Density-dependent selection and the limits of relative fitness

Jason Bertram [1,*]

Joanna Masel [1]

1. Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721.

∗ Corresponding author; e-mail: jbertram@email.arizona.edu.

Keywords: Lottery model, Wright-Fisher model, r/K-selection, α-selection, eco-evolutionary synthesis.

Abstract

Evolutionary biologists frequently describe selection using relative fitness, and relative fitness models are foundational in population genetics. Yet the population ecological basis of relative fitness is poorly understood, and the classical ecology literature on selection in crowded populations seems to be incompatible with widely-used relative fitness models such as the Wright-Fisher. Here we develop a generalization of the Wright-Fisher model in which density depends dynamically on the demographic rates of the types present. We then explore the population ecology of relative-fitness using this model as a reference point. Although density-dependent selection causes relative fitness to break down, we emphasize that relative fitness models are fairly robust in populations close to demographic equilbrium. In particular, relative fitness only breaks down if selection is strong and density-dependent, and density is also selection-dependent. Our generalized Wright-Fisher model clearly distinguishes these effects, whereas the classical literature on density-dependent selection confounds them. We argue that selection-independent density is ecologically plausible for many species given the prevalence of reproductive excesses, an important aspect of selection omitted in many ecological models. Our model also offers a natural alternative to relative fitness when the latter is untenable, as is likely the case far from demographic equilbrium.

Introduction

There are a variety of different measures of fitness. Some widely used examples in evolutionary ecology are expected lifetime reproductive ratio R_0, intrinsic growth rate r, saturation population density (often labeled "K") [Benton and Grant, 2000], and invasion fitness [Metz et al., 1992]. In addition, "relative fitness" is the standard in much of evolutionary biology, particularly evolutionary genetics, where attention is generally restricted to relative genotypic proportions [Barton et al., 2007, pp. 468]. The variety of fitness measures is not problematic in itself, because different measures may be more useful in different circumstances. But it should be clear how the measure being used is connected to the processes of birth and death which govern population biology [Doebeli et al., 2017, Metcalf and Pavard, 2007]. While such a connection is fairly clear for absolute fitness measures like r, relative fitness seems largely divorced from population biology. It has even been proposed that relative fitness be justified from more abstract measure-theoretical arguments, abandoning population biology altogether [Wagner, 2010].

In uncrowded populations, relative fitness simply represents differences in intrinsic growth rate. In discrete time, the change in frequency of type i is $\Delta p_i = \left(\frac{W_i}{\overline{W}} - 1 \right) p_i$, where W_i is the intrinsic absolute growth factor of type i, and $\overline{W} = \sum_i W_i p_i$ is the population mean W. Here we can rescale $\overline{W}$ however we please and replace it with "relative fitness" $\overline{w}$ without affecting the ratio $\frac{W_i}{\overline{W}} = \frac{w_i}{\overline{w}}$. In continuous time, the canonical selection equation is $\frac{dp_i}{dt} = (r_i - \overline{r})p_i$, where $\overline{W}$ is replaced by the intrinsic exponential growth rate $\overline{r}$ [Crow et al., 1970, pp. 26]. If there are two types present, a wildtype and a mutant for instance, then the continuous time canonical selection equation can be written as

$$\frac{dp_i}{dt} = sp_i(1 - p_i), \tag{1}$$

where the constant selection coefficient s is the difference in r between types. The corresponding adaptive sweeps follow a logistic curve.

The difficulty with Eq. (1) arises in crowded populations. Since crowded and uncrowded conditions are so different, s will often depend on density. Eq. (1) is then no longer a com-

plete description of selection — we would also need to specify a model for density. Note that frequency-dependent selection does not raise similar problems; Eq. (1) is still a complete description of selection even if its behavior is more complicated compared to the constant-s case. Standard population genetics evades the issue of density-dependent selection by simply assuming that total population density N has reached its equilibrium value, which is assumed to be a fixed constant. The selection coefficient s now parameterizes the rate at which selection changes relative frequencies, but no longer corresponds to differences in intrinsic growth rates r.

Counter to the assumption of constant N, MacArthur famously showed that when population growth is density-regulated, selection in crowded populations is intimately connected to the ability to keep growing at higher densities than other types can tolerate [MacArthur and Wilson, 1967]. The classic example is the logistic model, where the type with the greatest saturation population density "K" excludes the others (Fig. 1a). Similarly, the "R^* rule", a central tenet of resource competition theory, states that when growth is limited by a single homogeneous consumable resource, the type able to deplete the resource to the lowest equilibrium density R^* excludes the others [Grover, 1997]. Differences in R^* will often entail differences in saturation density. The Lotka-Volterra competition model also couples selection in crowded populations to changes in N except in a few special cases (we return to this in section "K-selection versus relative fitness"). In these examples, both N and s change during, and as a result of, adaptive sweeps. It would therefore seem that the ubiquitous constant-N, relative fitness description of selection is incompatible with a huge class of population ecological processes driving selection (Fig. 1b).

The relative fitness description has been justified in broadly two different ways for crowded populations (we do not discuss Wagner's [2010] measure-theoretical justification, which is explicitly independent of population biology and thus falls outside of our scope). The first is to assume that selection is independent of density but still allow density to be affected by selection [Barton et al., 2007, pp. 468] [Prout, 1980]. While this allows us to relax the assumption of constant N, it does not address the problem that s can, in reality, depend on density. In the examples from the
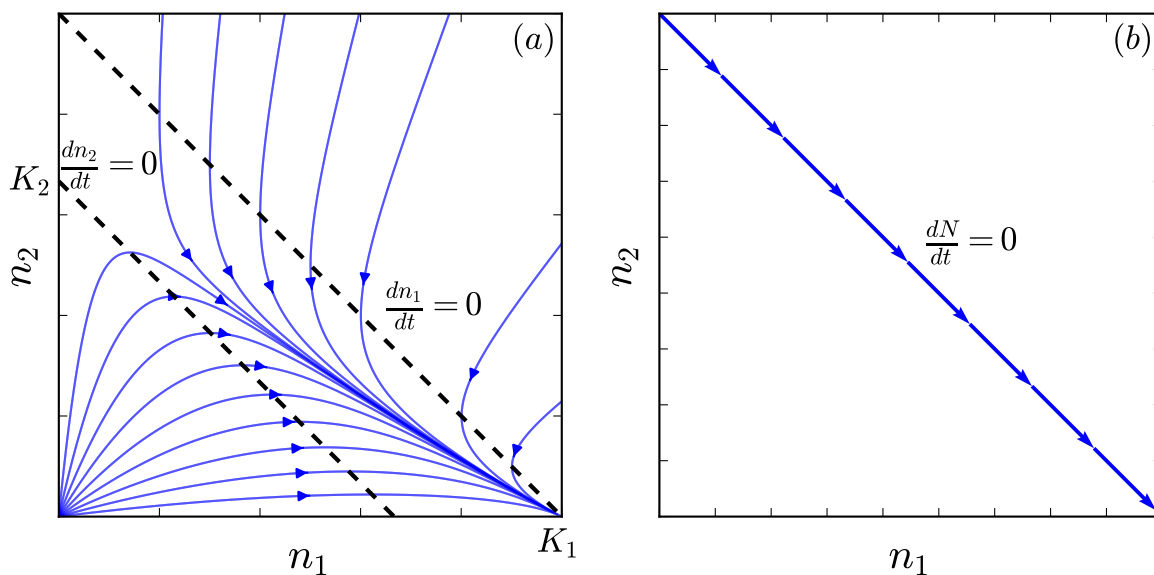
Figure 1: Selection in crowded environments shown as a phase diagram for the densities of two types n_1 and n_2. (a) The logistic model $\frac{dn_1}{dt} = r_1(1 - \frac{n_1+n_2}{K_1})n_1$ and $\frac{dn_2}{dt} = r_2(1 - \frac{n_1+n_2}{K_2})n_1$ with $r_1 = r_2$ and $K_1 > K_2$. (b) The constant-N, relative fitness description of selection.

previous paragraph, selection is density-dependent; indeed, the type-specific responses to density are at the center of MacArthur's argument and the density-dependent selection literature that grew out of it [Roughgarden, 1979].

