \frac{1}{\sqrt{2v}} \exp\left[-\sqrt{\frac{2(x-y)^2}{v}}\right] \tag{4}$$

and initialized each simulation with individuals present only in the center of space ($n_i(x, t) = 1$ for $i = 1, ...\tau$, $|x| < 0.5$; 0 otherwise). By scaling space and density, we can fix parameters b and v without losing generality of the results. We iterated the model forward 150 years ($t = 150$), recording the population density at each year (t) over space (x). To quantify the population spread rate, we first found the location of the population edge (the farthest point where the population density exceeded a threshold of 0.001) for each year t. Then, we took the difference in population edge location from one year to the next as the spread rate. We held most parameters constant (see Table S1 for all model variables, parameters and default values for simulations), and varied two parameters: the Allee threshold (a) and the dispersal mortality (μ). For each parameter combination we ran one simulation with spatial sorting (offspring inherit their parent's dispersal strategy) and one without spatial sorting (offspring inherit a dispersal strategy from a fixed distribution) and compared the results.

3.3 Analytic approximation

Second, we derived an analytic approximation of the model that is explicitly solvable. We summarize our approach here and give details in the Supporting Information. The approach is based on a separation of time scales. Without spatial sorting, we have only two time scales. We assume that reproduction is fast compared to dispersal. Then the population density in the next year is at carrying capacity (resp. at extinction) if it is above (resp. below) the Allee threshold in the current year (Kot *et al.*, 1996). Combined with the non-sorting

7

scenario that offspring traits are uniformly distributed, we obtain the growth function

$$f(N) = \begin{cases} 0, & \text{if } N < a \\ 1/\tau, & \text{otherwise} \end{cases} \tag{5}$$

for the integrodifference equation (3). (The carrying capacity is scaled to unity.)

The spreading behavior of this model can be analyzed completely (Lutscher, 2019). The population will spread if the average dispersal strategy $\bar{p} = \sum p_i/\tau$ satisfies $\bar{p}(1 - \mu) > 2a$, and the spread rate c^* is given implicitly by

$$F(c^*) = 1 - \frac{a}{\bar{p}(1 - \mu)}, \tag{6}$$

where F is the cumulative density function of k.

In the case with spatial sorting, we only consider two types for simplicity ($\tau = 2$). We assume that competition happens at an intermediate time scale: slower than reproduction but faster than dispersal. When dispersers spread from the current extent of the range, three zones emerge. Behind is the current extent of the range, whereas far ahead is the region where the total disperser density is below the Allee threshold. In between is the newly occupied region where the density is above the Allee threshold. In the far ahead region, the population density is zero in the next year because of the Allee effect. In the current extent, we consider a "winner takes all" competition where the lower disperser eventually wins because it does not move as much. In between, the total population reaches carrying capacity while the ratio of the two types corresponds to the ratio of the dispersers in this region. For this model, we can derive equations for the length of the in-between zone and the frequency of the high disperser there (see Supporting Information).

4   Results

First, imagine the case of a population with no Allee effect ($a = 0$) and no dispersal mortality ($\mu = 0$), and where all individuals disperse ($p = 1$), shown in Figure 1 (dashed line) in year $t = 0$. In the next year ($t = 1$),

8

in the absence of other factors, the population would spread further under a more dispersive strategy than under a less dispersive one (Figure 1, dark grey solid line vs light grey solid line). However, in the presence of an Allee effect, the low-density edge of the population is not viable; the edge of the viable population is defined by where the Allee threshold (Figure 1 horizontal dotted lines) intersects the population density (Figure 1 solid lines), marked in Figure 1 by the vertical dotted lines. For a small Allee threshold (a_{lo}) the population edge (e_{lo}) will occur where the more dispersive type is most abundant. In contrast, for a large Allee threshold (a_{hi}) the population edge (e_{hi}) will occur where the less dispersive type is most abundant. Dispersal mortality has a similar effect: more dispersive types by definition more often suffer from mortality during dispersal, which reduces their density at the population edge. Thus, dispersal mortality, by lowering the population density, alters where the population density edge intersects the Allee threshold. Overall then, the combination of the Allee threshold (a) and dispersal mortality (μ) determine where the population edge occurs.

Over time, one of two outcomes occurred in simulations with spatial sorting. When the population edge occurred where a more dispersive type was abundant (e.g., low enough μ for a given a), the most dispersive strategies became more abundant on the population edge (Figure 2a-c). But when the population edge occurred where a less dispersive type was abundant (e.g., high enough μ for a given a), the most dispersive strategies were removed from (selected out of) the population (Figure 2d-f). Thus, although the most dispersive individuals were always pushed to the population edge (via spatial sorting), these individuals were favored at the edge only under some conditions.

Which dispersal strategy was favored at the simulated population edge in turn scaled up to affect how fast the population spread (Figure 3). When higher dispersal strategies were favored at the edge, spatial sorting accelerated invasions, leading to faster spread than simulations without spatial sorting (Figure 3a). However, when lower dispersal strategies were favored at the edge (and higher dispersal strategies were removed from the population), spatial sorting decelerated invasions and led to slower spread than simulations without

9

spatial sorting (Figure 3b).

Which pattern emerged in simulations was determined by the specific values of the Allee threshold a and dispersal mortality μ. Simulations with (Figure 4a) and without (Figure 4b) spatial sorting both spread fastest for small Allee thresholds and low dispersal mortalities. In cases where a and μ were sufficiently low, more dispersive individuals were favored at the edge and population spreads faster with spatial sorting than without (Figure 4c-d). However, for slightly larger values of either a or μ, more dispersive individuals were disfavored at the edge and the population spread slower with spatial sorting than without (Figure 4c-d). Finally, if either a or μ were too big, the population did not spread (white regions in Figure 4). Thus, spatial sorting slowed down population spread on the 'edge of extinction' (Figure 4), i.e., when the population was on the edge of not being able to spread at all. However, in the absence of an Allee effect ($a = 0$), simulations with spatial sorting always spread faster. Our simulated results are quite robust: changing the number of dispersal strategies (τ), the type of dispersal kernel used (k), or initial conditions does not qualitatively change our results (Supporting Information Figure S1).

Our analytic results match these simulated results. When the population is close to extinction, the spread rate without spatial sorting is higher, but when the population is far from extinction, the spread rate with spatial sorting is higher (Figure 5). Thus, the comparison of the analytic spread rates with and without spatial sorting (Figure 5) shows the same pattern as with the simulation model (Figure 4c).

5   Discussion

Here we show that spatial sorting (of individuals by dispersal type) and spatial selection (with an Allee effect) combine in a feedback loop that can act either to accelerate or decelerate population spread; this finding contrasts with the current narrative in the literature of spatial sorting only as an accelerator of population spread. Specifically, we find that this feedback loop leads to slower population spread (compared to simulations without spatial sorting and thus without a feedback loop) only when there is a large enough

10

Allee threshold (Figure 4d). Spatial sorting leads to an accumulation of higher dispersers on the population edge, where they are unable to survive (due to the Allee effect) and are removed from the population, thus causing the overall rate of population spread to slow down over time. These results answer the recent call by Miller *et al.* (2020) for theory to understand under what conditions Allee effects override positive effects of spatial sorting on dispersal.

