bioRxiv preprint doi: https://doi.org/10.1101/2021.09.17.460820; this version posted September 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The response of a metapopulation to a changing environment

Nick Barton[1,§], Oluwafunmilola Olusanya[1]

[1]*Institute of Science and Technology Austria, Am Campus 1, Klosterneuburg 3400, Austria.*
§ Corresponding author

Abstract

A species that is distributed across heterogeneous environments may adapt to local conditions. Szep et al (Evolution, 2021) modelled this process in the infinite island model, finding the stationary distribution of allele frequencies and deme sizes. We extend this to ask how a metapopulation responds to changes in carrying capacity, selection strength, or migration rate, restricting attention to fixed deme size ("soft selection"). We develop a "fixed-state" approximation (accurate when migration is rare) which assumes that the loci are near fixation. Under this approximation, polymorphism is only possible for a narrow range of habitat proportions when selection is weak compared to drift, but for a much wider range otherwise. When local conditions (Ns or Nm) change in a single deme of the metapopulation, it takes the population a time of order $1/m$ to reach the new equilibrium. However, even with many loci, there can be substantial fluctuations in net adaptation, due to the bimodal allele frequency distributions at each locus. Thus, in a finite metapopulation, variation may gradually be lost by chance, even if it would persist if there were infinitely many demes. When conditions change across the whole metapopulation, there can be rapid change, which is predicted well by the fixed-state approximation when $Nm \ll 1$.

Keywords: metapopulation, local adaptation, species' range, diffusion, adaptive walk, changing conditions, soft selection.

Introduction

Species must adapt to varied environments, whilst drawing on a common pool of genetic variation. Thus, there is a tension between selection that favours different alleles in different places, and the maintenance of diversity across the whole species. Local populations can only sustain themselves if they are sufficiently well-adapted; conversely, adaptation to conditions beyond the current niche can extend the range of the species.

These issues, which lie at the interface between ecology and evolution, have only quite recently attracted sustained theoretical attention. This ranges from studies of "evolutionary rescue", typically of a single isolated deme [5, 6, 18], through to analyses of limits to a species' range in one or two spatial dimensions [9, 11, 14]. Here, we consider an idealised metapopulation; in this island model, there is no explicit spatial structure. Nevertheless, we can ask whether the species' range can extend over a variety of habitats, and examine how it responds dynamically to changing conditions – either in a single deme, or across the whole metapopulation.

This paper is an extension of [16], which analysed the joint evolution of allele frequencies and deme sizes, in an island model with explicit density-dependent regulation; a diffusion approximation gave explicit formulae for the stationary distribution of an infinite metapopulation. Here, we extend this treatment to consider the evolution of individual demes, and of the whole metapopulation, as conditions change; we also consider fluctuations in a metapopulation with

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.17.460820; this version posted September 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

a limited number of demes, where variation can be lost by chance. We simplify the problem by assuming that deme sizes are fixed, independent of adaptedness ("soft selection"), but believe that the methods we introduce can be extended to allow density regulation ("hard selection").

In principle, we can calculate the joint distribution of deme size and allele frequencies under the diffusion approximation. However, this is numerically challenging, since it involves a high dimensional partial differential equation; in any case, it can only be done for an infinite metapopulation, where the mean population size and allele frequencies across the population as a whole are fixed, even though population sizes and allele frequencies within any deme follow a distribution. In order to go beyond mere simulation, we use the approximation that loci are typically near fixation; this is accurate if the number of non-native alleles that enter per generation is small. It allows us to follow the distribution of states of a finite metapopulation through time, which depends only on the rates of substitutions in either direction. This "fixed-state" approximation is an extension of models of "adaptive walks" (e.g. Orr [12], Trubenova et al [17]) to structured populations.

We first consider an infinite metapopulation, and determine the accuracy of the fixed-state approximation. We then apply the approximation to calculate the dynamics of a finite metapopulation, and to find how its equilibria depend on the number of demes. (In order for a non-trivial equilibrium to exist, we must allow a low rate of mutation to maintain variation in the long term). Finally, we show how metapopulations respond to changing conditions, focusing on changes that take the system between qualitatively different regimes.

Model and Methods

We simulate a haploid population, assuming linkage equilibrium. Provided that selection is weak, this is accurate, and allows us to efficiently simulate large numbers of loci and demes; Szep et al (2021, SI C) examine the effects of linkage disequilibrium in this model, using individual-based simulations. We obtain analytical results by taking the diffusion limit, which also assumes weak selection, and then approximate this by assuming that demes are near fixation, which applies when there are few migrants ($Nm<1$). As is traditional in population genetics, we take the fundamental model to be the diffusion, since this captures the behaviour of a variety of particular life histories, and identifies the key dimensionless parameters.

Simulations

Our baseline island model assumes that demes each have carrying capacity N haploid individuals, and contribute equally to the migrant pool. A deme of size N is expected to lose a fraction m of individuals by emigration, and receives a Poisson distributed number of migrants, Nm^*, with expectation Nm. There are L biallelic loci, with the two alternative alleles labelled $X_{i,k} = 0$ or 1; i labels the deme, and k the locus. Deme i is described by $\{j_{i,1}, j_{i,2}, \ldots, j_{i,L}\}$, where $0 \leq j_{i,k} \leq N$ is the number of copies of the '1' allele at the k'th locus. That allele is favoured by selection s_i, which we assume to be the same across loci; the marginal relative fitnesses are $1{:}e^{s_i}$, and fitnesses multiply across loci. Under soft selection, loci evolve independently, and so it would be straightforward to extend to allow variation in selection across loci.

We assume linkage equilibrium (LE), and apply the Wright-Fisher model to each locus independently. After selection, allele frequencies are $p^*_{i,k} = j_{i,k} / \left((N - j_{i,k})\, e^{-s_i} + j_{i,k}\right)$, and after migration, $p^{**}_{i,k} = m\,\bar{p}_k + (1-m) p^*_{i,k}$ where $\bar{p}_k$ is the frequency averaged across all demes of the metapopulation. The new population in deme i consists of N individuals, the number of allele copies at locus k being binomially sampled with frequency $p^{**}_{i,k}$. This procedure is accurate provided that

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.17.460820; this version posted September 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

s is not too large (<0.2, say), so that recombination shuffles genes faster than selection, drift, or migration build up associations between them (Szep et al ., 2021, SI C).

A Mathematica notebook containing the simulation code and result is provided as a supplementary material.

Diffusion approximation

The diffusion approximation to this model describes the evolution of the joint distribution of allele frequencies across different demes, conditional on the mean allele frequency across the metapopulation [2]. A single deme follows a stochastic path governed by this distribution, whilst an infinite metapopulation represents the whole distribution, which evolves deterministically at the level of the whole metapopulation. The diffusion depends only on scaled parameters Ns, Nm.

Wright [19, 20] gave an explicit solution for the stationary distribution of allele frequencies:

$$\Psi\left[p\,|\overline{p}\right]=\frac{1}{Z}\prod_{k=1}^{L}p_k^{2\,N\,m\bar{p}_k-1}q_k^{2N\,m\bar{q}_k-1}e^{2Nsp_k} \tag{1}$$

where Z is a normalising constant. Under this simple model of directional selection, allele frequencies evolve independently across demes and across loci, conditional on the mean allele frequencies, $\bar{p}_k$. Equation (1) applies to a single deme; the subscript i was dropped for clarity. All demes that share the same parameters will follow the same distribution, in a given habitat, and so we can integrate over the distribution, and sum over habitats, to find the mean $\bar{p}_k$. This allows us to solve fully for the stationary state.