Second, constant N and s can both be seen as a short-term linear approximation [Ewens, 2004, pp. 277]. That is, within a sufficiently short time frame, N and s can be treated as constant. Provided that selection is sufficiently weak and stable over time, this "short-term" assumption is not a major restriction. Yet it is increasingly recognized that selection is not always weak, that it can fluctuate considerably over time, and that N can vary by orders of magnitude over a few generations as a routine feature of a population's ecology [Messer et al., 2016]. These are not rare exceptions, but occur widely in nature and the lab, including in wild *Drosophila* [Bergland et al., 2014]. Nevertheless, relative fitness models are the foundation for much of the population genetic literature, and are still widely used without considering the "short-term" restriction or the lack of integration with population ecology [Mallet, 2012]. Thus, it is important to understand the population ecological basis of relative fitness models, both to gain insight into their domain of

5

applicability, and as part of the broader challenge of synthesizing ecology and evolution.

Another issue with the constant-N relative fitness description of selection is that it precludes consideration of longer-term aspects of the interplay between evolution and ecology such as population extinction. [will make this sound less dismissive] Adaptive dynamics currently provides a powerful framework for addressing the complex feedbacks between evolutionary change and population density [Diekmann et al., 2004]. However, the focus of adaptive dynamics is trait evolution rather than the underlying genetics, and in particular, selective sweeps are typically subsumed into effectively-instantaneous "trait substitutions". We emphasize that our focus here is the description of the time-dependent process of selection itself, which is particularly critical for making sense of evolution at the genetic level for which we now have abundant data.

Here we analyze the population ecology of relative fitness using a novel model of density-dependent population growth based on territorial contests. Rather attempting to make sense of relative fitness in the standard models of population growth mentioned above (e.g. Mallet [2012]), we instead do the reverse, and attempt to make population ecological sense of the widely-used Wright-Fisher, constant-N, relative fitness model. Our starting point is the classic lottery model of territorial contest [Chesson and Warner, 1981, Sale, 1977]. Like the Wright-Fisher model, the classic lottery assumes a saturated population with constant N, and fitness involves a product of fertility and juvenile viability [Crow et al., 1970, pp. 185]. Unlike the Wright-Fisher model, generations can overlap in the lottery model. We generalize the lottery model to allow population density to vary, and show that this model can be interpreted as a density-dependent generalization of the Wright-Fisher model with overlapping generations.

We show that when this model reaches a demographic steady-state, the constant-N, relative fitness picture emerges. Futhermore, we show that our model is entirely consistent with MacArthur's analysis of selection in crowded populations. In particular, we emphasize that MacArthur's argument does not justify the widespread intuition that selection in crowded environments is necessarily connected to achieving greater densities. This is largely an artifact of the models historically used in the density-dependent selection literature, which ignore relative

111 contests.

112   Our first task is to analytically extend the classic lottery model to correctly account for low

113 density behavior (sections). We then...

Model

Model assumptions and definitions

116 We assume that reproductively mature individuals ("adults") each require their own territory to

117 survive and reproduce (Fig. 2). All territories are identical, and the total number of territories is

118 T. Time t advances in discrete iterations, each representing the time from birth to reproductive

119 maturity. In iteration t, the number of adults of the i'th type is $n_i(t)$, the total number of adults

120 is $N(t) = \sum_i n_i(t)$, and the number of unoccupied territories is $U(t) = T - N(t)$.

121   We assume that the n_i are large enough that stochastic fluctuations in the n_i ("drift") can

122 be ignored. In particular, we do not evaluate the initial stochastic behaviour of mutant lineages

123 while they are at low abundance. We derive equations for the expected change in the n_i over

124 time, leaving the evaluation of drift for future work.

125   Each iteration, adults produce new offspring ("propagules"), m_i of which disperse to un-

126 occupied territories. We assume that adults cannot be ousted from their territories, so that m_i

127 only includes propagules landing on unoccupied territories. Propagules disperse at random over

128 the unoccupied territories, regardless of distance from their parents, and independently of each

129 other. There is no interaction between propagules (e.g. avoidance of territories crowded with

130 propagules). Loss of propagules during dispersal is subsumed into m_i. We assume that each

131 adult produces a constant number b_i of successfully dispersing propagules; the loss of propag-

132 ules due to dispersal to occupied territories then implies $m_i = b_i(1 - N/T)n_i$. Note that due

133 to our assumption of uniform dispersal, the parameter b_i can be thought of as a measure of

134 "colonization ability", which combines fecundity and dispersal ability [Bolker and Pacala, 1999,

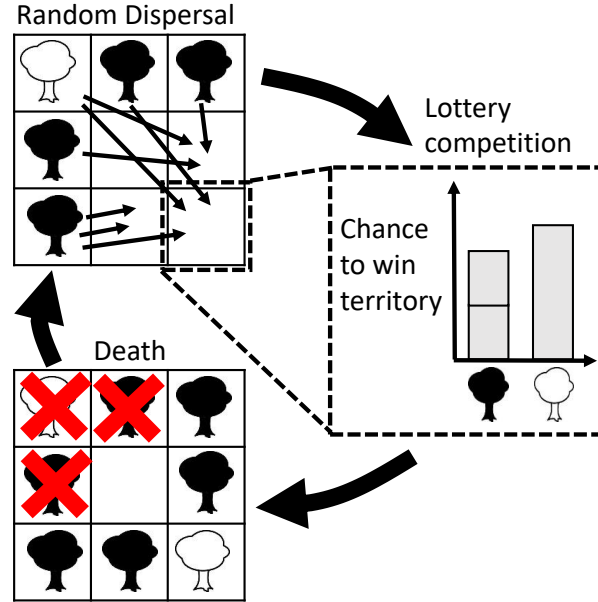135 Levins and Culver, 1971, Tilman, 1994].

Figure 2: Each iteration of our model has three elements. First, propagules are produced by adults and dispersed at random (only propagules landing on unoccupied territories are shown). Some territories may receive zero propagules. Lottery competition then occurs in each unoccupied territory that receives a propagule (only illustrated in one territory). Each type has a probability proportional to $c_i x_i$ of securing a given territory, where c_i measures competitive ability and x_i is the number of propagules that disperse there. In the illustrated territory, the black type disperses more propagules but is a poorer competitor. Territories are then made available by deaths among those adults present at the start of the iteration (red crosses).

The number of individuals of the i'th type landing in any particular territory is denoted x_i. We assume that x_i follows a Poisson distribution $p_i(x_i) = l_i^{x_i} e^{-l_i} / x_i!$, where $l_i = m_i/U$ is the mean territorial propagule density. This approximation becomes exact when the n_i are large enough that drift in n_i can be ignored (Appendix A).

When multiple propagules land on the same territory, the victor is determined by lottery competition: type i wins a territory with probability $c_i x_i / \sum_j c_j x_j$, where c_i is a constant representing relative competitive ability (Fig. 2). We expect that a fraction $p_1(x_1) \ldots p_G(x_G)$ of the U unoccupied territories will have the propagule composition $x_1, \ldots, x_G$. type i is expected to win $c_i x_i / \sum_j c_j x_j$ of these. Ignoring fluctuations about these two expectations (due to our no-drift, large T, large n_i approximation), type i's territorial acquisition is given by

$$\Delta_+ n_i(t) = U(t) \sum_{x_1, \ldots, x_G} \frac{c_i x_i}{\sum_j c_j x_j} p_1(x_1) \ldots p_G(x_G), \tag{2}$$

in our extended lottery model, where the sum only includes territories with at least one propagule present.