Our findings provide important nuance to the interaction between spatial sorting and spatial selection that was overlooked by past theory. Indeed, existing models of dispersal evolution during population spread have often assumed no Allee effects (Phillips *et al.*, 2008; Travis *et al.*, 2009; Burton *et al.*, 2010; Bénichou *et al.*, 2012; Deforet *et al.*, 2019) despite their ubiquity in biological populations. Consistently building models with same assumptions limits our ability to understand biological phenomena (Shaw, 2022); a diversity of modeling assumptions brings greater understanding than each alone can provide (Levins, 1966). As a result of the theory, we have ended up with a framing in the literature that faster population spread will always arise either from spatial sorting alone or in the feedback loop formed with spatial selection. Moving beyond this narrow view requires separately considering the processes of spatial sorting and spatial selection and the diversity of ways they can combine in feedback loops.

We suggest that a broader definition of 'spatial selection' be adopted, based on our findings. To date, spatial sorting has been used as meaning that selection benefits individuals at the low-density edge of the population. The logic here is that individuals benefit at low density, population density varies spatially, and so selection varies spatially. With an Allee effect, selection no longer benefits low-density individuals. However, since Allee effects are felt most acutely at low density, and population density varies spatially, Allee effects should thus be considered a form of spatial selection, although in the opposite direction of that traditionally considered in the literature. Thus, we argue that it is more intuitive to define spatial selection broadly as cases where selection varies spatially, regardless of in which direction.

Our results also tie in to past theory on life history tradeoffs and spreading populations. We find that

11

spatial sorting slows population spread when being at the low-density population edge is sufficiently bad for individuals (i.e., a high Allee threshold). One could also imagine that spatial sorting might slow population spread if getting to the low-density edge is sufficiently costly. This could be captured by a mortality cost to dispersal (as in our model); the process of dispersal itself is costly, in terms of energy, time, risk, and opportunity (Bonte *et al.*, 2012). The costs of dispersal can scale with distance traveled, such that dispersing further comes with higher cost (Rousset & Gandon, 2002; Johnson *et al.*, 2009). Alternatively, this could be captured by a tradeoff between dispersal and other life history traits like fecundity and competitive ability (as in Burton *et al.* 2010; Deforet *et al.* 2019; Ochocki *et al.* 2020).

Intriguingly, the outcome that spatial sorting and spatial selection can slow the rate of population spread has not, to our knowledge, been reported in any empirical study (Miller *et al.*, 2020). There are several reasons why this may be the case. First, survivorship bias: if evolutionary processes substantially slow down an invasion, they may prevent the population from spreading altogether. Failed invasions are a common outcome that we still are unable to understand or predict. Indeed, in our model we were most likely to see that spatial sorting slowed down populations that were already spreading quite slow and near the 'edge of extinction'; perhaps empirical systems are pushed over this edge by other factors. Second, parameter values for lab systems. All lab studies that have disrupted spatial sorting directly have found that doing so slows down population spread (Williams *et al.*, 2016; Ochocki & Miller, 2017; Szűcs *et al.*, 2017; Weiss-Lehman *et al.*, 2017). Most of these systems used have 'weedy' life histories and thus reside far from the 'edge of extinction' and in parameter space where spatial sorting should indeed typically speed up spread. Similarly, many of the sources of dispersal mortality present in field conditions (e.g., predation) are absent from lab systems. Third, mechanisms: it is challenging to control for de novo mutation in field studies; a factor we excluded from our model. Our results suggest that if spatial sorting can indeed slow down population spread in empirical systems, it would do so in those that have slower growing life histories, a large Allee threshold, and/or high dispersal mortality.

12

Spatial sorting may indeed be close to universal, but it will not always lead to faster invasions over time when combined with spatial selection, as we show here. Our focus has been evolution from standing variation; future theory is needed to understand how the feedback loops between spatial sorting and spatial selection are affected when mutation is included and there is de novo evolution. Future empirical work targeting systems with Allee effects and high dispersal mortality are needed to determine whether evolutionary feedbacks that lead to slower spread are seen in empirical systems. Finally, we call for future studies to consider spatial selection as capturing the idea that selection can vary spatially across a population, rather than just the idea that selection favors individuals at the low-density population edge.

6 Acknowledgments

References

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

Burton, O.J., Phillips, B.L. & Travis, J.M.J. (2010). Trade-offs and the evolution of life-histories during range expansion. *Ecology Letters*, 13, 1210–1220.

Bénichou, O., Calvez, V., Meunier, N. & Voituriez, R. (2012). Front acceleration by dynamic selection in Fisher population waves. *Physical Review E*, 86, 041908.

Courchamp, F., Clutton-Brock, T.H. & Grenfell, B.T. (1999). Inverse density dependence and the Allee effect. *Trends in Ecology and Evolution*, 14, 405–410.

Cwynar, L.C. & MacDonald, G.M. (1987). Geographical variation of lodgepole pine in relation to population history. *The American Naturalist*, 129, 463–469.

Deforet, M., Carmona-Fontaine, C., Korolev, K.S. & Xavier, J.B. (2019). Evolution at the edge of expanding populations. *The American Naturalist*, 194, 291–305.

Dobzhansky, T. (1973). Nothing in biology makes sense except in the light of evolution. *The American Biology Teacher*, 35, 125–129.

Edmonds, C.A., Lillie, A.S. & Cavalli-Sforza, L.L. (2004). Mutations arising in the wave front of an expanding population. *Proceedings of the National Academy of Sciences*, 101, 975–979.

Gascoigne, J., Berec, L., Gregory, S. & Courchamp, F. (2009). Dangerously few liaisons: a review of mate-finding Allee effects. *Population Ecology*, 51, 355–372.

Hanski, I., Breuker, C.J., Schöps, K., Setchfield, R. & Nieminen, M. (2002). Population history and life history influence the migration rate of female Glanville fritillary butterflies. *Oikos*, 98, 87–97.

Johnson, C.A., Fryxell, J.M., Thompson, I.D. & Baker, J.A. (2009). Mortality risk increases with natal dispersal distance in American martens. *Proceedings of the Royal Society B: Biological Sciences*, 276, 3361–3367.

Klopfstein, S., Currat, M. & Excoffier, L. (2006). The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution*, 23, 482–490.

14

298 Korolev, K.S. (2015). Evolution arrests invasions of cooperative populations. *Physical Review Letters*, 115, 208104.

300 Kot, M., Lewis, M.A. & van den Driessche, P. (1996). Dispersal data and the spread of invading organisms. *Ecology*, 77, 2027–2042.