Fixed-state approximation

If the number of incoming alleles is small ($Nm\ll 1$) then the distribution of allele frequencies will be sharply peaked around 0 and 1. To a good approximation, populations are near fixation for one or other allele, and their state is determined by the rates of substitution in either direction. Since we will later be considering the stationary state of a finite metapopulation, we must include mutation, which we assume to be symmetric at rate μ. Then, the rate at which demes currently fixed for allele 0 substitute allele 1, $\lambda_{0\to 1}$ (or vice versa, $\lambda_{1\to 0}$) is the product of the number of '1' (or '0') alleles entering the population, and their individual fixation probability. Thus:

$$\lambda_{0\to 1}=\frac{2s\,(N\mu+Nm\bar{p})}{1-e^{-2N\,s}},\qquad \lambda_{1\to 0}=\frac{2s\,(N\mu+Nm\bar{q})}{e^{2N\,s}-1} \tag{2}$$

Different loci evolve independently, conditional on the numbers of migrants coming into the deme ($Nm\bar{p}$), ($Nm\bar{q}$).

For an infinite metapopulation, and two habitats with selection s_1, s_2, with deme sizes fixed at N (i.e., soft selection), we can just follow the proportion of demes fixed for the '1' allele in each habitat. Neglecting mutation:

$$\begin{aligned}\partial_t P_1&=\frac{2s_1\,Nm}{1-e^{-2N\,s_1}}\left(\bar{p}Q_1-\bar{q}e^{-2Ns_1}P_1\right)\\ \partial_t P_2&=\frac{2s_2\,Nm}{1-e^{-2Ns_2}}\left(\bar{p}Q_2-\bar{q}e^{-2Ns_2}P_2\right)\\ \bar{p}&=\rho\,P_1+(1-\rho)P_2\end{aligned} \tag{3}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.17.460820; this version posted September 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The first two equations involve the difference in net rates of substitution in each direction. The fraction of loci near fixation for 0, 1 are Q, P; the fraction of migrants with allele 1 vs 0, which can contribute to a substitution, are $\bar{p}, \bar{q}$; and the fixation probabilities in each direction are in the ratio $1{:}e^{-2Ns_1}$. Finally, the mean allele frequencies are a weighted average across habitats, which are in the proportions $\rho{:}1-\rho$.

These equations apply separately to each locus, but for simplicity, in numerical examples we will assume symmetric initial conditions, so that P_1, P_2 are the same for all loci, and correspond to the proportion of loci fixed for the '1' allele in each deme.

If the '1' allele is favoured in habitat 1, but disfavoured in habitat 2 (i.e. $s_2{<}0{<}s_1$), and if neither habitat is too rare, then polymorphism is possible, with equilibrium frequency given by:

$$\bar{p}{=}\frac{\rho\left(e^{2N(s_1-s_2)}-1\right)-\left(e^{-2Ns_2}-1\right)}{\left(e^{2Ns_1}-1\right)\left(e^{-2Ns_2}-1\right)}\ ,\qquad \frac{e^{-2Ns_2}-1}{e^{2N(s_1-s_2)}-1}{<}\rho{<}\frac{\left(e^{-2Ns_2}-1\right)}{e^{2N(s_1-s_2)}-1}e^{2Ns_1} \tag{4}$$

If selection is weak relative to drift, polymorphism is possible only for a very narrow range of habitat proportions (left of fig. 1), whereas if it is strong, polymorphism is possible over a wide range (right of fig. 1).

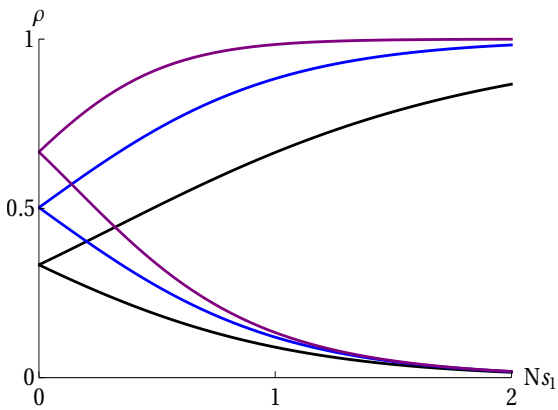

Figure 1: Bounds on the proportions of habitat 1, ρ, between which polymorphism is possible, as a function of the strength of selection in that habitat, Ns_1. The three sets of bounds correspond to $Ns_2/Ns_1 = 0.5, 1, 2$ (black, blue and purple respectively). These results apply in the limit of low migration, and soft selection.

Suppose now that there are a finite number of demes, with d_i having habitat i. At any one locus, the state of the metapopulation is described by the number of demes fixed for the '1' allele, $0{\leq}k_i{\leq}d_i$. For example, with two habitats, there are $(d_1{+}1)\,(d_2{+}1)$ possible values for the state $\{k_1, k_2\}$. The probability of transitions between these states depends on the mean allele frequency across the metapopulation. With soft selection, where all demes have the same size N, this mean is just $\bar{p}{=}(k_1{+}k_2)\,/\,(d_1{+}d_2)$. We can therefore calculate the transition matrix that governs the stochastic evolution of the metapopulation; the stationary state is given by the leading eigenvector of this matrix. With soft selection, each locus evolves independently, governed by this matrix, and so we can easily calculate the stochastic evolution of the metapopulation.

In the Appendix, we examine the accuracy of the fixed-state approximation under soft selection. This approximation applies in the limit of low migration, and identifies the failure of adaptation due to random drift.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.17.460820; this version posted September 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Results

Evolution of a single deme

Suppose in a metapopulation containing two habitats we focus on the rare habitat so that the numbers and adaptedness of the whole metapopulation remains constant. Suppose also that the adaptive allele is favored in the rare habitat and disfavored in the common habitat with selective strength s_1 and s_2 respectively, we can look at how the mean allele frequency (averaged across loci) in the rare habitat responds to changes in local conditions such as Nm or local conditions such as changes in the population size or local selection pressure in a single deme (Ns) as shown below. We see in fig. 2a that with Nm too low, genetic variation is lost and with Nm too high, the adaptive allele is swamped. There is an intermediate value of Nm that maximizes the mean allele frequency of the adaptive allele in the rare habitat. In fig. 2b, with Ns_1 too small, drift overwhelms selection and the adaptive allele is lost from the rare habitat.

Consider a metapopulation, where Nm is small enough that populations are near fixation. If Ns_1=1 in a rare habitat, represented in ρ=0.2 of the demes, and Ns_2=−2 in the common habitat, then polymorphism will be maintained with $\bar{p}$=0.079 overall (eq. (4)). We begin by considering how a single deme responds to changes in its local conditions, for fixed $\bar{p}$, and so in fact, all that matters is the value of $\bar{p}$. In the focal deme, allele frequencies will be in the ratio $\bar{q}$:$\bar{p}e^{2Ns_1}$ when $Nm \ll 1$, since that is the ratio of substitution rates in either direction; hence, the expected allele frequency in the rare habitat is 0.386 (fig. 2a, left). As Nm increases, the expected allele frequency decreases, approaching $\bar{p}$=0.079 (fig. 2a, right). For given Nm, the expected allele frequency in the focal deme increases with Ns_1 from $\bar{p}$ to 1, as selection becomes more effective (fig. 2b).