Finally, we assume that mortality only occurs in adults (Fig. 2; setting aside the juvenile deaths implicit in territorial contest), and at a constant, type-specific per-capita rate $0 \le d_i \le 1$, so that the overall change in type abundances is

$$\Delta n_i(t) = \Delta_+ n_i(t) - d_i n_i(t). \tag{3}$$

Connection to the Wright-Fisher and classic lottery models

In the classic lottery model [Chesson and Warner, 1981], unoccupied territories are assumed to be saturated with propagules from every type $l_i \gg 1$. From the law of large numbers, the composition of propagules in each territory will then not deviate appreciably from the mean composition $l_1, l_2, \ldots, l_G$ (G is the number of types present), and so the probability that type i wins any particular unoccupied territory is approximately $c_i l_i / \sum_j c_j l_j$. Then the numbers of territories

9

won by each type $\Delta_+ n_1, \Delta_+ n_2, \ldots, \Delta_+ n_G$ follow a multinomial distribution with U trials and success probabilities $\frac{c_1 l_1}{\sum_j c_j l_j}, \frac{c_2 l_2}{\sum_j c_j l_j}, \ldots, \frac{c_G l_G}{\sum_j c_j l_j}$, respectively. Type i is expected to win $c_i l_i / \sum_j c_j l_j$ of the U available territories, and deviations from this expected outcome are small (since T is large by assumption), giving

$$\Delta_+ n_i(t) = \frac{c_i l_i}{\sum_j c_j l_j} U(t) = \frac{c_i l_i}{\bar{c} L} U(t), \tag{4}$$

where $\bar{c} = \sum_j c_j m_j / M$ is the mean propagule competitive ability for a randomly selected propagule, $L = M/U$ is the total propagule density and $M = \sum_j m_j$ is the total number of propagules.

Eq. (4) breaks down for types with low propagule density ($l_i \ll 1$) because territorial acquisition is then not correctly represented by a lottery in each territory with the mean propagule density. Instead, a rare type's propagules only make it to a few territories where at least one of its propagule present. In our extension of the classic lottery model, we correct (Eq. 3) to account for this.

There is a close connection between the classic lottery model and the Wright-Fisher model of genetic drift [Svardal et al., 2015]. In the Wright-Fisher model, type abundances are sampled each generation from a multinomial distribution with success probabilities $w_i n_i / \sum_j w_j n_j$, where w is relative fitness and the n_i are type abundances in the preceding generation. Population size N remains constant. This is equivalent to the classic lottery model with non-overlapping generations ($d_i = 1$ for all i) and relative fitness given by $w_i = b_i c_i$ i.e. a product of fecundity and viability [Crow et al., 1970, pp. 185]. Thus, the classic lottery model is essentially the Wright-Fisher model extended to allow overlapping generations, but ignoring drift. This means that our extension of the classic lottery model to arbitrary densities represents a density-dependent generalization of the Wright-Fisher model.

10

Results

Analytical approximation of the density-dependent lottery

Eq. (2) involves an expectation over the time-dependent dispersal distributions p_i, and is thus too complicated to give intuition about the dynamics of density-dependent lottery competition. We now evaluate this expectation.

Similarly to the high-l_i approximation of classic lottery model, we replace the x_i with appropriate mean values, although we cannot simply replace x_i with l_i. For a type with low propagule density $l_i \ll 1$, we have $x_i = 1$ in the territories where its propagules land, and so its growth comes entirely from territories which deviate appreciably from l_i. To account for this, we separate Eq. (2) into $x_i = 1$ and $x_i > 1$ parts. Our more general approximation only requires that there are no large discrepancies in competitive ability (i.e. we do not have $c_i/c_j \gg 1$ for any two types). We obtain (details in Appendix B)

$$\Delta_+ n_i(t) \approx \left[e^{-L} + (R_i + A_i)\frac{c_i}{\bar{c}} \right] l_i U(t), \tag{5}$$

where

$$R_i = \frac{\bar{c}e^{-l_i}\left(1 - e^{-(L-l_i)}\right)}{c_i + \frac{\bar{c}L - c_i l_i}{L - l_i}\frac{L - 1 + e^{-L}}{1 - (1+L)e^{-L}}},$$

and

$$A_i = \frac{\bar{c}\left(1 - e^{-l_i}\right)}{\frac{1 - e^{-l_i}}{1 - (1+l_i)e^{-l_i}}c_i l_i + \frac{\bar{c}L - c_i l_i}{L - l_i}\left(L\frac{1 - e^{-L}}{1 - (1+L)e^{-L}} - l_i\frac{1 - e^{-l_i}}{1 - (1+l_i)e^{-l_i}}\right)}.$$

Comparing Eq. (5) to Eq. (4), the classic lottery per-propagule success rate $c_i/\bar{c}L$ has been replaced by three separate terms. The first, e^{-L}, accounts for propagules which land alone on unoccupied territories; these territories are won without contest. The second, $R_i c_i/\bar{c}$ represents competitive victories when the i type is a rare invader in a high density population, determining its invasion fitness [Metz et al., 1992]. The third term, $A_i c_i/\bar{c}$, represents competitive victories when the i type is abundant. The relative importance of these three terms varies with both the

11

overall propagule density L and the relative propagule frequencies m_i/M. If $l_i \gg 1$ for all types, we recover the classic lottery model (only the $A_i c_i/\bar{c}$ term remains, and $A_i \to 1/L$). Note that not all unoccupied territories are claimed each iteration, since under Poisson dispersal a fraction e^{-L} remain unoccupied; total population density thus obeys

$$\Delta N = (1 - e^{-L})U - \sum_i d_i n_i \tag{6}$$

Fig. 3 shows that Eq. (5) and its components closely approximate simulations of the density-dependent lottery model over a wide range of propagule densities. Two types are present, one of which is at low frequency. The growth of the low-frequency type relies crucially on the low-density competition term $R_i c_i/\bar{c}$. On the other hand, $R_i c_i/\bar{c}$ is negligible for the high-frequency type, which depends instead on high density territorial victories. Fig. 3 also shows the break-down of the classic lottery model at low propagule densities.

K-selection versus relative fitness

We now compare the density-dependent lottery model from the previous section with MacArthur's analysis of selection in crowded environments [MacArthur and Wilson, 1967]. MacArthur considers a population with two types that have densities n_1 and n_2 subject to density-dependent growth described by

$$\frac{dn_1}{dt} = f_1(n_1, n_2) \qquad \frac{dn_2}{dt} = f_2(n_1, n_2). \tag{7}$$

The environment is assumed to remain constant apart from the type densities. The functions f_1 and f_2 must decline to zero if n_1 or n_2 are sufficiently large, because no population has unlimited resources. This defines the nullclines $f_1(n_1, n_2) = 0$ and $f_2(n_1, n_2) = 0$ in (n_1, n_2) space. The outcome of selection is then determined by the relationship between these nullclines. Specifically, a type will be excluded if its nullcline is completely contained in the region bounded by the other type's nullcline. In other words, for a type to have the possibility of persisting, it must be able to keep growing to higher densities than the other type can tolerate in some region of (n_1, n_2) space
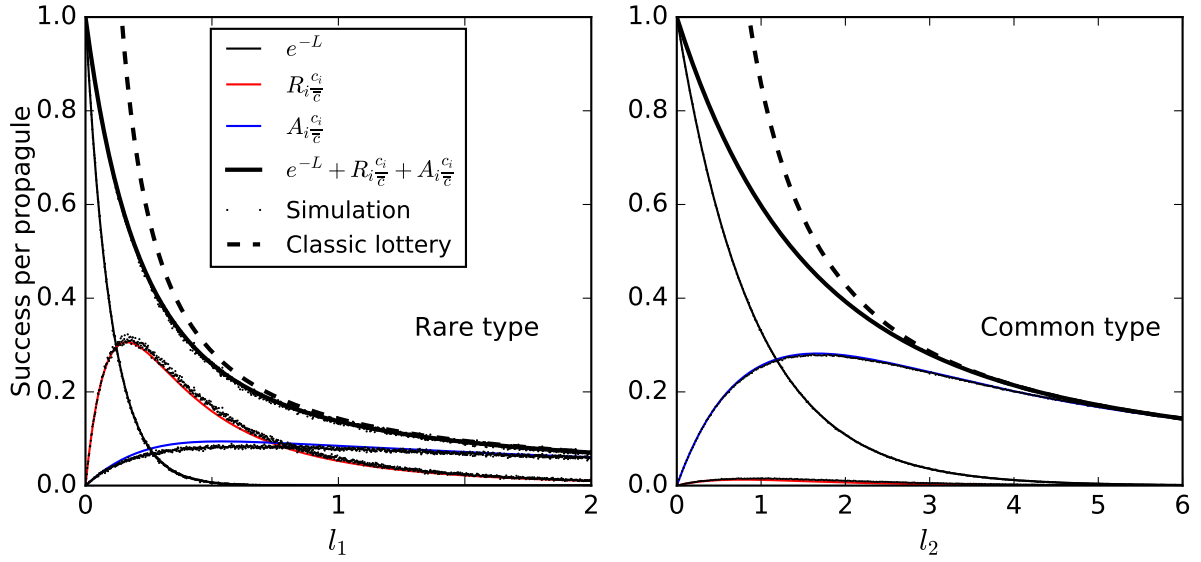
12

