Supporting Information for:

Pushed to the edge: Spatial sorting can slow down invasions

Authors:

Allison K. Shaw, 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

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

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

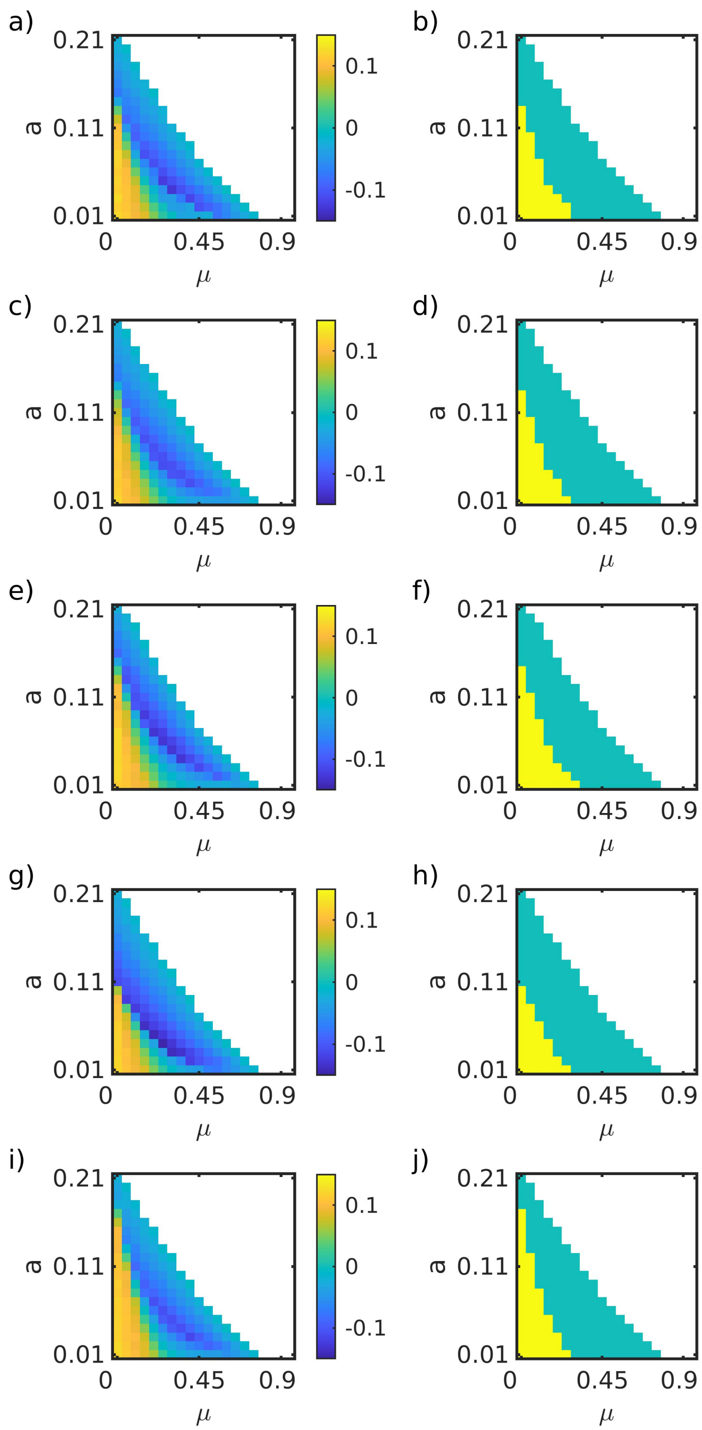

Figure S1: Caption on the next page.

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, and we present the analytical results.

Spread models with Allee effect are notoriously difficult to analyze, and explicit results are almost never 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 (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 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 the speed for our extended model.

The case without spatial sorting

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 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 $K = 1$ 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.

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); 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.

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

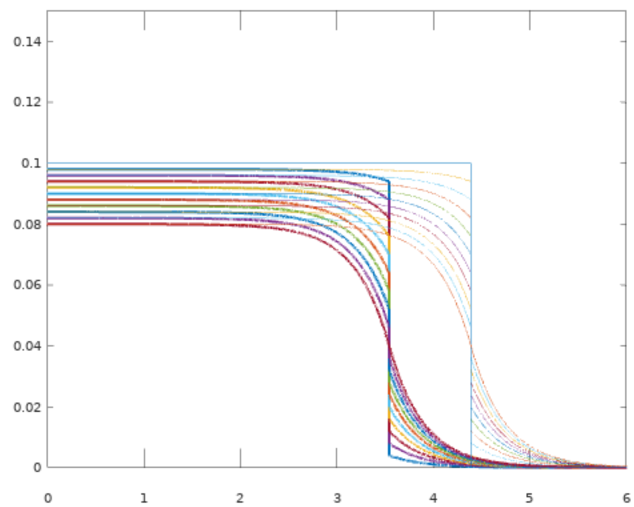

Figure S2: The advance of a population of ten types in one year. Strong colours show year t, weaker colours show year $t+1$. The lowest density behind the front corresponds to the highest profile ahead of the front and belongs to the strategy with the highest p_i (here $p_{10} = 1$). The kernel is the 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