302 Kubisch, A., Fronhofer, E.A., Poethke, H.J. & Hovestadt, T. (2013). Kin competition as a major driving force for invasions. *The American Naturalist*, 181, 700–706.

304 Levins, R. (1966). The strategy of model building in population biology. *American Scientist*, 54, 421–431.

Lutscher, F. (2019). *Integrodifference Equations in Spatial Ecology.* vol. 49 of *Interdisciplinary Applied*
306 *Mathematics*. Springer, Cham, Switzerland.

Miller, T.E.X., Angert, A.L., Brown, C.D., Lee-Yaw, J.A., Lewis, M., Lutscher, F., Marculis, N.G., Mel-
308 bourne, B.A., Shaw, A.K., Szűcs, M., Tabares, O., Usui, T., Weiss-Lehman, C. & Williams, J.L. (2020). Eco-evolutionary dynamics of range expansion. *Ecology*, 101, e03139.

310 Ochocki, B.M. & Miller, T.E.X. (2017). Rapid evolution of dispersal ability makes biological invasions faster and more variable. *Nature Communications*, 8, 14315.

312 Ochocki, B.M., Saltz, J.B. & Miller, T.E.X. (2020). Demography-Dispersal Trait Correlations Modify the Eco-Evolutionary Dynamics of Range Expansion. *The American Naturalist*, 195, 231–246.

314 Perkins, T.A., Boettiger, C. & Phillips, B.L. (2016). After the games are over: life-history trade-offs drive dispersal attenuation following range expansion. *Ecology and Evolution*, 6, 6425–6434.

316 Perkins, T.A., Phillips, B.L., Baskett, M.L. & Hastings, A. (2013). Evolution of dispersal and life history interact to drive accelerating spread of an invasive species. *Ecology Letters*, 16, 1079–1087.

318 Phillips, B.L. (2015). Evolutionary processes make invasion speed difficult to predict. *Biological Invasions*, 17, 1949–1960.

Phillips, B.L., Brown, G.P., Travis, J.M.J. & Shine, R. (2008). Reid's paradox revisited: the evolution of dispersal kernels during range expansion. *The American Naturalist*, 172, S34–S48.

Rousset, F. & Gandon, S. (2002). Evolution of the distribution of dispersal distance under distance-dependent cost of dispersal. *Journal of Evolutionary Biology*, 15, 515–523.

Shaw, A.K. (2022). Diverse perspectives from diverse scholars are vital for theoretical biology. *Theoretical Ecology*, 15, 143–146.

Shaw, A.K. & Kokko, H. (2015). Dispersal evolution in the presence of Allee effects can speed up or slow down invasions. *American Naturalist*, 185, 631–639.

Shine, R., Brown, G.P. & Phillips, B.L. (2011). An evolutionary process that assembles phenotypes through space rather than through time. *Proceedings of the National Academy of Sciences*, 108, 5708–5711.

Stephens, P.A. & Sutherland, W.J. (1999). Consequences of the Allee effect for behaviour, ecology and conservation. *Trends in Ecology and Evolution*, 14, 401–405.

Szűcs, M., Vahsen, M.L., Melbourne, B.A., Hoover, C., Weiss-Lehman, C. & Hufbauer, R.A. (2017). Rapid adaptive evolution in novel environments acts as an architect of population range expansion. *Proceedings of the National Academy of Sciences*, 114, 13501–13506.

Travis, J.M.J. & Dytham, C. (2002). Dispersal evolution during invasions. *Evolutionary Ecology Research*, 4, 1119–1129.

Travis, J.M.J., Mustin, K., Benton, T.G. & Dytham, C. (2009). Accelerating invasion rates result from the evolution of density-dependent dispersal. *Journal Of Theoretical Biology*, 259, 151–158.

Weiss-Lehman, C., Hufbauer, R.A. & Melbourne, B.A. (2017). Rapid trait evolution drives increased speed and variance in experimental range expansions. *Nature Communications*, 8, 14303.

Williams, J.L., Kendall, B.E. & Levine, J.M. (2016). Rapid evolution accelerates plant population spread in

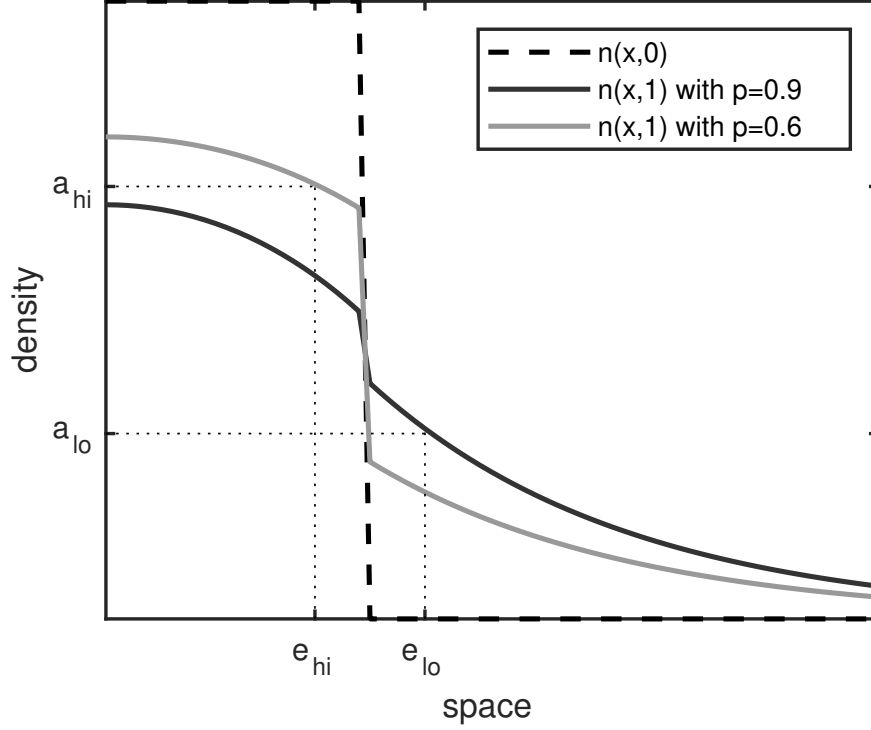342   fragmented experimental landscapes. *Science*, 353, 482–485.