Figure 3: Comparison of the analytical approximation Eq. (5) with simulations. Per-propagule success probability $\Delta_+ n_i / l_i U$ from the classic lottery model, individual-based simulations of random dispersal and lottery competition, and Eq. (5) and its three components. Two types are present, a rare type with $c_1 = 1.5$, and a common type with $c_2 = 1$. Simulation points are almost invisible in for the common type due to near exact agreement with Eq. (5). Dashed lines in show the breakdown of the classic lottery model. Parameters: $m_1 = 10^4$ and $m_2 = 9 \times 10^4$ and U varies between 5×10^3 and 10^6.

(Fig. 1a).

To formalize the relationship between nullclines, MacArthur used the symbol "K" to label the four intersection points of the nullclines with the n_1 and n_2 axes, specifically $f_1(K_{11}, 0) = 0$, $f_1(0, K_{12}) = 0$, $f_2(0, K_{22}) = 0$ and $f_2(K_{21}, 0) = 0$. These K values determine whether a region of higher-density growth exists for each type, provided that the nullclines are close to being straight lines. Note that only K_{11} and K_{22} are saturation densities akin to the K parameter in the logistic model; following widespread convention, we will refer to selection on these saturation densities as "K-selection" (Fig. 1a). The other intersection points, K_{12} and K_{21}, are related to competition between types. For instance, in the Lotka-Volterra competition model we have

$$f_1(n_1, n_2) = r_1(1 - \alpha_{11}n_1 - \alpha_{12}n_2)n_1$$

$$f_2(n_1, n_2) = r_2(1 - \alpha_{22}n_1 - \alpha_{21}n_2)n_2 \tag{8}$$

where $\alpha_{11} = 1/K_{11}$ and $\alpha_{22} = 1/K_{22}$ measure competitive effects within types, while $\alpha_{12} = 1/K_{12}$ and $\alpha_{21} = 1/K_{21}$ measure competitive effects between types (Fig. 4a).

Thus, when MacArthur concludes that "fitness is K" in crowded populations [MacArthur and Wilson, 1967, pp. 149], the meaning is that selection either favors the ability to keep growing at ever higher densities (moving a type's own nullcline outwards), or the ability to suppress the growth of competitors at lower densities (moving the nullcline of competitors inwards) [Gill, 1974]. This general idea applies even if the nullclines are nonlinear to such an extent that the "K" values themselves do not give much information about the regions of high-density growth.

It is obvious from Eq. (8) that selection can favor a superior competitor in a crowded population even if its saturation density is the same as, or lower than that of the other types present. However, note that the Lotka-Volterra model still closely couples selection to population density [Smouse, 1976]. Fig. 4a shows Lotka-Volterra selection between two types with the same saturation density ($\alpha_{11} = \alpha_{22}$, $\alpha_{21} > \alpha_{12}$). Even though the initial and final densities of a sweep are the same, density is not constant during a sweep. Only a highly restricted subset of r and α values
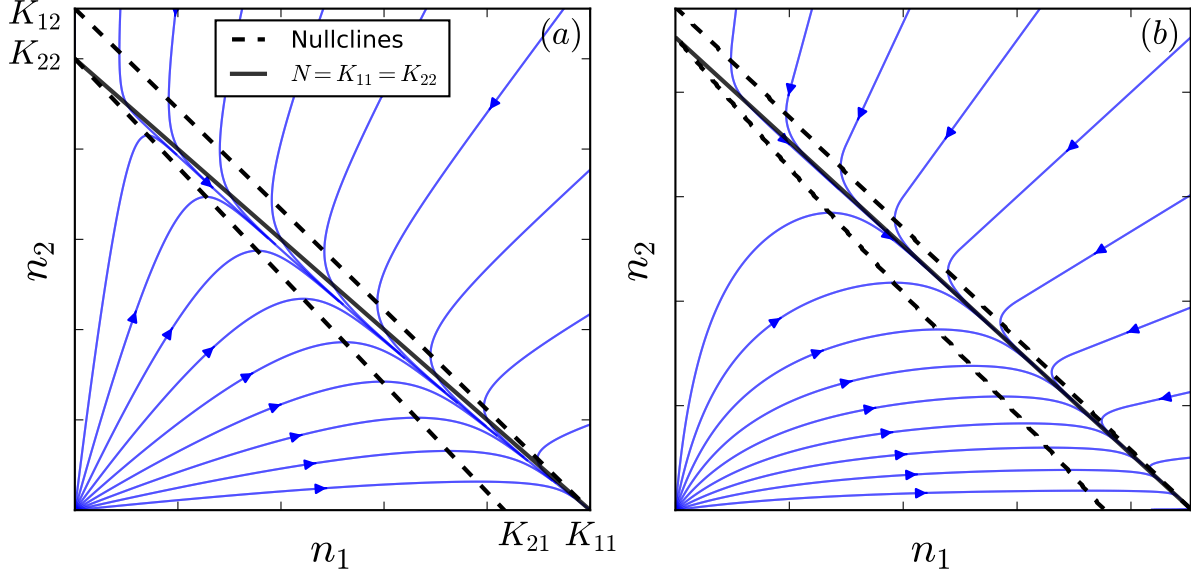
Figure 4: Selection between types with identical saturation density but different inter-type competitive ability. (a) Lotka-Volterra competition (Eq. 8) with $r_1 = r_2 = 1$, $\alpha_{11} = \alpha_{22} = 1$, $\alpha_{12} = 0.9$ and $\alpha_{21} = 1.2$. Trajectories do not follow the line $N = K_{11} = K_{22}$. (b) Lottery competition (Eq. 5) with $b_1 = b_2 = 5$, $d_1 = d_2 = 0.1$ and $c_1/c_2 = 5$. Trajectories converge on the line $N = K_{11} = K_{22}$.

will keep N constant over a selective sweep (further details in Appendix C). Intuitively, for one type to exclude another with the same saturation density, inter-type competitive effects must be stronger than intra-type competitive effects, causing a dip in N over the sweep.

By contrast, if one type in our density-dependent lottery model has a c advantage but birth and death rates are identical (and hence so is saturation density), the density trajectories converge on the line of constant density equal to the saturation density (Fig. 4b). Selection then occurs purely along this line, uncoupled from the regulation of density. In other words, once the population reaches demographic equilibrium, it behaves indistinguishably from a constant-N relative fitness model. More generally, the competitive ability trait c does not directly affect population density (this can be seen formally in Eq. (6)), since c only affects which type wins a territory, not whether a territory is won at all. While quite different from classical growth models like the Lotka-Volterra, this is all perfectly consistent with MacArthur's general argument.

15

Density-dependence and the strength of selection

We are now in a position to analyze the validity of Eq. (1) more explicitly. In the previous section we showed that selection and the regulation of population density can be completely independent of each other even if population growth is density-regulated, and moreover that MacArthur's general argument [MacArthur and Wilson, 1967] for the r/K scheme never precluded this possibility. Nevertheless, selection and density regulation *are* intimately connected in widely used models of population growth. To understand why this poses a problem for Eq. (1), consider the simple birth-death model [Kostitzin, 1939, pp. 20]

$$\frac{dn_i}{dt} = (b_i - \delta_i N)n_i \tag{9}$$

where δ_i is the per-capita increase in type i's mortality rate due to crowding (for simplicity, there are no deaths when uncrowded). Then, starting from a monomorphic population, the frequency of an adaptive δ-variant $\delta_i \to \delta_i(1 - \epsilon)$ obeys

$$\frac{dp_i}{dt} = \epsilon \delta_i N p_i(1 - p_i). \tag{10}$$

The selection coefficient $s = \epsilon \delta_i N$ thus depends on density (compare [Crow et al., 1970, pp. 29]). On the other hand, the frequency of an adaptive b-variant $b_i \to b_i(1 + \epsilon)$ will exactly obey Eq. (1) with $s = \epsilon b_i$, independent of density.