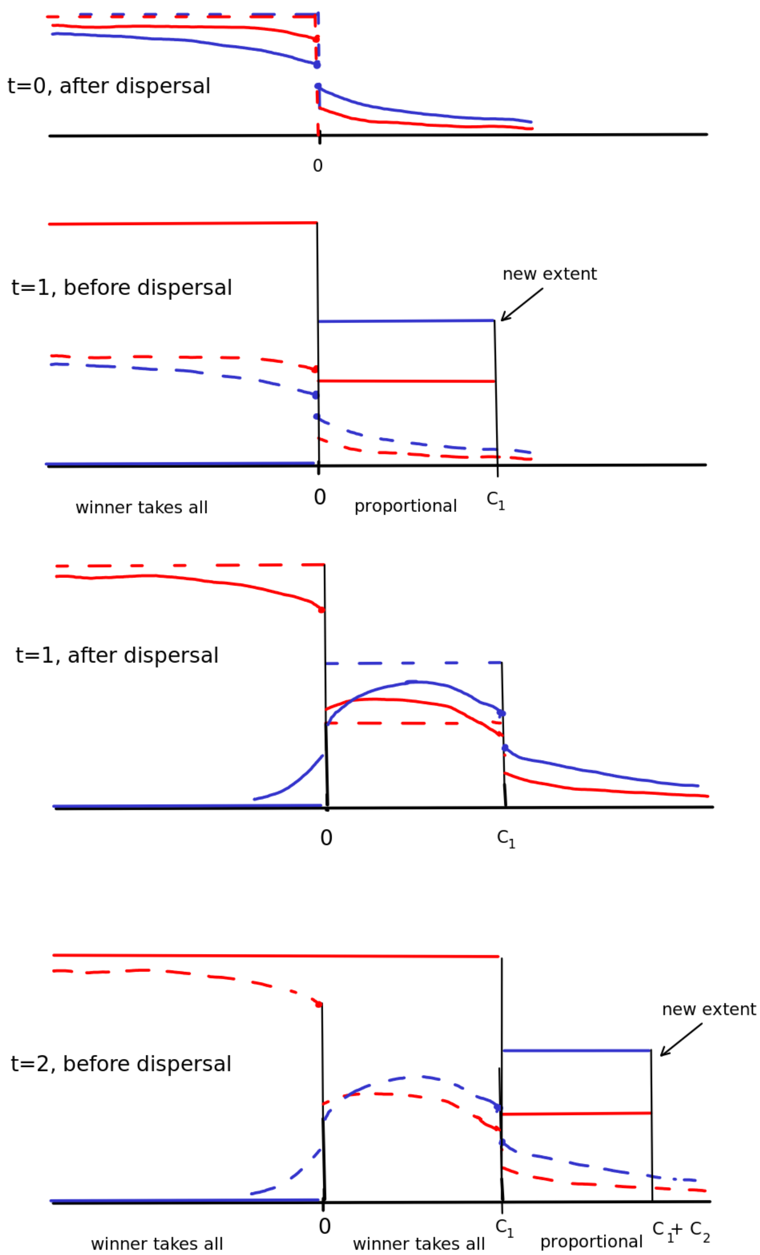

Figure S3: Schematic illustration of the model with spatial sorting for two types. In each panel, the solid (dashed) lines represent the current (preceding) densities of the low (red) and high (blue) disperser. In the first dispersal step, the population advances by c_1 space units, in the second by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

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

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

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 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

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

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

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

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

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

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

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

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.

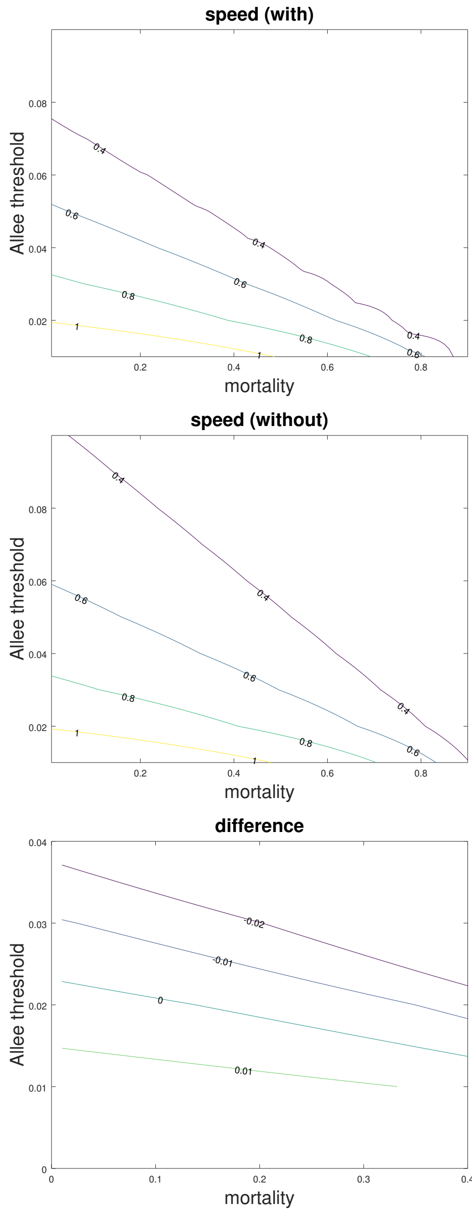

Figure S4: Comparing the speed with (top) and without (middle) sorting. The difference is plotted in the bottom panel. The analysis gives straight lines. Parameters are $p_1 = 0.6$, $p_2 = 0.7$ and $v = 0.25$, which gives $b = \sqrt{v/2} \approx 0.35355$.

References

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

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

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.