Figure 1: Schematic showing the how the Allee effect shapes the population edge. The dashed line shows population distribution in year $t = 0$, and the solid lines shows what the population distribution would be in year $t = 1$ under two different dispersal strategies, a higher dispersal probability ($p = 0.9$, dark grey) and a lower dispersal probability ($p = 0.6$, light grey). The Allee threshold a (horizontal dotted lines) determines the edge of the viable population e (vertical dotted lines). Two scenarios are shown: for Allee threshold a_{lo}, high dispersers are more abundant at the population edge e_{lo}, but for Allee threshold a_{hi}, low dispersers are more abundant at the population edge e_{hi}. Both have no dispersal mortality ($\mu = 0$).
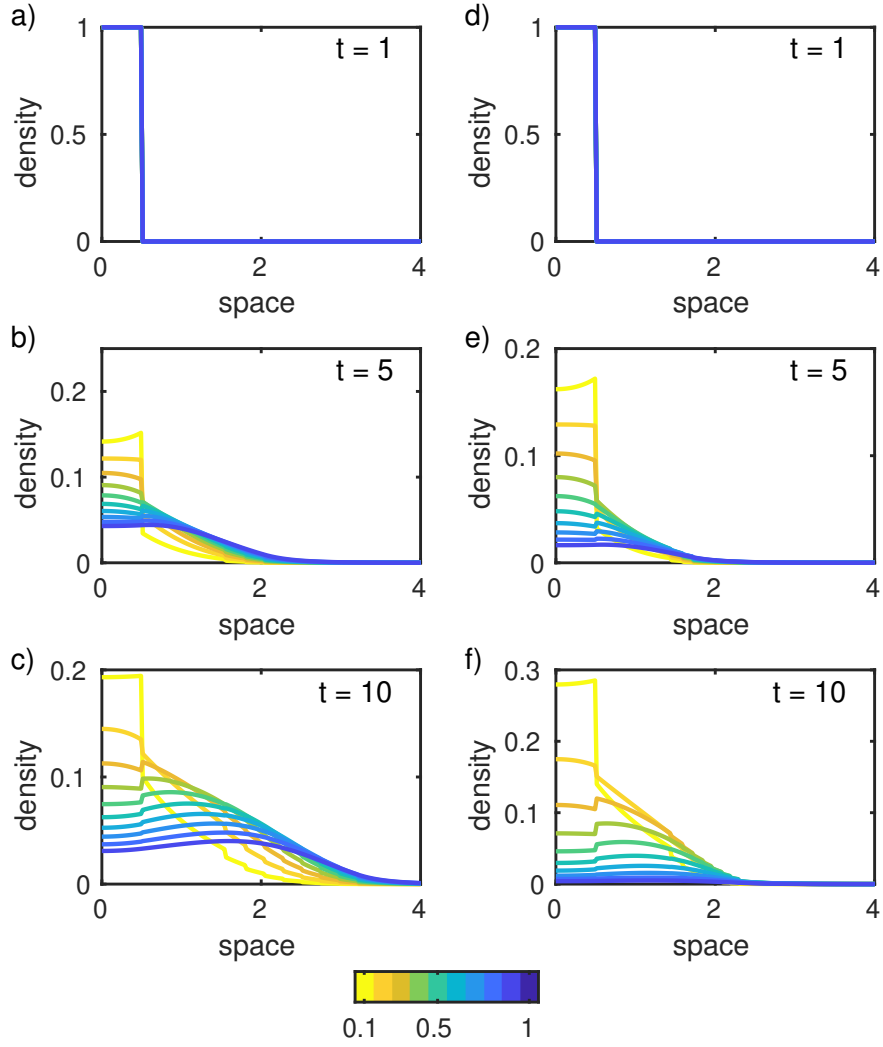
Figure 2: Simulations. Two examples of how the distribution of dispersal strategies (shown immediately after dispersal) across space changes over time ($t = 1, 5, 10$) for simulations with spatial sorting with Allee threshold $a = 0.02$ and (a-c) low dispersal mortality ($\mu = 0.2$) and (d-f) high dispersal mortality ($\mu = 0.4$). Colors indicate the fraction dispersing from 0.1 (yellow) to 1 (dark blue).
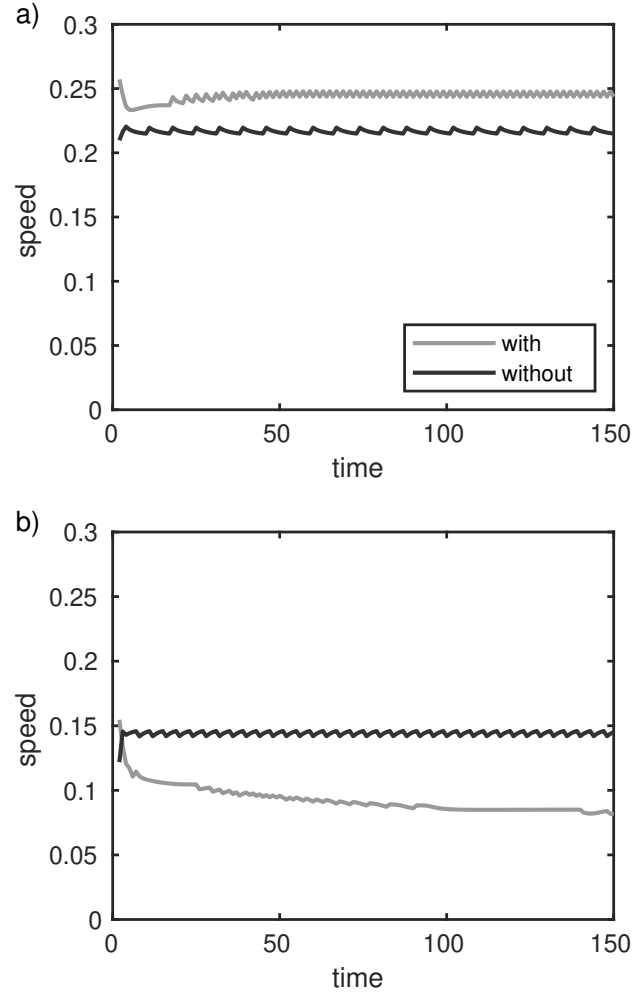
19