In practice the density dependence in Eq. (10) only matters if N changes substantially during a sweep. This can easily occur if a population is far from demographic equilibrium (we return to this scenario in the Discussion). A more serious objection to Eq. (1) is that it is internally inconsistent even when N has reached equilibrium. To quantify how serious this objection is, we need to account for how much N changes as a result of a sweep involving a density-regulating trait. In Eq. (9), the saturation density is $N_{\text{initial}} = b_i/\delta_i$, which increases to $N_{\text{final}} = b_i/\delta_i(1 - \epsilon) = N_{\text{initial}}/(1 - \epsilon)$ over a δ sweep. Thus, the selection coefficient increases from $s_{\text{initial}} = \epsilon b_i$ to
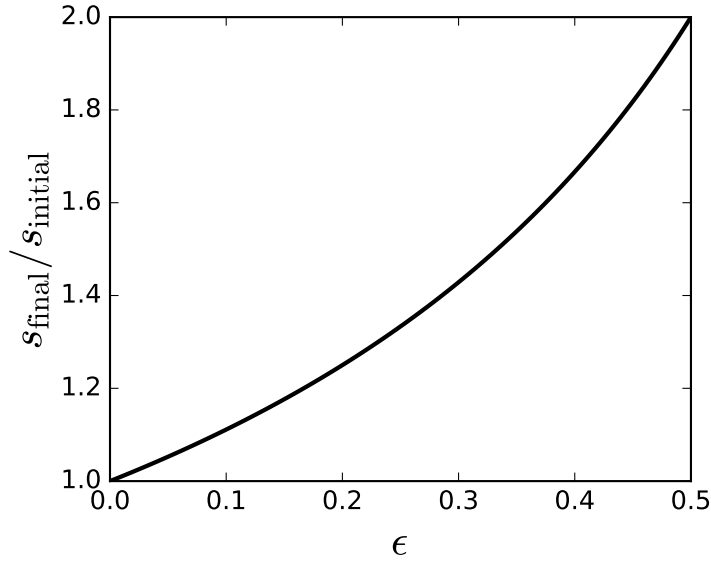
Figure 5: Proportional change in the selection coefficient over a "*K*-selection"-type sweep for a type that experiences proportionally $1 - \epsilon$ fewer deaths induced by crowding. The population is in demographic equilibrium at the start of the sweep.

$s_{\text{final}} = s_{\text{initial}}/(1 - \epsilon)$ over a δ sweep (Fig. 5). Consequently, Eq. (1) breaks down if selection is sufficiently strong on δ_i, with proportional effects of $\epsilon > 0.2$ inducing substantial deviations from constant selection. The cause of this effect is that the mortality rate $\delta_i N$ is both density-regulating and subject to density-dependent selection; thus selection on δ induces a change in N but also changes in strength in response to N.

Let us now contrast the simple linearly density-dependent model Eq. (9) with our density-dependent lottery. As we have seen in our lottery model, the importance of c depends on density, yet since c has no effect on density, s is exactly constant for c sweeps. For b and d sweeps, s is also exactly constant, because the overall regulation of population density applies equally to all types [Prout, 1980]. To see this formally, we set c_i to be the same for all types in Eq. (5) which yields

$$\Delta n_i = \frac{m_i}{M}(1 - e^{-L})U - d_i n_i = \left(b_i \frac{1 - e^{-L}}{L} - d_i \right) n_i. \qquad (11)$$

Thus, the density-regulation of population size occurs via the juvenile recruitment fraction $(1 -$

17

$e^L)/L$. This applies uniformly to all types and therefore does not induce density-dependence in selective advantages related to b or d. Note that our lottery model accounts for two key aspects of the interaction between selection and demography: selection is density-dependent, and density will change as a result of selection on absolute fitness traits. And yet pure b, c and d sweeps all obey Eq. (1) exactly. [Need to flesh this out a little more and address some subtleties with discrete vs. continuous time]

Discussion

It is widely recognized that the relative fitness description of selection widely used by evolutionists is poorly integrated with the ecological literature on population growth [Mallet, 2012]. This is not simply a lack of ecological realism on the part of relative fitness models. In many population growth models, only one life-history stage is represented, and the competitive effects resulting from crowding appear as a reduction in absolute fitness that only depends on the type densities at this life-history stage (e.g. the n_i^2 and $n_i n_j$ terms in the Lotka Volterra equation). However, many species have a "reproductive excess" [thanks for the suggestion, still need to do references] of juveniles that are considerably more fragile than their adult counterparts [Chesson, 1983]. As a result, competition and the selection it induces can concentrated at the juvenile phase, where a large reproductive excess allows for for most juveniles to be competitively eliminated. In our lottery model, this excess appears when the number of propagules is greater than the number of available territories. Then only $\approx 1/L$ of the juveniles contesting available territories survive to adulthood. Note that unlike the role of adult density n_i in single-life-stage models, it is the propagule densities l_i that represent the crowding drives competition. In general, reproductive excesses will tend to produce strictly-relative lottery-type contests in which fitter types can grow at the expense of others by preferentially filling the available adult "slots". The number of slots can remain fixed or change independently of selection at the juvenile stage. By ignoring reproductive excesses, single life-stage models are biased to have total population density be sensitive

18

to ongoing selection, artificially invalidating Eq. (1). In this respect, the Wright-Fisher model and similar "viability selection" models actually capture an important ecological process.

We now turn to the breakdown of Eq. (1). We first discuss the problem shown in Fig. 5, which occurs when strong selection changes (adult) population density and is also density-dependent. In our lottery model, this can only occur if types differ in c as well as b or d. Why does our model need joint trait variation when the classical density-dependent selection literature is built on traits that behave like the δ in Eq. (1) [Christiansen, 2004]? To briefly review: Based on a diploid, bi-allelic variant of the logistic model, the r/K scheme posited a dichotomy between r-selection (uncrowded) and K-selection (crowded), with the latter taken to mean selection for greater saturation density [Gill, 1974]. A more general Lotka-Volterra model introduces the inter-type α_{ij} competition coefficients, with selection on these termed "α-selection" [Gill, 1974, Joshi et al., 2001]. Thus we are left with r-selection (uncrowded) versus K- or α-selection (crowded). The latter two forms of selection in crowded populations behave like selection on δ in that they are both density-dependent and cause density to change over a sweep (although N only dips temporarily during an α-sweep).

By contrast, the density-dependent (c) and density-determining traits (b and d) in our lottery model are independent. This trait separation highlights the two distinct directions in which density and selection interact: selection may depend on density, and density may depend on selection [Prout, 1980]. Both are needed to pose a threat to Eq. (1) (recall that we are still assuming N is not far from equilibrium). In this respect the simple linear models that have become standard in discussions of density-dependent selection [Christiansen, 2004, Mallet, 2012, Roughgarden, 1979] actually represent a complicated form of the density-selection interaction that confounds a key underlying issue.

Note that while this is a conceptual reason to be wary of the classical density-dependent selection models, it is not clear how we should expect the trait variation in nature to align. For instance, should we expect mutations to generally affect b, c and d independently of each other, or pleiotropically such that δ-like selection is prevalent? In the case of well-mixed consumable

19

resources, the R^* rule suggests that δ-like selection will be prevalent. However, for many populations consumable resources are not well-mixed. Spatial localization of consumable resources (e.g. due to restricted movement of nutrients through soils) will tend to create a territorial situation similar to the lottery model, where resource competition only occurs locally and is subsumed into the competitive ability c, which does not affect N.