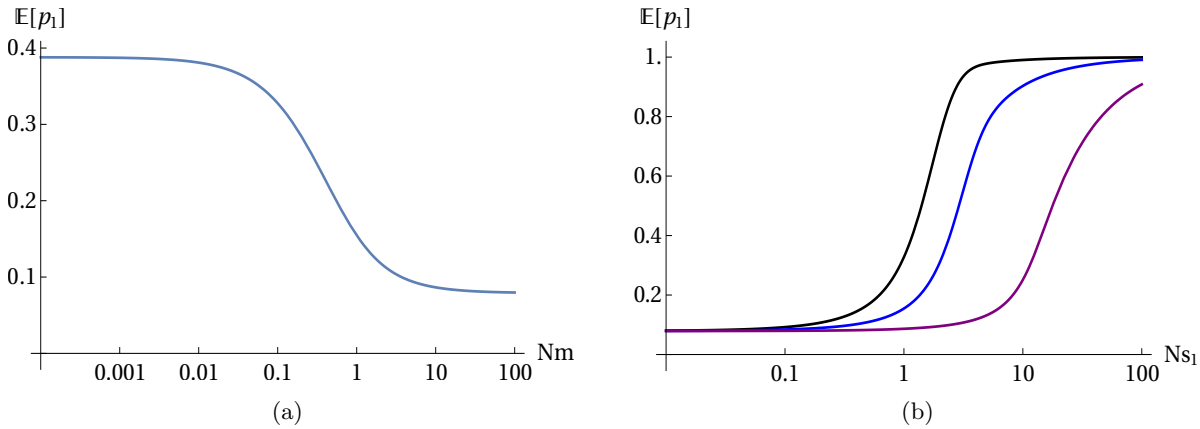

Figure 2: Left: Expected allele frequency vs Nm with $\bar{p}$=0.079, Ns_1=1. Right: Expected allele frequency vs. Ns, for Nm=0.1, 1, 10 (black, blue, purple).

Figure 3a shows how the distribution of allele frequencies changes as Nm changes. If all loci start close to the frequency in the gene pool ($\bar{p}$=0.079) then with a low migration rate (Nm=0.05), even weak selection (Ns=1) can raise the mean substantially, to 0.355. However, this increase is slow, taking ∼5000 generations, because it occurs through occasional substitutions, at a rate is proportional to m=5×10^{-4} (eq. (3)). The population does mostly flip between fixation of one or other allele, giving a U-shaped frequency distribution (e.g. grey trajectory in fig. 3), and so the fixed-state approximation is quite close to the exact mean (orange vs. red at left). However, the average across even 100 loci fluctuates substantially (blue), implying that population fitness will fluctuate randomly, even when adaptation is highly polygenic.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.17.460820; this version posted September 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

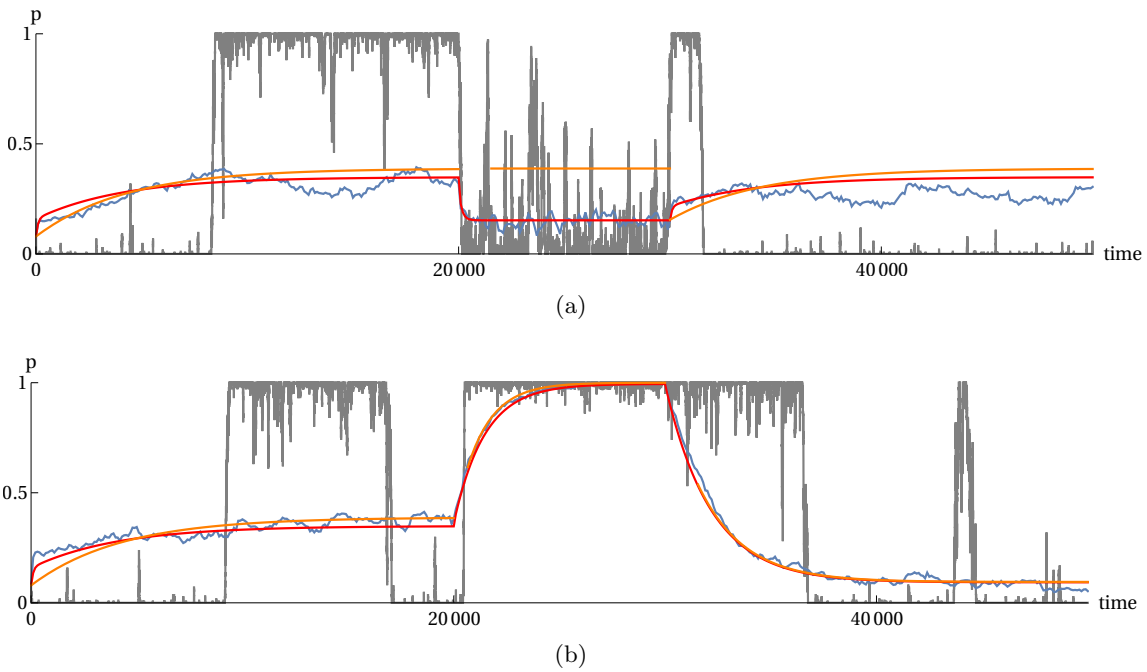

Figure 3: (a) Evolution of a single deme as Nm changes; Ns=1, $\bar{p}$=0.079, L=100 loci. Initially, Nm=0.05, and all loci are at $\bar{p}$. After 20,000 generations, Nm increases to 1, and after another 10,000 generations, it returns to Nm=0.05. The grey line shows allele frequency at a single locus, and the blue line shows the average over 100 loci. The red curve is the mean of the probability distribution, calculated exactly using the Wright-Fisher transition matrix. The orange curve is the fixed-state approximation (eq. (3)), which is accurate only for $Nm \ll 1$. (b) The same, but for Ns changing from 1 to 10 at 20,000 generations, and then to 0.1 at 30,000 generations; Nm=0.05 throughout.

At 20,000 generations, the number of migrants increases to Nm=1, and the mean allele frequency is quickly pulled down towards that in the gene pool, to 0.155. The fixed-state approximation is the limit of low migration, and so is independent of Nm (see fig. A1). Indeed, allele frequencies are now often intermediate, and so this approximation fails (orange vs. red, fig. 3, middle). Nevertheless, it does give the important intuition that rates of change are proportional to migration, which is now m=0.01, implying a ∼100 generation timescale for response of the population mean. In this model, variance is maintained by migration, and so the response to selection is proportional to m. After Nm returns to the original low value at 30,000 generations, there is a slow return to the original bimodal distribution, again captured by the fixed-state approximation (orange vs red at right of fig. 3).

Figure 3b shows the response to changes in Ns, which could arise through changes in selection strength, and/or changes in effective local population size. In this example, Nm=0.05 throughout, and so the fixed-state approximation is accurate (orange vs red curves). The timescale is again set by m, which determines the rate at which variation is introduced into local demes. Since $m=5\times10^{-4}$, it takes thousands of generations for the proportion of loci fixed for the '1' allele to respond to changes in selection strength.