Figure 3: Simulations. The instantaneous rate of population spread over time for simulations with (light grey) and without (dark grey) spatial sorting for two cases, where the presence of spatial sorting (a) speeds up spread and (b) slowing down spread. Parameters: $a = 0.02$, (a) $\mu = 0.2$ and (b) $\mu = 0.4$.
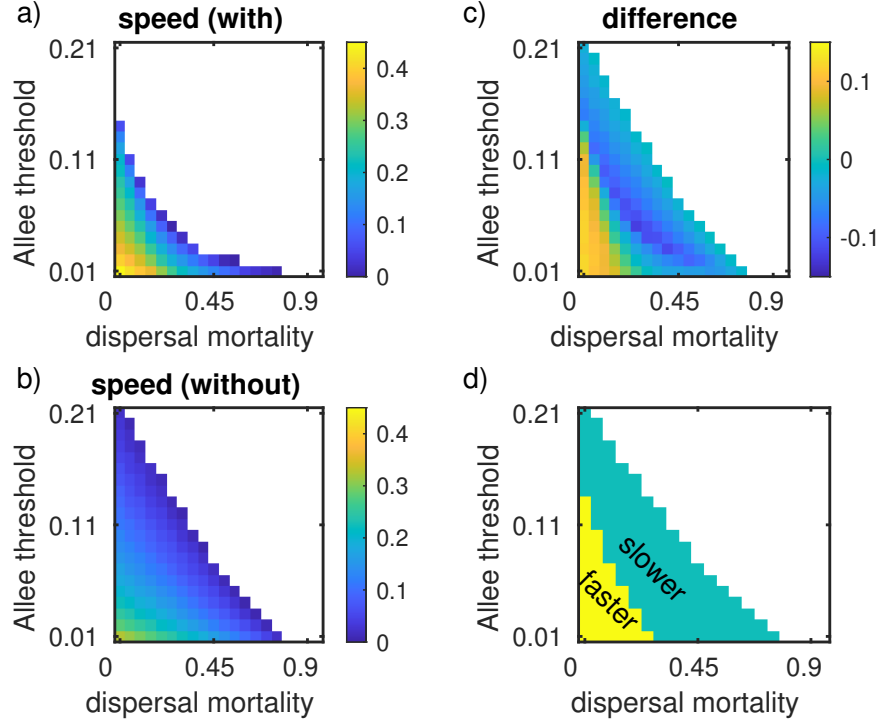
Figure 4: Simulations. The rate of population spread for simulations (a) with and (b) without spatial sorting, as a function of dispersal mortality (μ; x-axes) and Allee threshold (a; y-axes); white regions indicate where the populations failed to spread. The (c) difference in spread rate for simulations with spatial sorting minus without spatial sorting, and (d) the overall effect of spatial sorting on spread rate.
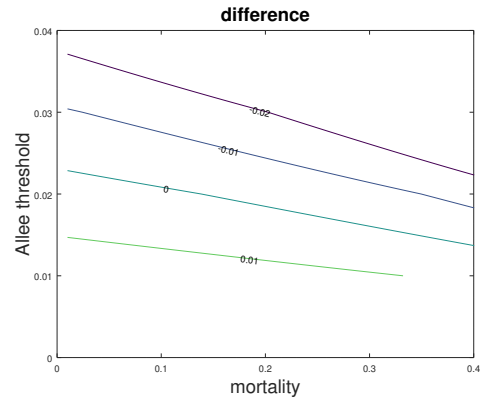
Figure 5: Analytic. The analytic difference in spread rate with spatial sorting minus without spatial sorting. Negative values indicate that spread rate without spatial sorting is faster. Parameters are $p_1 = 0.6$, $p_2 = 0.7$ and $v = 0.25$, which gives $b = \sqrt{v/2} \approx 0.35355$.
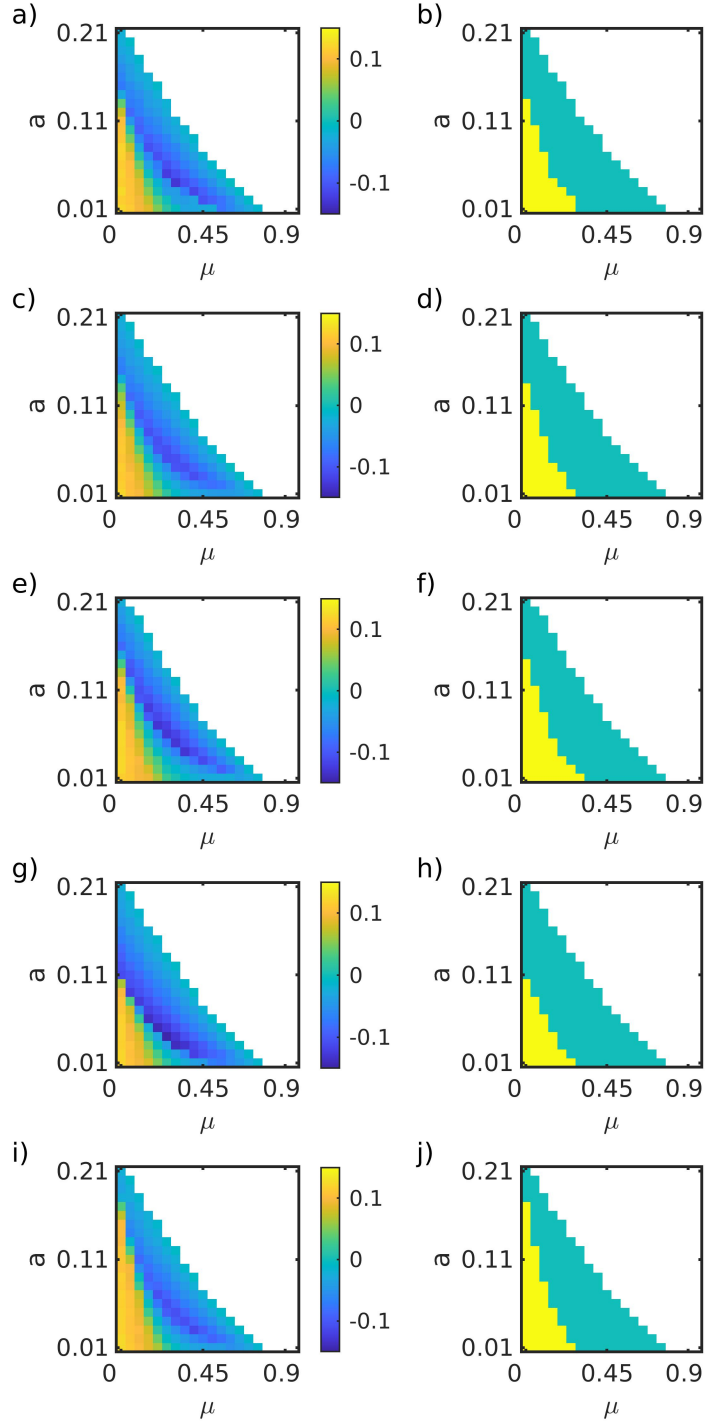
Supporting Information

Figure S1: Caption on the next page.

Figure S1: (Previous page.) Parameter variants – same as Fig. 4c-d but with different parameter values. The (left column) difference in spread rate for simulations with spatial sorting minus without spatial sorting, and (right column) the overall effect of spatial sorting on spread rate, as a function of dispersal mortality (μ; x-axes) and Allee threshold (a; y-axes). White regions indicate where the populations failed to spread. In the right column, yellow indicates where spatial sorting speeds up spread, teal indicates where spatial sorting slows down spread. Parameters: (a-b) half as many dispersal types ($\tau = 5$), (c-d) twice as many dispersal types ($\tau = 20$), (e-f) a different (gaussian) dispersal kernel, (g-h) different initial conditions: skewed towards the lowest dispersal type $n_1(x,t) = 5$ and $n_i(x,t) = 5/9$ for for $i = 2, ...\tau$, $|x| < 0.5$; 0 otherwise, (i-j) different initial conditions: skewed towards the highest dispersal type $n_{10}(x,t) = 5$ and $n_i(x,t) = 5/9$ for for $i = 1, ...\tau - 1$, $|x| < 0.5$; 0 otherwise.