Relative fitness models truly break down if N is far from equilibrium and selection is density-dependent. For example, wild *Drosophila* experience huge seasonal boom-bust cycles in population density coupled to strong selection that drives large swings in allele frequency [Bergland et al., 2014]. In this case there is no choice but to abandon relative fitness. Whether or not our density-dependent lottery model is a good description of *Drosophila* ecology, the close connection between our model and Wright-Fisher is useful, because drift in our model should behave broadly similarly. Thus, our model it should provide a useful starting point for analyzing evolution in this and other from-from-equilibrium situations.

Acknowledgments

References

NK Arenbaev. Asymptotic behavior of the multinomial distribution. *Theory of Probability & Its Applications*, 21(4):805–810, 1977.

NH Barton, DEG Briggs, JA Eisen, DB Goldstein, and NH Patel. *Evolution*. NY: Cold Spring Harbor Laboratory Press, 2007.

354 TG Benton and A Grant. Evolutionary fitness in ecology: comparing measures of fitness in stochastic, density-dependent environments. *Evolutionary ecology research*, 2(6):769–789, 2000.

356 Alan O. Bergland, Emily L. Behrman, Katherine R. O'Brien, Paul S. Schmidt, and Dmitri A. Petrov. Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in drosophila. *PLOS Genetics*, 10(11):1–19, 11 2014. doi: 10.1371/journal.pgen.1004775.

359 Benjamin M. Bolker and Stephen W. Pacala. Spatial moment equations for plant competition: Understanding spatial strategies and the advantages of short dispersal. *The American Naturalist*, 153(6):575–602, 1999. doi: 10.1086/303199.

362 Peter L. Chesson. *Coexistence of Competitors in a Stochastic Environment: The Storage Effect*, pages 188–198. Springer Berlin Heidelberg, Berlin, Heidelberg, 1983. ISBN 978-3-642-87893-0.

364 Peter L Chesson and Robert R Warner. Environmental variability promotes coexistence in lottery competitive systems. *American Naturalist*, pages 923–943, 1981.

366 F Christiansen. Density dependent selection. In *Evolution of Population Biology: Modern Synthesis*, pages 139–155. Cambridge University Press, 2004.

368 James F Crow, Motoo Kimura, et al. An introduction to population genetics theory. *An introduction to population genetics theory.*, 1970.

370 Odo Diekmann et al. A beginner's guide to adaptive dynamics. *Banach Center Publications*, 63: 47–86, 2004.

372 Michael Doebeli, Yaroslav Ispolatov, and Burt Simon. Towards a mechanistic foundation of evolutionary theory. *eLife*, 6:e23804, feb 2017. ISSN 2050-084X. doi: 10.7554/eLife.23804.

374 Warren J Ewens. *Mathematical Population Genetics 1: Theoretical Introduction*. Springer Science & Business Media, 2004.

376 Douglas E Gill. Intrinsic rate of increase, saturation density, and competitive ability. ii. the evolution of competitive ability. *American Naturalist*, 108:103–116, 1974.

James P Grover. *Resource competition*, volume 19. Springer Science & Business Media, 1997.

Amitabh Joshi, NG Prasad, and Mallikarjun Shakarad. K-selection, α-selection, effectiveness, and tolerance in competition: density-dependent selection revisited. *Journal of Genetics*, 80(2):63–75, 2001.

Vladimir Alexandrovitch Kostitzin. *Mathematical biology*. George G. Harrap And Company Ltd.; London, 1939.

Richard Levins and David Culver. Regional coexistence of species and competition between rare species. *Proceedings of the National Academy of Sciences*, 68(6):1246–1248, 1971.

Robert H MacArthur and Edward O Wilson. *Theory of Island Biogeography*. Princeton University Press, 1967.

James Mallet. The struggle for existence. how the notion of carrying capacity, k, obscures the links between demography, darwinian evolution and speciation. 2012.

Philipp W Messer, Stephen P Ellner, and Nelson G Hairston. Can population genetics adapt to rapid evolution? *Trends in Genetics*, 32(7):408–418, 2016.

C. Jessica E. Metcalf and Samuel Pavard. Why evolutionary biologists should be demographers. *Trends in Ecology and Evolution*, 22(4):205 – 212, 2007. ISSN 0169-5347. doi: https://doi.org/10.1016/j.tree.2006.12.001.

Johan AJ Metz, Roger M Nisbet, and Stefan AH Geritz. How should we define 'fitness' for general ecological scenarios? *Trends in Ecology & Evolution*, 7(6):198–202, 1992.

Timothy Prout. Some relationships between density-independent selection and density-dependent population growth. *Evol. Biol*, 13:1–68, 1980.

Jonathan Roughgarden. Theory of population genetics and evolutionary ecology: an introduction. 1979.

Peter F. Sale. Maintenance of high diversity in coral reef fish communities. *The American Naturalist*, 111(978):337–359, 1977.

Peter E Smouse. The implications of density-dependent population growth for frequency-and density-dependent selection. *The American Naturalist*, 110(975):849–860, 1976.

Hannes Svardal, Claus Rueffler, and Joachim Hermisson. A general condition for adaptive genetic polymorphism in temporally and spatially heterogeneous environments. *Theoretical Population Biology*, 99:76 – 97, 2015. ISSN 0040-5809. doi: http://dx.doi.org/10.1016/j.tpb.2014.11.002.

David Tilman. Competition and biodiversity in spatially structured habitats. *Ecology*, 75(1):2–16, 1994.

Günter P Wagner. The measurement theory of fitness. *Evolution*, 64(5):1358–1376, 2010.

Appendix A: Poisson approximation

For simplicity of presentation, we have assumed a Poisson distribution for the x_i as our model of dispersal. Strictly speaking, the total number of i propagules $\sum x_i$ (summed over unoccupied territories) is then no longer a constant m_i, but fluctuates between generations for a given mean m_i, which is more biologically realistic. Nevertheless, since we do not consider the random fluctuations in type abundances here, and for ease of comparison with the classic lottery model, we ignore the fluctuations in m_i. Instead we focus, on Poisson fluctuations in propagule composition in each territory.

   In the exact model of random dispersal, the counts of a type's propagules across unoccupied territories follows a multinomial distribution with dimension U, total number of trials equal to m_i, and equal probabilities $1/U$ for a propagule to land in a given territory. Thus, the x_i in different territories are not independent random variables. However, for sufficiently large U and m_i, this multinomial distribution for the x_i across territories is closely approximated by a product of independent Poisson distributions for each territory, each with rate parameter l_i

23

[Arenbaev, 1977, Theorem 1]. Since we are ignoring finite population size effects, we effectively have $T \to \infty$, in which case U can be only be small enough to violate the Poisson approximation if there is vanishing population turnover, and then the dispersal distribution is irrelevant anyway. Likewise, in ignoring stochastic finite population size for the n_i, we have effectively already assumed that m_i is large enough to justify the Poisson approximation (the error scales as $1/\sqrt{m_i}$; Arenbaev 1977).

Appendix B: Derivation of growth equation

We separate the right hand side of Eq. (2) into three components $\Delta_+ n_i = \Delta_u n_i + \Delta_r n_i + \Delta_a n_i$ which vary in relative magnitude depending on the propagule densities l_i. Following the notation in the main text, the Poisson distributions for the x_i (or some subset of the x_i) will be denoted p, and we use P as a general shorthand for the probability of particular outcomes.

Growth without competition

The first component, $\Delta_u n_i$, accounts for territories where only one focal propagule is present $x_i = 1$ and $x_j = 0$ for $j \neq i$ (u stands for "uncontested"). The proportion of territories where this occurs is $l_i e^{-L}$, and so

$$\Delta_u n_i = U l_i e^{-L} = m_i e^{-L}. \tag{12}$$

Competition when rare

The second component, $\Delta_r n_i$, accounts for territories where a single focal propagule is present along with at least one non-focal propagule (r stands for "rare") i.e. $x_i = 1$ and $X_i \geq 1$ where $X_i = \sum_{j \neq i} x_j$ is the number of nonfocal propagules. The number of territories where this occurs is $U p_i(1) P(X_i \geq 1) = b_i n_i e^{-l_i}(1 - e^{-(L-l_i)})$. Thus

$$\Delta_r n_i = m_i e^{-l_i}(1 - e^{-(L-l_i)}) \left\langle \frac{c_i}{c_i + \sum_{j \neq i} c_j x_j} \right\rangle_{\tilde{p}}, \tag{13}$$

24

where $\langle\rangle_{\tilde{p}}$ denotes the expectation with respect to $\tilde{p}$, and $\tilde{p}$ is the probability distribution of nonfocal propagule abundances x_j *after* dispersal, in those territories where exactly one focal propagule, and at least one non-focal propagule, landed.