Figure 4 shows the time taken for a population to respond to changes in Nm (fig. 4a) or Ns (fig. 4b), as a function of the other parameter. As we saw in fig. 3a, an increase in Nm causes a much faster response than a decrease, simply because high gene flow introduces more genetic variance. However, if selection is very strong, the response time becomes similar in either direction, and decreases in proportion to Ns (right of fig. 4a). The response to changes in Ns take somewhat longer for an increase than a decrease (fig. 4b), but the main pattern here is that the response time decreases in proportion to Nm.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.17.460820; this version posted September 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

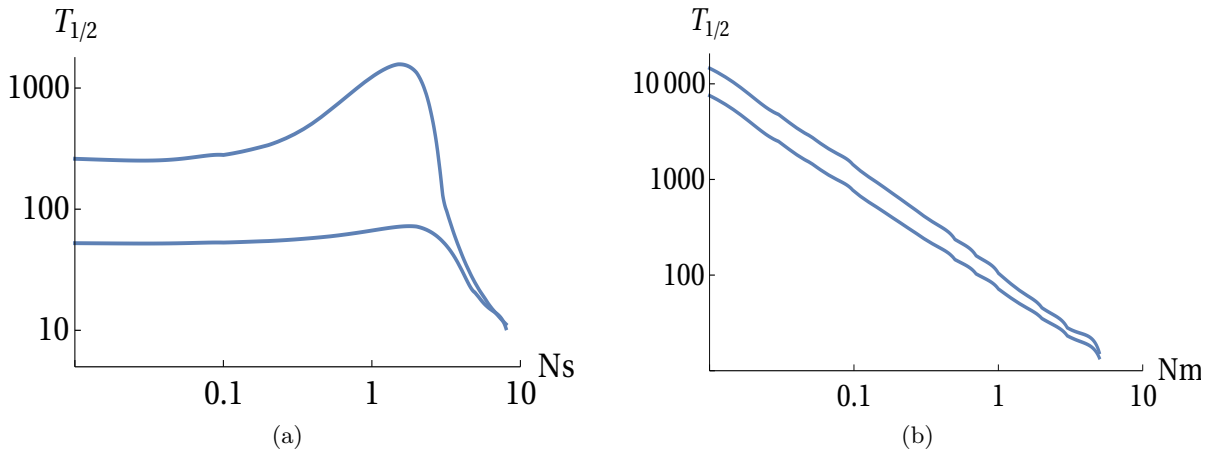

Figure 4: The time to make half of the response to a change in parameters. For both plots, $\bar{p}$=0.079. Values were calculated using a transition matrix with N=100. (a) Nm shifts from 0.05 to 1 or from 1 to 0.05 (lower, upper curves, resp.), for varying Ns. (b) Ns shifts from 0.1 to 1 or from 1 to 0.1 (upper, lower curves, resp.) for varying Nm.

Evolution of a metapopulation

We begin by considering the stationary state of a metapopulation, extending Szep et al [16] by allowing a finite number of demes – in which case, a low rate of mutation is required to maintain variation in the long term. We then give an example that shows how variation is lost, as loci fixed across the whole metapopulation. Finally, we give examples (analogous to fig. 3), showing the response when parameters change across the whole metapopulation.

Stationary state of a finite metapopulation in the limit of small Nm

Szep et al (2021, Fig. 2) show that with soft selection, polymorphism can be maintained in an infinite metapopulation, provided that selection is sufficiently strong. With symmetric selection (s_1=s_2), this requires Ns>Ns_{crit}=$1/2\log\left[\frac{1-\rho}{\rho}\right]$+$Nm(1-2\rho)$; the first term is derived from the fixed-state approximation, in the limit $Nm\ll 1$, and the second from the deterministic model, which requires s>$m(1-2\rho)$ for polymorphism. In a metapopulation with a finite number of demes, variation must ultimately be lost: we must include mutation to allow a non-trivial stationary state. In this section, we examine how the outcome depends on the relative rates of selection and drift (Ns) and on the relative rates of mutation and migration (μ/m). In particular, we show that with sufficiently many demes, the outcome is insensitive to the mutation rate.

Figure 5 shows the stationary state in the limit of small Nm, derived using the fixed-state approximation. The top row of fig. 5 shows how the fraction of demes fixed at equilibrium depends on the strength of selection; the focal allele is favoured in 20% of demes (blue), and disfavored twice as strongly in 80% of demes (i.e. Ns_2=$-2Ns_1$; red). When mutation is appreciable (μ/m=0.05, fig. 5a), the allele is unlikely to be lost by chance, and so the equilibrium is insensitive to the number of demes, and close to the solution for an infinite population: results for $50,\ldots,400,\infty$ demes are superimposed, and almost indistinguishable. When selection is strong (right of fig. 5a and 5b), all demes are fixed for the favoured allele, whereas when selection is negligible, on average half of the demes are fixed for each allele. (Mutation is assumed symmetric). In-between (0.1<Ns<1), the allele favoured in the rare habitat becomes rare, being pulled to low frequency by migration from the commoner habitat, where it is more

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.17.460820; this version posted September 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

strongly disfavoured. When mutation is weak relative to migration (as is likely in nature), this pattern is exaggerated (μ/m=0.0005; fig. 5b). Above a critical value, $Ns_{crit}\sim(1/2)\log\left[\frac{1-\rho}{\rho}\right]\sim$ 0.7, polymorphism can be maintained by divergent selection, despite drift and gene flow. The equilibrium for an infinite population (purple) gives an upper bound, but stochastic loss from a finite set of demes reduces the expected frequency, and increases the critical Ns_{crit} (dashed lines around $Ns\sim1$, for $50, 100, \ldots$ demes). There is a wide region ($0.03<Ns<0.7$) where the allele is almost absent, being swamped by gene flow. However, for very weak selection, the frequency of the allele increases towards the symmetric neutral equilibrium at 0.5. (In this regime, the frequencies in the two habitats are almost identical, and cannot be distinguished in the figure). In this regime (left of fig. 5a), although selection is negligible within demes ($Ns<0.1$), migration is much faster than mutation, and so selection over the whole metapopulation is effective in eliminating the allele that is deleterious in most demes. When $\mu\ll m$, $Ns<0.1$ (left of fig. 5a), selection is more effective at the level of the whole metapopulation in the habitat which has more demes.