Table S1: Model variables and parameters, meaning, and default values for simulations (where applicable).

	Meaning	Default value
a	Allee threshold	varied
b	density-dependence parameter	1
f	growth function	eqn. 1
g	density-dependence function	eqn. 2
i	dispersal strategy	$i = 1, ..., \tau$
k	dispersal kernel function	eqn. 4
p_i	proportion of individuals with strategy i that disperse	$0 \leq p_i \leq 1$
v	variance of dispersal kernel	0.25
t	time (year)	-
x	space	-
y	space	-
N	population density	-
λ	growth rate	2
μ	dispersal mortality	varied
τ	number of dispersal strategies	10

Derivation and analysis of the analytical approximation

In this section, we give a detailed derivation and description of the analytical approximation of our model,

346 and we present the analytical results.

Spread models with Allee effect are notoriously difficult to analyze, and explicit results are almost never

348 available. A notable exception is the integrodifference model in Kot *et al.* (1996). The authors consider

a step function to model the Allee effect: the population density in the next year is at carrying capacity

350 (resp. at extinction) if it is above (resp. below) the Allee threshold in the current year. Their model can be

explicitly solved and the speed of a spreading population can be determined by using the cumulative density

352 function of the dispersal kernel (Lutscher, 2019). We begin by showing that the model by Kot *et al.* (1996)

can be understood in terms of a time-scale separation and then use the ideas in Lutscher (2019) to calculate

354 the speed for our extended model.

The case without spatial sorting

356 When the offspring dispersal strategy is uniformly distributed, independent of the parental strategy, the

reproduction function f in the IDE Eq. (3) (main text) depends only on the total density N. We can rescale

358 Eq. (2b) to read

$$f(N) = \begin{cases} \frac{1}{\tau} \frac{RN}{1+(R-1)N/K} & N \geq a \\ \\ 0 & N < a, \end{cases} \tag{S1}$$

where $a > 0$ is again the Allee threshold, R is a growth rate and K the carrying capacity. We further scale

360 $K = 1$ and let R tend to infinity. Then the function becomes the step function

$$f(N) = \begin{cases} \frac{1}{\tau} & N \geq a \\ \\ 0 & N < a, \end{cases} \tag{S2}$$

where $a > 0$ is still the Allee threshold.

362 We first study the IDE in Eq. (3) with step function in Eq. (S2) for a single type ($\tau = 1$), which we

denote by $n = n_1 = N$. If all individuals disperse ($p = 1$), this is exactly the case in Kot *et al.* (1996);

27

Lutscher (2019). It is useful to introduce the density after the reproduction phase, $\hat{n} = f(n)$. This density satisfies the equation

$$\hat{n}(x, t+1) = f((1-p)\hat{n} + p(1-\mu)k * \hat{n}). \tag{S3}$$

We use the shorthand notation $*$ for the convolution integral in (3).

We begin with the step function $\hat{n}(x, 0)$, which is equal to 1 for $x \leq 0$ and equal to zero for $x > 0$. We calculate

$$k * \hat{n} = 1 - F(x), \tag{S4}$$

where F is the cumulative density function of k. In particular, F is a non-decreasing function, and, hence, $1 - F$ is a non-increasing function. Therefore,

$$\tilde{n} = (1-p)\hat{n} + p(1-\mu)k * \hat{n} \tag{S5}$$

is also a non-increasing function. Hence, there exists a largest value $\tilde{x}$ where $\tilde{n}(x) \geq a$. This implies that $n(x, t+1) = f(\tilde{n})$ is again a step function. Hence, if we start with a step function in one year, then the population density remains a step function in following years. It turns out that we can calculate how far the front moves in one year (namely $\tilde{x}$), which is the precisely the speed that we are interested in (which we denote by c). The calculations are only a slight extension of those given when $\mu = 0$ and $p = 1$ in Lutscher (2019).

Lemma 1 *If*

$$p(1-\mu) > 2a \tag{S6}$$

then there is a traveling wave with positive speed c (distance per year), which is given implicitly by

$$F(c) = 1 - \frac{a}{p(1-\mu)}, \tag{S7}$$

where F is the cumulative density function of k. If (S6) is reversed, the population will not spread. In that case, if $(1-p) > a$, the population will not retreat, i.e., there is a pinned wave of speed zero, otherwise the population will retreat in a wave with negative speed.

28

The reasoning above extends directly to the case of two or more dispersal strategies ($\tau \geq 2$) with movement probabilities p_i. We obtain the following result.

Lemma 2 *If the average dispersal strategy $\bar{p} = \frac{1}{\tau} \sum_{i=1}^{\tau} p_i$ satisfies*

$$\bar{p}(1 - \mu) > 2a \tag{S8}$$

then there is a traveling wave with positive speed c (distance per year), which is given implicitly by

$$F(c) = 1 - \frac{a}{\bar{p}(1 - \mu)}, \tag{S9}$$

where F is again the cumulative density function of k.

For an explicit example, we consider the Laplace kernel in Eq. (4). Its cumulative density function for $x > 0$ is $F(x) = 1 - \frac{1}{2}\exp(-\sqrt{2/v}\,x)$. The explicit formula for the speed then becomes

$$c = \sqrt{\frac{v}{2}} \ln\left(\frac{\bar{p}(1 - \mu)}{2a}\right). \tag{S10}$$

We illustrate the profile of an advancing population with 10 types in Fig. S2.

Remarks.

1. The statements in the two lemmas hold under very general assumptions on the dispersal kernel: it has to be symmetric and integrable. It does not have to be exponentially bounded.

2. As is usual with strong Allee effects, if the initial density of a population is too low and/or the initial spatial extent is too small, then the population will not spread but go extinct. Here, we always choose initial conditions that spread spatially.