Our "mean field" approximation is to replace x_j with its mean in the last term in Eq. (13),

$$\left\langle \frac{c_i}{c_i + \sum_{j \neq i} c_j x_j} \right\rangle_{\tilde{p}} \approx \frac{c_i}{c_i + \sum_{j \neq i} c_j \langle x_j \rangle_{\tilde{p}}}. \tag{14}$$

Below we justify this replacement by arguing that the standard deviation $\sigma_{\tilde{p}}(\sum_{j \neq i} c_j x_j)$ (with respect to $\tilde{p}$), is much smaller than $\langle \sum_{j \neq i} c_j x_j \rangle_{\tilde{p}}$.

We first calculate $\langle x_j \rangle_{\tilde{p}}$. Let $X = \sum_j x_j$ denote the total number of propagules in a territory and $\mathbf{x}_i = (x_1, \ldots, x_{i-1}, x_{i+1} \ldots, x_G)$ denote the vector of non-focal abundances, so that $p(\mathbf{x}_i) = p_1(x_1) \ldots p_{i-1}(x_{i-1}) p_{i+1}(x_{i+1}) \ldots p_G(x_G)$. Then, $\tilde{p}$ can be written as

$$\begin{aligned}
\tilde{p}(\mathbf{x}_i) &= p(\mathbf{x}_i | X \geq 2, x_i = 1) \\
&= \frac{P(\mathbf{x}_i, X \geq 2 | x_i = 1)}{P(X \geq 2)} \\
&= \frac{1}{1 - (1 + L)e^{-L}} \sum_{X=2}^{\infty} P(X) p(\mathbf{x}_i | X_i = X - 1),
\end{aligned} \tag{15}$$

and so

$$\begin{aligned}
\langle x_j \rangle_{\tilde{p}} &= \sum_{\mathbf{x}_i} \tilde{p}(\mathbf{x}_i) x_j \\
&= \frac{1}{1 - (1 + L)e^{-L}} \sum_{X=2}^{\infty} P(X) \sum_{\mathbf{x}_i} p(\mathbf{x}_i | X_i = X - 1) x_j.
\end{aligned} \tag{16}$$

The inner sum over $\mathbf{x}_i$ is the mean number of propagules of a given nonfocal type j that will be found in a territory which received $X - 1$ nonfocal propagules in total, which is equal to

$\frac{l_j}{L-l_i}(X-1)$. Thus,

$$\langle x_j\rangle_{\tilde{p}} = \frac{l_j}{1-(1+L)e^{-L}}\frac{1}{L-l_i}\sum_{k=2}^{\infty}P(X)(X-1)$$

$$= \frac{l_j}{1-(1+L)e^{-L}}\frac{L-1+e^{-L}}{L-l_i}, \tag{17}$$

where the last line follows from $\sum_{X=2}^{\infty}P(X)(X-1) = \sum_{X=1}^{\infty}P(X)(X-1) = \sum_{X=1}^{\infty}P(X)X - \sum_{X=1}^{\infty}P(X)$.

The exact analysis of the fluctuations in $\sum_{j\neq i}c_jx_j$ is complicated because the x_j are not independent with respect to $\tilde{p}$. These fluctuations are part of the "drift" in type abundances which we leave for future work. Here we use the following approximation to give some insight into the magnitude of these fluctuations and also the nature of the correlations between the x_j. We replace $\tilde{p}$ with $\tilde{q}$, defined as the $\mathbf{x}_i$ Poisson dispersal probabilities conditional on $X_i \geq 1$ (which are independent). The distinction between $\tilde{p}$ with $\tilde{q}$ will be discussed further below. The $\tilde{q}$ approximation gives $\langle x_j\rangle_{\tilde{q}} = \langle x_j\rangle_p/C = l_j/C$,

$$\sigma_{\tilde{q}}^2(x_j) = \langle x_j^2\rangle_{\tilde{q}} - \langle x_j\rangle_{\tilde{q}}^2$$

$$= \frac{1}{C}\langle x_j^2\rangle_p - \frac{l_j^2}{C^2}$$

$$= \frac{1}{C}(l_j^2 + l_j) - \frac{l_j^2}{C^2}$$

$$= \frac{l_j^2}{C}\left(1-\frac{1}{C}\right) + \frac{l_j}{C}, \tag{18}$$

and

$$\sigma_{\tilde{q}}(x_j, x_k) = \langle x_jx_k\rangle_{\tilde{q}} - \langle x_j\rangle_{\tilde{q}}\langle x_k\rangle_{\tilde{q}}$$

$$= \frac{1}{C}\langle x_jx_k\rangle_p - \frac{l_jl_k}{C^2}$$

$$= \frac{l_jl_k}{C}\left(1-\frac{1}{C}\right), \tag{19}$$

26

where $C = 1 - e^{-(L-l_i)}$ and $j \neq k$.

The exact distribution $\tilde{p}$ assumes that exactly one of the propagules present in a given site after dispersal belongs to the focal type, whereas $\tilde{q}$ assumes that there is a focal propagule present before non-focal dispersal commences. As a result, $\tilde{q}$ predicts that the mean propagule density is greater than L (in sites with only one focal propagule is present) when the focal type is rare and the propagule density is high. This is erroneous, because the mean number of propagules in every site is L by definition. Specifically, if $L - l_i \approx L \gg 1$, then the mean propagule density predicted by $\tilde{q}$ is approximately $L + 1$. The discrepancy causes rare invaders to have an intrinsic rarity disadvantage (territorial contests under $\tilde{q}$ are more intense than they should be). In contrast, Eq. (17) correctly predicts that there are on average $\sum_{j \neq i} \langle x_j \rangle_{\tilde{p}} \approx L - 1$ nonfocal propagules because $\tilde{p}$ accounts for potentially large negative covariances between the x_j "after dispersal". By neglecting the latter covariences, $\tilde{q}$ overestimates the fluctuations in $\sum_{j \neq i} c_j x_j$; thus $\tilde{q}$ gives an upper bound on the fluctuations. The discrepancy between $\tilde{q}$ and $\tilde{p}$ will be largest when L is of order 1 or smaller, because then the propagule assumed to already be present under $\tilde{q}$ is comparable to, or greater than, the entire propgaule density.

Decomposing the variance in $\sum_{j \neq i} c_j x_j$,

$$\sigma_{\tilde{q}}^2 \left(\sum_{j \neq i} c_j x_j \right) = \sum_{j \neq i} \left[c_j^2 \sigma_{\tilde{q}}^2(x_j) + 2 \sum_{k > j, k \neq i} c_j c_k \sigma_{\tilde{q}}(x_j, x_k) \right], \tag{20}$$

and using the fact that $\sigma_{\tilde{q}}(x_j, x_k)$ and the first term in Eq. (18) are negative because $C < 1$, we obtain an upper bound on the relative fluctuations in $\sum_{j \neq i} c_j x_j$,

$$\frac{\sigma \left(\sum_{j \neq i} c_j x_j \right)}{\langle \sum_{j \neq i} c_j x_j \rangle} = C^{1/2} \frac{\left(\sum_{j \neq i} c_j^2 l_j + (1 - 1/C) \left(\sum_{j \neq i} c_j l_j \right)^2 \right)^{1/2}}{\sum_{j \neq i} c_j l_j} < C^{1/2} \frac{\left(\sum_{j \neq i} c_j^2 l_j \right)^{1/2}}{\sum_{j \neq i} c_j l_j}. \tag{21}$$

Suppose that the c_j are all of similar magnitude (their ratios are of order one). Then Eq. (21) is $\ll 1$ for the case when $L - l_i \ll 1$ (due to the factor of $C^{1/2}$), and also for the case when at least some of the nonfocal propagule densities are large $l_j \gg 1$ (since it is then of order $1/\sqrt{L - l_i}$).