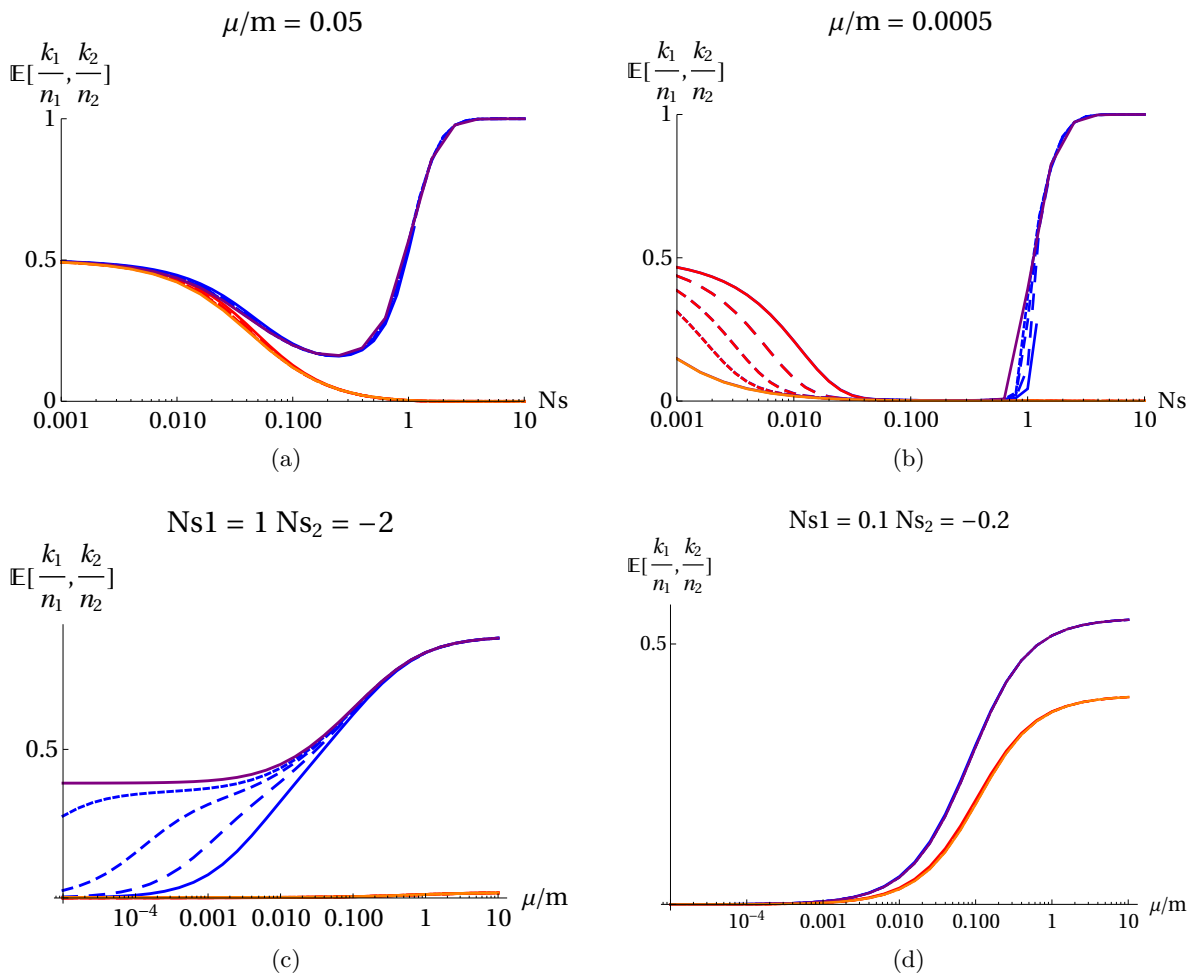

Figure 5: The fraction of demes fixed in the two habitats (blue, red), as a function of selection strength (Ns, top row) and the rate of mutation, relative to migration (μ/m, bottom row). The focal allele is favoured by selection Ns_1 in 20% of demes (blue), and disfavoured by selection $Ns_2=-2Ns_1$ in 80% of demes. In each plot, equilibria for 50, 100, 200 and 400 demes are superimposed (solid, dashed,... dotted lines), together with the limit for an infinite metapopulation (purple, orange).

The bottom row of fig. 5 shows the dependence on the relative rates of mutation versus

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.17.460820; this version posted September 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

migration, μ/m. With high mutation rates, the equilibrium approaches a fraction $\mathbb{E}[k/n]=1/(1+e^{-2Ns})$, given by the fixed-state approximation. There is strong divergence when $Ns_1=1$ (right of fig. 5c), and weaker divergence when selection is weak (fig. 5d, $Ns_1=0.1$). With moderately strong selection (fig. 5c), the allele that is less favoured overall is lost from the common habitat, independent of the number of demes and mutation rate (orange line). In the rare habitat, with weak mutation (left of fig. 5c), the locally favoured allele can be fixed in nearly half the demes in an infinite metapopulation (purple), but tends to be lost by chance from finite metapopulations, even with several hundred demes (blue). When selection is weak relative to local deme size (fig. 5d), selection can still be effective over the whole metapopulation, eliminating the allele that is disfavoured overall (left of fig. 5d). However, when mutation becomes comparable with migration, polymorphism is maintained by mutation pressure, with some bias between habitats caused by weak selection (right of fig. 5d).

We focus on the regime with moderately strong selection, comparable to drift ($Ns_1\sim1$), and weak mutation. This corresponds to the right half of fig. 5b ($0.1<Ns_1$), and the middle of fig. 5c ($10^{-4}<\mu/m<0.1$). Then, as long as mutation is not extremely small, and there are enough demes, the stationary state is close to that in an infinite metapopulation (compare blue dashed with purple lines in fig. 5c). However, note that with weak mutation ($\mu/m\sim10^{-4}$–10^{-3}, say), the locally favoured allele tends to be lost even when there are several hundred demes.

Loss of diversity in a finite metapopulation

When deme sizes are fixed, and numbers of migrants are low enough that loci are typically fixed for one or other allele, the state of the metapopulation at each locus can be described by the number of demes, k_i, in each habitat, i, that are fixed for the '1' allele. The distribution of k_i evolves according to a transition matrix, and each locus follows an independent realisation of the same stochastic process. Figure 6 compares the dynamics of this fixed-state approximation with simulations, to illustrate the accuracy of the fixed-state approximation. For the low migration rate $Nm=0.05$ assumed here, there is reasonable agreement; with $Nm=0.01$, agreement is very close (fig. A2). Variation is lost faster than predicted by the fixed-state approximation, because migration tends to swamp adaptive divergence. The timescale is inversely proportional to m, and is therefore slow. Here, we are focussing on the slow loss of adaptation through random drift in small populations; with higher migration rates, swamping by gene flow causes additional, faster, degradation.

Note that because the number of demes is limited, and because each deme flips between fixation for alternative alleles, there is substantial variability in average allele frequency between loci (grey lines). Therefore, adaptation is lost slightly faster in a finite than in an infinite metapopulation (compare black and magenta lines in fig. 6, which both derive from the fixed-state approximation). Nevertheless, the overall mean, averaged over 40 loci, changes smoothly and predictably (red curves in fig. 6). We assume no mutation, and so all variation will inevitably be lost. However, because the total population is large ($100\times50=5000$ individuals), and because the very low migration rate increases the effective size of the whole metapopulation, loss across the whole metapopulation is extremely slow: none of the 40 loci fix during the 10^4 generations shown here.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.17.460820; this version posted September 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

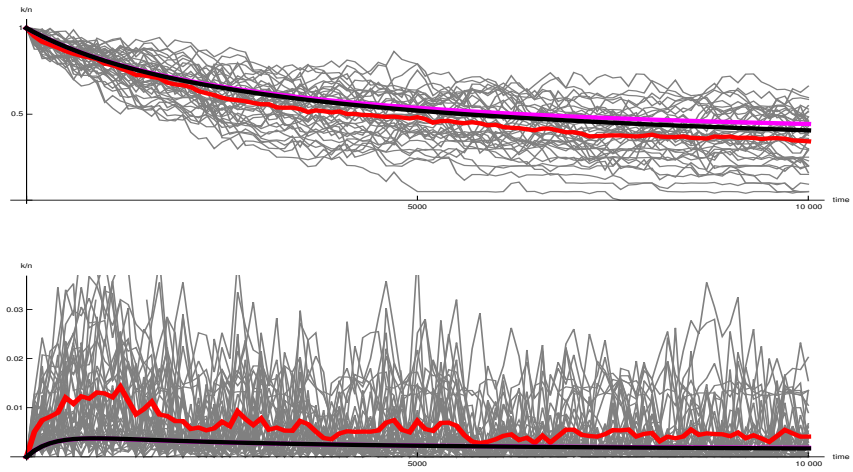

Figure 6: Loss of diversity in a metapopulation of 100 demes, which is initially perfectly adapted. Mean allele frequency is plotted against time, in the 20 demes where the focal allele is favoured (top), and the 80 demes where it is not (bottom). Thin grey lines show allele frequencies at 40 loci, averaged over demes; the red line shows the overall mean. The black curve shows the fixed-state approximation, for a finite metapopulation, and the magenta line, for an infinite metapopulation. Simulations are for $N = 50$, $Nm = 0.05$, $s_{1,2} = \{0.02, -0.04\}$; thus, $Ns_{1,2} = \{1, -2\}$, so that selection and drift are of similar magnitude.