3. The two lemmas can be generalized considerably in that the different types can have different dispersal-induced mortality (i.e., we can replace μ by a different μ_i for each type). Another possible extension is that rather than drawing the dispersal strategy from a uniform distribution, it can be drawn from
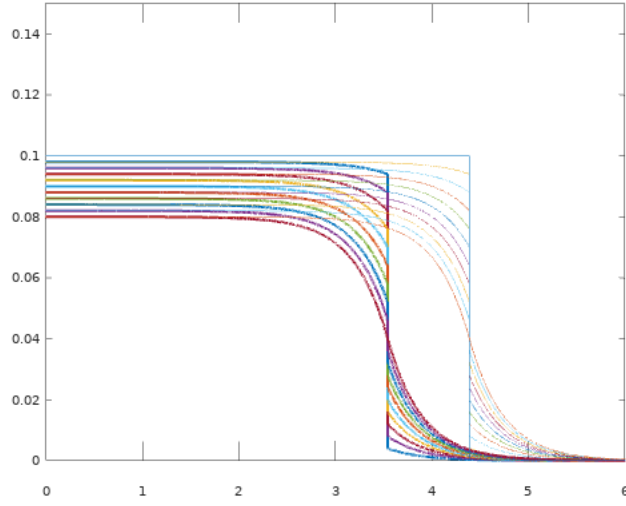
29

Figure S2: The advance of a population of ten types in one year. Strong colours show year t, weaker colours show year $t+1$. The lowest density behind the front corresponds to the highest profile ahead of the front and belongs to the strategy with the highest p_i (here $p_{10} = 1$). The kernel is the Laplace kernel with $v = 0.25$. The initial condition has each type uniformly distributed on $[-1, 1]$. The ten types have $p_i = i/10$ for $i = 1, \ldots, 10$. Other parameters are $\mu = 0.2$ and $a = 0.02$. The resulting speed is $c = 0.848$.

any fixed distribution, i.e., a fraction m_i of offspring has dispersal strategy i in each year. The only change in the statement of Lemma 2 is that the average of p_i is replaced by the weighted average of $p_i(1 - \mu_i)$ with weights m_i.

The case with spatial sorting

With spatial sorting, we have a third time scale, namely the competition between types. We first describe the model assumptions verbally (see Figure S3), then we formalize it in equations. For simplicity, we consider only two types: a low disperser (red) and a high disperser (blue). Initially, both types are equally present for $x < 0$ and absent for $x > 0$ (top panel, dashed profiles). After dispersal (top panel, solid curves), the higher disperser has the higher density ahead of the original population extent and the lower density behind. The new extent of the population in year 1 (vertical line at c_1) is given where the combined density of the two types exceeds the Allee threshold (second panel). Between the old and the new extent, the population grows to carrying capacity immediately and the relative frequency of the two types after reproduction is the same as after dispersal (lottery competition). Throughout the old extent ($x < 0$), the lower disperser (red line) wins the competition since fewer of its individuals disperse; see e.g. Perkins *et al.* (2016). After the subsequent dispersal phase, the high disperser has again the higher density ahead of c_1 (solid curves, third panel). As before, the combined density determines the new extent ($c_1 + c_2$) in year 2 (bottom panel). Behind c_1, the low disperser takes over. In the newly occupied region between c_1 and $c_1 + c_2$, the total density is at carrying capacity whereas the frequency of the two types reflects that after dispersal.

Before we formulate the above verbal description in mathematical terms, we make one more simplifying assumption. To calculate the relative densities of the two types in the newly extended range, we take the relative density at the range edge as representative for the entire region (rather than taking it at every point in the region). With this, we are ready to formulate equations.

We denote the density of the low (high) disperser by $n_1(x,t)$ ($n_2(x,t)$). The initial condition (dashed
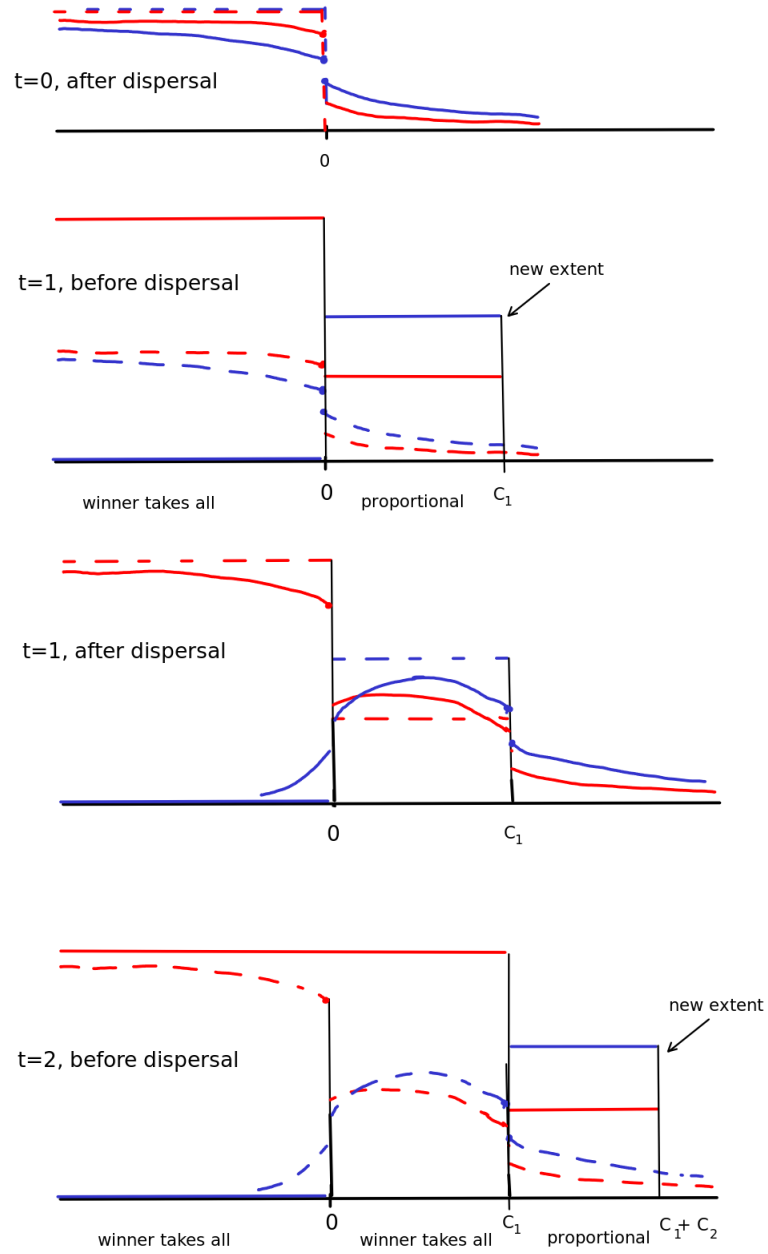
31

Figure S3: Schematic illustration of the model with spatial sorting for two types. In each panel, the solid (dashed) lines represent the current (preceding) densities of the low (red) and high (blue) disperser. In the first dispersal step, the population advances by c_1 space units, in the second by c_2. See text for details.