474 The worst case scenario occurs when $L - l_i$ is of order one. Then Eq. (21) gives a relative error of

475 approximately 50%, which from our earlier discussion we know to be a substantial overestimate

476 when L is of order 1. Our numerical results (Fig. 3) confirm that the relative errors are indeed

477 small.

478   However, the relative fluctuations in $\sum_{j \neq i} c_j x_j$ can be large if some of the c_j are much larger

479 than the others. Specifically, in the presence of a rare, extremely strong competitor ($c_j l_j \gg c_{j'} l_{j'}$

480 for all other nonfocal types j', and $l_j \ll 1$), then the RHS of Eq. (21) can be large and we cannot

481 make the replacement Eq. (14).

482   Substituting Eqs. (14) and (17) into Eq. (13), we obtain

$$\Delta_r n_i \approx m_i R_i \frac{c_i}{\bar{c}}, \tag{22}$$

483 where R_i is defined in Eq. (6).

Competition when abundant

485 The final contribution, $\Delta_a n_i$, accounts for territories where two or more focal propagules are

486 present (a stands for "abundant"). Similarly to Eq. (13), we have

$$\Delta_a n_i = U(1 - (1 + l_i)e^{l_i}) \left\langle \frac{c_i x_i}{\sum_j c_j x_j} \right\rangle_{\hat{p}} \tag{23}$$

487 where $\hat{p}$ is the probability distribution of both focal and nonfocal propagule abundances *after*

488 dispersal in those territories where at least two focal propagules landed.

489   Again, we argue that the relative fluctuations in $\sum c_j x_j$ are much smaller than 1 (with respect

490 to $\hat{p}$), so that,

$$\left\langle \frac{c_i x_i}{\sum_j c_j x_j} \right\rangle_{\hat{p}} \approx \frac{c_i \langle x_i \rangle_{\hat{p}}}{\sum_j c_j \langle x_j \rangle_{\hat{p}}}. \tag{24}$$

Following a similar procedure as for $\Delta_r n_i$, where the vector of propagule abundances is denoted

x, the mean focal type abundance is,

$$
\begin{aligned}
\langle x_i \rangle_{\hat{p}} &= \sum_{\mathbf{x}} x_i p(\mathbf{x}|x_i \geq 2) \\
&= \sum_{x_i} x_i p(x_i|x_i \geq 2) \\
&= \frac{1}{1 - (1 + l_i)e^{-l_i}} \sum_{x_i \geq 2} p(x_i) x_i \\
&= l_i \frac{1 - e^{-l_i}}{1 - (1 + l_i)e^{-l_i}}.
\end{aligned}
\tag{25}
$$

For nonfocal types $j \neq i$, we have

$$
\begin{aligned}
\langle x_j \rangle_{\hat{p}} &= \sum_{\mathbf{x}} x_j p(\mathbf{x}|x_i \geq 2) \\
&= \sum_{X} P(X|x_i \geq 2) \sum_{\mathbf{x}} x_j p(\mathbf{x}|x_i \geq 2, X) \\
&= \sum_{X} P(X|x_i \geq 2) \sum_{x_i} p(x_i|x_i \geq 2, X) \sum_{x_i} x_j p(\mathbf{x}_i|X_i = X - x_i) \\
&= \sum_{X} P(X|x_i \geq 2) \sum_{x_i} p(x_i|x_i \geq 2, X) \frac{l_j (X - x_i)}{L - l_i} \\
&= \frac{l_j}{L - l_i} \left[\sum_{X} P(X|x_i \geq 2)X - \sum_{x_i} p(x_i|x_i \geq 2)x_i \right] \\
&= \frac{l_j}{L - l_i} \left(L \frac{1 - e^{-L}}{1 - (1 + L)e^{-L}} - l_i \frac{1 - e^{-l_i}}{1 - (1 + l_i)e^{-l_i}} \right).
\end{aligned}
\tag{26}
$$

To calculate the relative fluctuations in $\sum_{j \neq i} c_j x_j$, we use a similar approximation as for $\Delta_r n_i$: $\hat{p}$ is approximated by $\hat{q}$, defined as the **x** dispersal probabilities in a territory conditional on $x_i > 2$ (that is, treating the x_j as indepenent). All covariances between nonfocal types are now zero, so that $\sigma_{\hat{q}}^2(\sum c_j x_j) = \sum c_j^2 \sigma_{\hat{q}}^2(x_j)$, where $\sigma_{\hat{q}}^2(x_j) = l_j$ for $j \neq i$, and

$$
\sigma_{\hat{q}}^2(x_i) = \frac{l_i}{D} \left(l_i + 1 - e^{-l_i} - \frac{l_i}{D} \left(1 - e^{-l_i} \right)^2 \right),
\tag{27}
$$

29

where $D = 1 - (1 + l_i)e^{-l_i}$, and

$$\frac{\sigma_{\hat{q}}(\sum c_j x_j)}{\langle \sum c_j x_j \rangle} = \frac{\left(\sum_{j \neq i} c_j^2 l_j + c_i^2 \sigma_{\hat{q}}^2(x_i) \right)^{1/2}}{\sum_{j \neq i} c_j l_j + c_i l_i (1 - e^{-l_i})/D}. \tag{28}$$

Similarly to Eq. (21), the RHS of Eq. (28) is $\ll 1$ for the case that $L \ll 1$ (due to a factor of $D^{1/2}$), and also for the case when at least some of the propagule densities (focal or nonfocal) are large — provided that c_i and the c_j are all of similar magnitude. Again, the worst case scenario occurs when l_i and $L - l_i$ are of order 1, in which case Eq. (28) is around 35%, which is again where the $\hat{q}$ approximation produces the biggest overestimate of the fluctuations in **x**. Similarly to Eq. (21), the RHS of (28) will not be $\ll 1$ in the presence of a rare, extremely strong competitor.

Combining Eqs. (23) and (24), we obtain

$$\Delta_a n_i = m_i A_i \frac{c_i}{\bar{c}}, \tag{29}$$

where A_i is defined in Eq. (6).

Appendix C: Total density in the Lotka-Volterra competition model

Here we show that under the Lotka-Volterra model of competition, total density N does not in general remain constant over a selective sweep in a crowded population even if the types have the same saturation density (for a related discussion on the density- and frequency-dependence of selection in the Lotka-Volterra mode, see [Mallet, 2012, Smouse, 1976]).

We assume equal effects of crowding within types $\alpha_{11} = \alpha_{22} = \alpha_{\text{intra}}$ and $N = 1/\alpha_{\text{intra}}$ and check whether it is then possible for $\frac{dN}{dt}$ to be zero in the sweep ($n_1, n_2 \neq 0$). Substituting these

conditions into Eq. (8), we obtain

$$\frac{dn_1}{dt} = r_1(\alpha_{11} - \alpha_{12})n_1 n_2$$
$$\frac{dn_2}{dt} = r_2(\alpha_{22} - \alpha_{21})n_1 n_2 \tag{30}$$

509 Adding these together, $\frac{dN}{dt}$ can only be zero if

$$r_1(\alpha_{\text{intra}} - \alpha_{12}) + r_2(\alpha_{\text{intra}} - \alpha_{21}) = 0. \tag{31}$$

510 To get some intuition for Eq. (31), suppose that a mutant arises with improved competitive

511 ability but identical intrinsic growth rate and saturation density ($r_1 = r_2$ and $\alpha_{11} = \alpha_{22}$). This

512 could represent a mutation to an interference competition trait, for example [Gill, 1974]. Then,

513 according the above condition, for N to remain constant over the sweep, the mutant must find

514 the wildtype more tolerable than itself by exactly the same amount that the wildtype finds the

515 mutant less tolerable than itself.

516   Even if we persuaded ourselves that this balance of inter-type interactions is plausible in some

517 circumstances, when multiple types are present the requirement for constant N becomes

$$\sum_{ij} r_i(\alpha_{\text{intra}} - \alpha_{ij})p_i p_j = 0, \tag{32}$$

518 which depends on frequency and thus cannot be satisfied in general for constant inter-type

519 coefficients α_{ij}. We conclude that selection in the Lotka-Volterra competition model will generally

520 involve non-constant N.