Response to changing conditions across the metapopulation

Figure 7 shows some examples of the response to a change in conditions across the whole metapopulation. We use the same baseline case as above, with the rarer habitat in ρ=0.2 of the demes. We consider global changes across the whole metapopulation; however, for the parameters we consider, the allele favoured in the rare habitat is always rare in the common habitat, and so results would be essentially the same if parameters changed only in the rarer habitat.

Figure 7a shows the consequences of a change in Nm from 0.05, to 1, and then back to 0.05, as in fig. 3a. Initially, conditions are the same as in fig. 6, except that we start with a fraction of demes fixed for the locally favoured allele, in the proportions predicted for an infinite metapopulation. This allele gradually declines, somewhat faster than predicted in the limit of small Nm (compare red vs black lines at left). After 10^4 generations, Nm increases to 1, and the rarer allele is rapidly swamped, over a timescale of $\sim 1/m$=50 generations. After 500 generations, variation persists at only 6 of the 40 loci, so that when Nm then decreases back to its original level, only those loci can recover. The dynamics at the remaining polymorphic loci are the same as before, with recovery over $\sim 1/m$=1000 generations, but there has been a drastic loss of variation during the brief period of swamping. This contrasts with fig. 3a, where variation could be restored at all loci, because only one deme had been swamped.

Figure 7b shows the response to changes in Ns, with Nm=0.05 throughout. The initial 10^4 generations have the same parameters as in fig. 7a, and so again, diversity is gradually lost, somewhat faster than predicted in the limit $Nm \to 0$. After 10^4 generations, selection increases by a factor 10, and all loci quickly shift to near-perfect adaptation. Despite the strong selection, there is still considerable variation in the rates of increase across loci, though the overall equilibrium is approached quite smoothly (red curve, $10^4 - 1.5 \times 10^4$ generations). After $15,000$ generations, selection weakens by a factor 100, to near-neutrality, and the rarer allele is lost. In the two later stages, the dynamics are closely predicted by the small Nm limit (black curves),

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.17.460820; this version posted September 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

which is based simply on the rates of substitutions in either direction.

Figure 7c shows a similar scenario, but with strong migration, Nm=1, which couples evolution in the different demes. We double the initial selection strength (i.e., Ns_1=2, Ns_2=−4), to ensure that polymorphism is maintained, despite high Nm. Mean allele frequencies fluctuate around the deterministic equilibrium, but there is sufficient drift that, occasionally, the allele is lost from all the demes (grey lines). Thus, after 10^4 generations, only 11 of the 40 loci are still polymorphic. Selection then strengthens ten-fold, and these polymorphic loci rapidly approach fixation, over a timescale of just a few generations (∼$1/s$=2.5). However, because only 11/40 loci remain polymorphic, the overall mean is ∼0.25 (red, middle). After 5,000 generations, selection returns to its initial value; as before, loci fluctuate around a metastable equilibrium, but occasionally fix. Thus, after 20,000 generations, only one locus remains polymorphic.

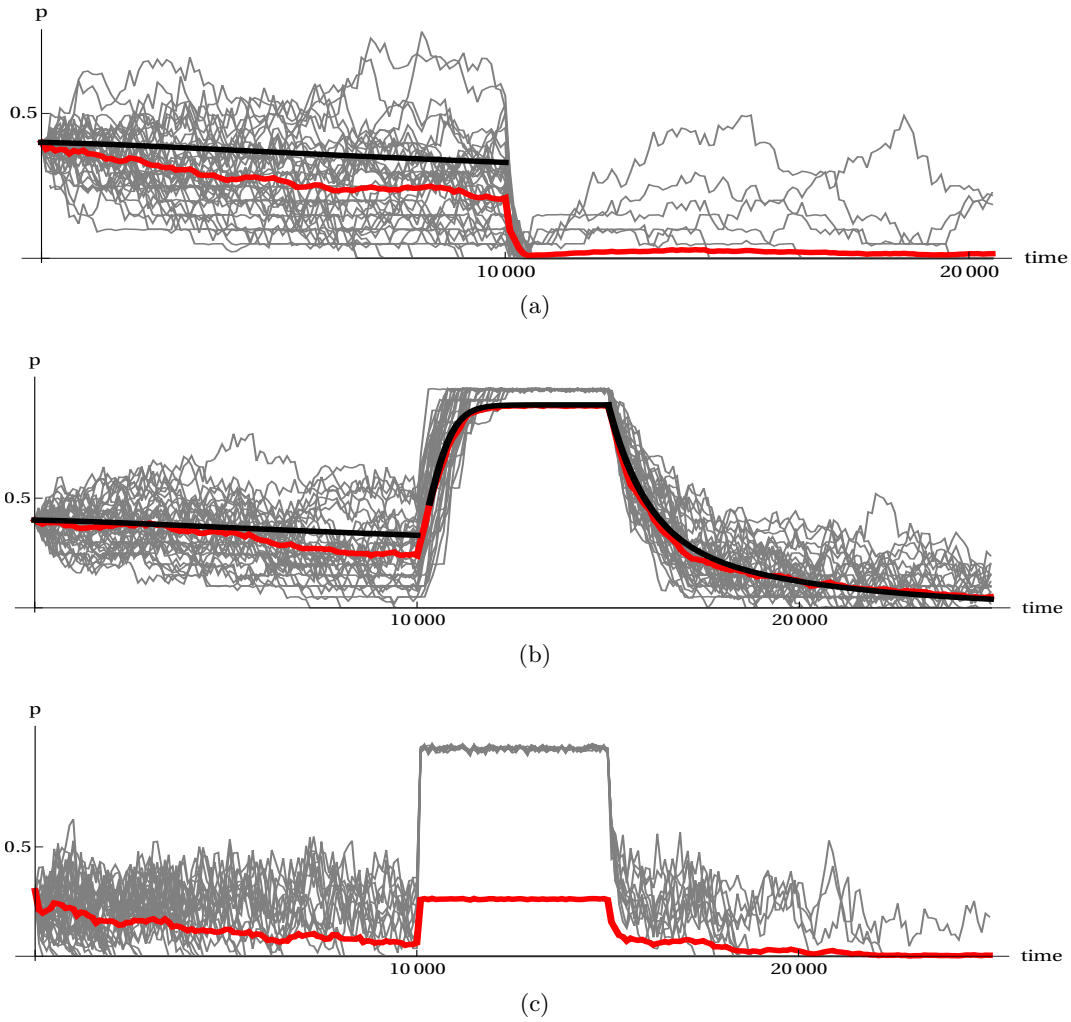

Figure 7: Response of a metapopulation to changing conditions. Grey lines show the allele frequencies, averaged over the 20 demes in the rare habitat, at each of 40 loci; the red line shows the overall mean in the rare habitat. The black line shows the prediction in the limit of small Nm. (a) Changing Nm, with Ns_1=1, Ns_2=−2. Initially, Nm=0.05; it increases to 1 after 10000 generations, and after a further 500 generations, returns to its initial value. (b) Changing Ns, with Nm=0.05. Initially, Ns_1=1, Ns_2=−2; after 10^4 generations, selection increases by a factor 10, and after a further 5000 generations, it decreases by a factor 100, to Ns_1=0.1, Ns_2=−0.2. (c) Changing Ns, with Nm=1. Initially, Ns_1=2, Ns_2=−4; after 10^4 generations, selection increases by a factor 10, and after a further 5000 generations, it decreases to the initial value, Ns_1=2, Ns_2=−4. As in fig. 6, simulations are for 100 demes of N=50, with 20 occupying one habitat, and 80 another.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.17.460820; this version posted September 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

If we compare the first 10^4 generations in fig. 7b and 7c, we see two effects of migration. Increasing migration from Nm=0.05 to Nm=1 somewhat reduces the mean allele frequency, because migration is now comparable to selection, and swamps local adaptation. It also couples together the allele frequencies in different demes, which makes it more likely that the allele is lost by chance form all the demes. These two effects increase the rate of stochastic loss, but the stronger selection in fig. 7c counters this loss. fig. A3 compares these distributions.