422 lines, top panel, Fig. S3) are

$$n_1(x,0) = n_2(x,0) = \begin{cases} 1/2, & x \leq 0, \\ \\ 0, & x > 0. \end{cases} \tag{S11}$$

After the dispersal phase (solid curves, same panel), the densities are given by

$$n_i(x,1) = (1 - p_i)n_i(x,0) + p_i(1 - \mu_i)(1 - F(x)), \tag{S12}$$

424 where F is again the cumulative density function of the dispersal kernel (identical for both types) and $*$

denotes the convolution integral. We calculate the new extent of the population exactly as before by finding

426 c_1 such that

$$n_1(c_1, 1) + n_2(c_1, 1) = a. \tag{S13}$$

The new extent in the first year, c_1, is given implicitly by

$$1 - F(c_1) = \frac{2a}{p_1(1 - \mu_1) + p_2(1 - \mu_2)}. \tag{S14}$$

428 When k is the Laplace kernel, we have the explicit expression

$$c_1 = \sqrt{\frac{v}{2}} \ln \left(\frac{p_1(1 - \mu_1) + p_2(1 - \mu_2)}{4a} \right). \tag{S15}$$

The percentage of high dispersers in the first year at the population edge is

$$H_1 = \frac{n_2(c_1, 1)}{n_1(c_1, 1) + n_2(c_1, 1)}. \tag{S16}$$

430 After the reproduction phase, the high disperser is present only in the new extent, whereas the low disperser

has taken over the previous extent and is proportionally present in the new extent (solid lines, second panel).

432 This can be expressed by using the indicator function $\chi_{(a,b]}$ (which is equal to 1 on $(a,b]$ and zero elsewhere)

as

$$f(n_1(x,1)) = \chi_{(-\infty,0]} + (1 - H_1)\chi_{(0,c_1]}, \quad f(n_2(x,1)) = H_1\chi_{(0,c_1]}. \tag{S17}$$

33

Now we look at the next dispersal phase. We only need to calculate the densities at locations $x > c_1$ to determine the advance in the second year, c_2. We find

$$n_1(x, 2) = p_1(1 - \mu_1)(1 - F(x)) + p_1(1 - \mu_1)(1 - H_1)(F(x) - F(x - c_1)),$$

$$n_2(x, 2) = p_2(1 - \mu_2)H_1(F(x) - F(x - c_1)).$$

From these expressions, we calculate the new extent, $c_1 + c_2$, and the fraction of high dispersers in the new extent, H_2 as before. If we use the Laplace kernel, we can obtain explicit expressions for the distance gained in each year (c_t) and the percentage of high dispersers (H_t) in terms of the previous year. After some tedious calculations, we find the following.

Lemma 3 *Let k be the Laplace kernel with parameter v, let c_t and H_t denote the distance advanced in the t-th year and the fraction of high dispersers at the front. Then we have the recursion equations*

$$c_{t+1} = \sqrt{\frac{v}{2}} \ln\left(\frac{p_1(1 - \mu_1) + [p_1(1 - \mu_1)(1 - H_t) + p_2(1 - \mu_2)H_t]\left(e^{c_t\sqrt{\frac{2}{v}}} - 1\right)}{2a} \right) \tag{S18}$$

and

$$H_{t+1} = \frac{p_2(1 - \mu_2)H(t)[F(c_{t+1} + c_t) - F(c_t)]}{p_1(1 - \mu_1)(1 - F(c_{t+1} + c_t)) + p_2(1 - \mu_2)H_t[F(c_{t+1} + c_t) - F(c_t)]}. \tag{S19}$$

These formulas look unwieldy, but they turn out to be much faster to simulate than the spatial system with the convolution integral and have some special properties that we summarize in the next lemma.

Lemma 4 *The updating functions in the previous lemma, i.e., the right-hand sides of (S18) and (S19) are monotone functions with respect to c_t and H_t. This implies that the solution of the recursion is monotone and, since it is also bounded, it converges to a fixed point, (c^*, H^*), given by the expressions*

$$E(1 - E) = \frac{2a}{p_2(1 - \mu_2)}, \qquad E = e^{-c^*\sqrt{\frac{2}{v}}} \tag{S20}$$

and

$$H^* = \frac{1 - \frac{p_1(1 - \mu_1)}{p_2(1 - \mu_2)} - p_1(1 - \mu_1)\frac{E^2}{2a}}{1 - \frac{p_1(1 - \mu_1)}{p_2(1 - \mu_2)}}. \tag{S21}$$

34

This result is remarkable for several reasons. First, it allows us to calculate explicitly the asymptotic speed c^* and the corresponding fraction of high dispersers at the front, H^*. Second, it says that the asymptotic speed c^* depends only on the movement behavior and mortality of the high disperser, not on that of the low disperser (since the equation for E does not contain parameters p_1 and μ_1). The caveat is that there can be two solutions for E, and therefore for c^*, because the equation is quadratic. However, in most simulations, only one of the two solutions for c^* has a positive value of H^* associated with it. That is the relevant one.

The analysis of the approximate model shows exactly the same qualitative behavior as the simulation model in the main text (Fig. S4). In analogy with Fig. 4 (a)–(c), we plot the speed with (top) and without (middle) sorting and their difference (bottom). As for the simulation model, we find that spatial sorting slows down range expansion near the extinction limit and speeds them up far from it. We conclude that these findings are robust with respect to model details.
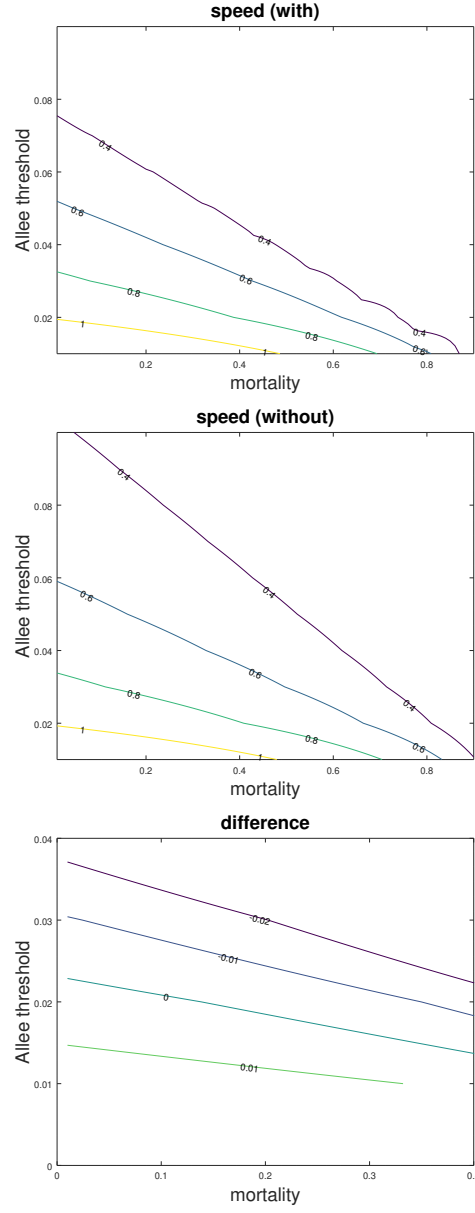
Figure S4: Comparing the speed with (top) and without (middle) sorting. The difference is plotted in the bottom panel. The analysis gives straight lines. Parameters are $p_1 = 0.6$, $p_2 = 0.7$ and $v = 0.25$, which gives $b = \sqrt{v/2} \approx 0.35355$.