Discussion

Our analysis uses simulation, the diffusion approximation and the "fixed-state" approximation to understand how a finite metapopulation changes through time, as it responds to changes in both local and global conditions. The "fixed state" approximation applies either where variation is due to mutation (when it is plausible that $N\mu<1$ within local demes, or even for the whole populations), or when variation is maintained by divergent selection across the whole metapopulation, but migration is low relative to drift $Nm<1$.

When selection is weaker than drift (i.e. $Ns\ll 1$), polymorphism can only be maintained for a very narrow range of habitat proportions (fig. 1). However with strong selection, this range becomes much wider. When conditions in a single deme of the metapopulation change, the population responds on a short time scale of order $1/m$, simply because in the regime we study, local genetic variance is maintained by migration. Variation may be temporarily lost as local conditions change, but can quickly be recovered. On the other hand, when conditions change across the metapopulation, variation that was maintained by divergent selection can be permanently lost, and is only slowly recovered by mutation. Even under constant conditions, variation at a locus can be lost by chance, unless there are a very large number of demes.

To simplify our analysis, we assumed an island model, with a large number of spatially equivalent demes. This is unlikely to be the case in nature, but may nevertheless capture the behaviour of spatially extended populations if there is long-range migration, which can introduce locally adaptive alleles from a distant habitat. It may be that a leptokurtic dispersal distribution can allow efficient adaptation, if locally favoured alleles are not swamped, and yet can be recovered by occasional long-range migration [1, 10].

Our analysis can be further extended to hard selection, by including explicit density regulation; Szep et al [16] show that one can still apply the diffusion approximation, provided that growth rates are not too high. With hard selection, substitution rates depend on deme size through Ns, and through the number of immigrant alleles, $m\overline{Np}$, $m\overline{Nq}$. This dependence can be approximated by assuming that the population size is determined by the genetic load. Sachdeva et al [15] and Szep et al [16] refer to this as the "semi-deterministic" approximation which is accurate when demographic stochasticity is weak. One can apply the "fixed-state" approximation by further assuming that there are enough loci that the mean load is proportional to the mean across loci of the number of demes fixed for one or the other allele. The transition matrix can then be calculated as before, but is now a function of the population sizes in the two habitats, $\{N_1, N_2\}$ which both depend on the current state via the load. The key assumption here is that with enough loci, the population sizes change almost deterministically, following the distribution of states across loci. One complication with hard selection is the existence of multiple stable equilibria: changing conditions would not just cause equilibria to shift but also changes the rates of transitions between equilibria.

A key assumption in our analysis is that selection is directional: in a given environment, alleles experience a fixed selection pressure, which tends to drive out variation. More often, selection

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.17.460820; this version posted September 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

may favour an intermediate optimum for a quantitative trait, such that when the mean is well-adapted, alleles are close to neutral. Our modelling framework can describe this case, but it is much more complex, since many different allele combinations can achieve the same optimum. However, if selection on each allele is weak ($Ns<1$), then the infinitesimal model [3] applies, and can also describe the population dynamics [4]. Local adaptation may be possible under higher migration rates in such a regime.

In this work, we have introduced a novel approach to understanding the dynamical evolution of metapopulations. Although the full behaviour requires simulation, the diffusion approximation allows the stationary state to be calculated, and identifies the key dimensionless parameters. Moreover, when migration is rare, we can use a fixed state approximation that connects population genetics with models of adaptive walks [12].

Funding

This research was partly funded by the Austrian Science Fund (FWF) [FWF P-32896B].

References

[1] Atkinson, R. P. D., Rhodes, C.J, Macdonald, D.W., Anderson, R.M.. "Scale-Free Dynamics in the Movement Patterns of Jackals." Oikos, vol. 98, no. 1, July 2002, pp. 134–40. DOI.org (Crossref), https://doi.org/10.1034/j.1600-0706.2002.980114.x.

[2] Barton, N. H., and Rouhani, S. "Adaptation and the 'Shifting Balance.'" Genetical Research, vol. 61, no. 1, Feb. 1993, pp. 57–74. DOI.org (Crossref), https://doi.org/10.1017/S0016672300031098.

[3] Barton, N.H., Etheridge, A.M., Veber,A . "The Infinitesimal Model: Definition, Derivation, and Implications." Theoretical Population Biology, vol. 118, Dec. 2017, pp. 50–73. DOI.org (Crossref), https://doi.org/10.1016/j.tpb.2017.06.001.

[4] Barton, N.H., and Etheridge, A.M. "Establishment in a New Habitat by Polygenic Adaptation." Theoretical Population Biology, vol. 122, July 2018, pp. 110–27. DOI.org (Crossref), https://doi.org/10.1016/j.tpb.2017.11.007.

[5] Bell, G., and A. Gonzalez. "Adaptation and Evolutionary Rescue in Metapopulations Experiencing Environmental Deterioration." Science, vol. 332, no. 6035, June 2011, pp. 1327–30. DOI.org (Crossref), https://doi.org/10.1126/science.1203105.

[6] Bell, G., "Evolutionary rescue." Annual Review of Ecology, Evolution, and Systematics, vol 48, Nov 2017, pp. 605-627.

[7] Bialozyt, R., Ziegenhagen, B., Petit, R. "Contrasting Effects of Long Distance Seed Dispersal on Genetic Diversity during Range Expansion." Journal of Evolutionary Biology, vol. 19, no. 1, Jan. 2006, pp. 12–20. DOI.org (Crossref), https://doi.org/10.1111/j.1420-9101.2005.00995.x.

[8] Breed, M.F., Ottewell, K.M, Gardner, M.G., Lowe, A.J. "Clarifying Climate Change Adaptation Responses for Scattered Trees in Modified Landscapes: Adaptation and Scattered Trees." Journal of Applied Ecology, vol. 48, no. 3, June 2011, pp. 637–41. DOI.org (Crossref), https://doi.org/10.1111/j.1365-2664.2011.01969.x.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.17.460820; this version posted September 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[9] Case, T.J., and Taper M.L. "Interspecific Competition, Environmental Gradients, Gene Flow, and the Coevolution of Species' Borders." The American Naturalist, vol. 155, no. 5, May 2000, pp. 583–605. DOI.org (Crossref), https://doi.org/10.1086/303351.

[10] Fric, Z., and Martin K. "Dispersal Kernels of Butterflies: Power-Law Functions Are Invariant to Marking Frequency." Basic and Applied Ecology, vol. 8, no. 4, July 2007, pp. 377–86. DOI.org (Crossref), https://doi.org/10.1016/j.baae.2006.06.005.

[11] Kirkpatrick, M., and Barton, N.H. "Evolution of a Species' Range." The American Naturalist, vol. 150, no. 1, July 1997, pp. 1–23. DOI.org (Crossref), https://doi.org/10.1086/286054.

[12] Orr, H. A. "The Population Genetics of Adaptation: The Distribution of Factors Fixed during Adaptive Evolution." Evolution, vol. 52, no. 4, Aug. 1998, p. 935. DOI.org (Crossref), https://doi.org/10.2307/2411226.

[13] Paulose, J., and Oskar, H. "The Impact of Long-Range Dispersal on Gene Surfing." Proceedings of the National Academy of Sciences, vol. 117, no. 14, Apr. 2020, pp. 7584–93. DOI.org (Crossref), https://doi.org/10.1073/pnas.1919485117.

[14] Polechová, J. "Is the Sky the Limit? On the Expansion Threshold of a Species' Range." PLOS Biology, edited by Pedro Jordano, vol. 16, no. 6, June 2018, p. e2005372. DOI.org (Crossref), https://doi.org/10.1371/journal.pbio.2005372.

[15] Sachdeva, H., Olusanya, O., Barton, N.H. 2021. Genetic load and extinction in peripheral populations: the roles of migration, drift and demographic stochasticity. Phil. Trans. Roy. Soc. B (subm.)

[16] Szép, E., Sachdeva, H., Barton, N.H. "Polygenic Local Adaptation in Metapopulations: A Stochastic Eco-evolutionary Model." Evolution, vol. 75, no. 5, May 2021, pp. 1030–45. DOI.org (Crossref), https://doi.org/10.1111/evo.14210.

[17] Trubenová, B., Krejca, M.S., Lehre, P.K., Kötzing, T. "Surfing on the Seascape: Adaptation in a Changing Environment." Evolution, vol. 73, no. 7, July 2019, pp. 1356–74. DOI.org (Crossref), https://doi.org/10.1111/evo.13784.

[18] Uecker, H., Otto, S.P, Hermisson, J. "Evolutionary Rescue in Structured Populations." The American Naturalist, vol. 183, no. 1, Jan. 2014, pp. E17–35. DOI.org (Crossref), https://doi.org/10.1086/673914.

[19] Wright, S. "The Distribution of Gene Frequencies in Populations." Genetics, vol. 23, no. 6, June 1937, pp. 307–20. DOI.org (Crossref), https://doi.org/10.1073/pnas.23.6.307.

[20] Wright, S. "The Distribution of Gene Frequencies in Populations." Science, vol. 85, no. 2212, May 1937, pp. 504–04. DOI.org (Crossref), https://doi.org/10.1126/science.85.2212.504.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.17.460820; this version posted September 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Appendix

Accuracy of the fixed-state approximation

Here, we compare the the mean allele frequency in an infinite metapopulation under the diffusion approximation with the fixed-state approximation for different Nm values. As expected, the accuracy of the fixed state approximation holds only for small Nm.

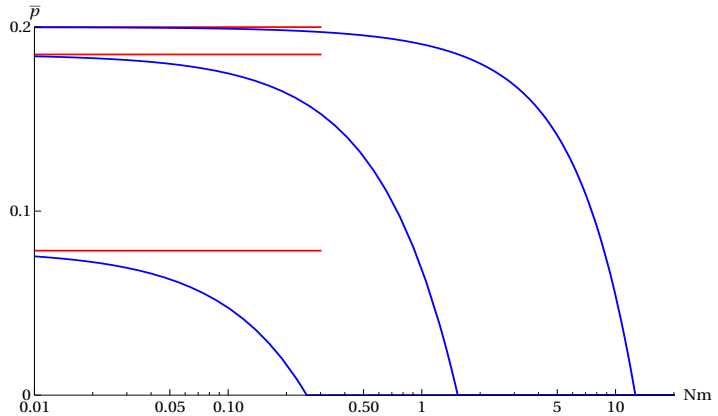

Figure A1: The mean allele frequency in an infinite metapopulation, plotted against Nm; $\rho = 0.2$, Ns_1, $Ns_2 =$ 1, -2 (lower curve) 2, -4 (middle curve) or 10, -20 (upper curve). The fixed-state approximation, which applies for small Nm, is shown by the red lines.

Loss of diversity from a finite population, with Nm=0.01

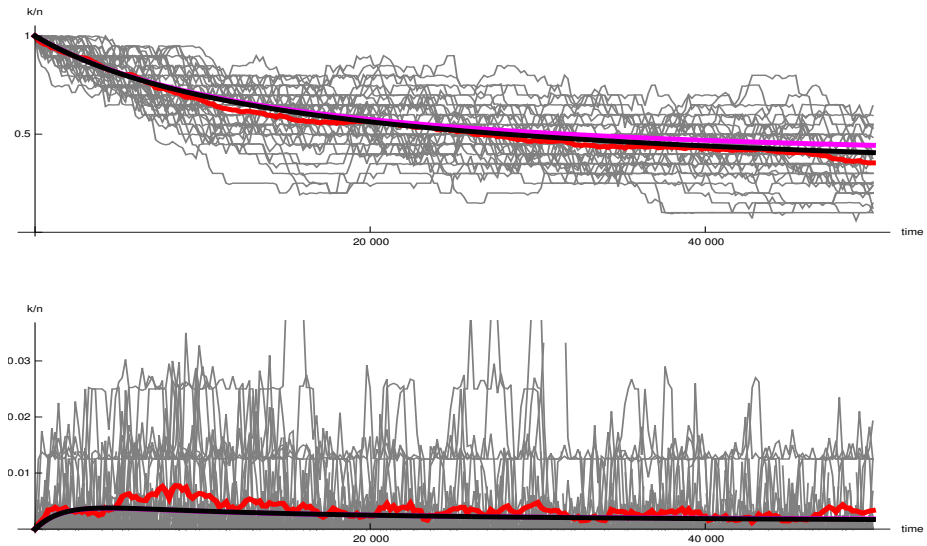

Figure A2: This is identical to fig. 6, except that Nm=0.01, and the timescale is correspondingly longer. The fixed-state approximation is more accurate with a lower number of migrants.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.17.460820; this version posted September 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Distribution of mean allele frequency

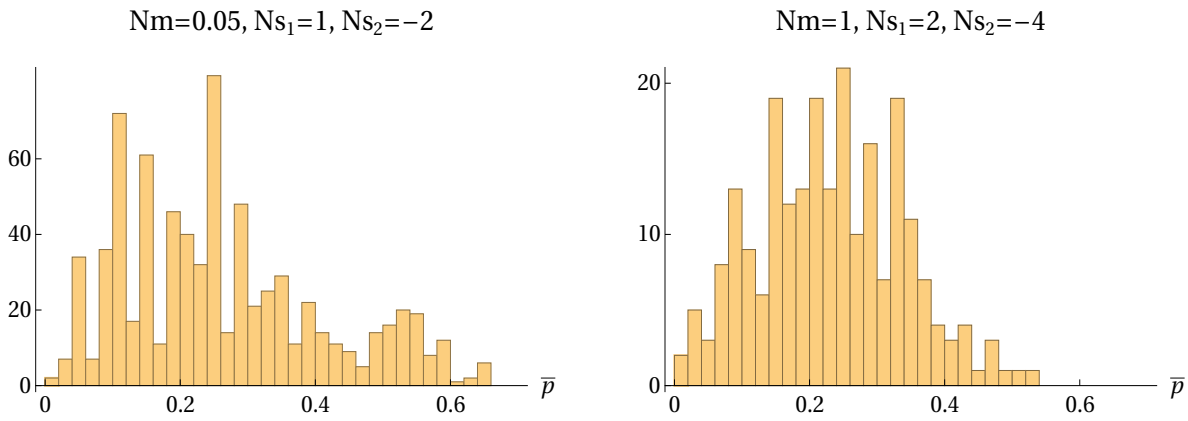

Figure A3: The distribution of allele frequencies, averaged over the 20 demes in the rare habitat, conditional on polymorphism, and accumulated over generations $8,000$, $8,100$, to $10,000$; taken from the simulations in fig. 7b, 7c.