Journal of Evolutionary Biology, 2024, **37**, 605–615
https://doi.org/10.1093/jeb/voae052
Advance access publication 29 April 2024
Review

JOURNAL OF
Evolutionary Biology

eseb

Limits to species' range: the tension between local and global adaptation

Nicholas Barton

Institute of Science and Technology Austria, Klosterneuburg, Austria

Corresponding author: Nicholas Barton, Institute of Science and Technology Austria Am Campus 1, Klosterneuburg 3400, Austria. Email: nick.barton@ist.ac.at

Abstract

We know that heritable variation is abundant, and that selection causes all but the smallest populations to rapidly shift beyond their original trait distribution. So then, what limits the range of a species? There are physical constraints and also population genetic limits to the effectiveness of selection, ultimately set by population size. Global adaptation, where the same genotype is favoured over the whole range, is most efficient when based on a multitude of weakly selected alleles and is effective even when local demes are small, provided that there is some gene flow. In contrast, local adaptation is sensitive to gene flow and may require alleles with substantial effect. How can populations combine the advantages of large effective size with the ability to specialise into local niches? To what extent does reproductive isolation help resolve this tension? I address these questions using eco-evolutionary models of polygenic adaptation, contrasting discrete demes with continuous space.

Keywords: species' range, eco-evolutionary model, information, migration load, reproductive isolation, local adaptation, metapopulation, cline, polygenic traits, gene flow

Introduction

"What limits a species range?" is a large question, which involves several fundamental concepts and cuts across many overlapping fields (Antonovics, 1976). Although we have a detailed understanding of the principles by which populations adapt via natural selection, together with the ancillary processes of mutation, random drift, recombination, and gene flow, we do not know what limits the efficiency of natural selection. As argued below, asking what limits a species' range requires that we ask about what limits adaptation in general.

The extraordinarily complex and varied adaptations that we see have accumulated over four billion years of evolution. The core biochemical machinery that accurately replicates DNA, transcribes it into RNA, and then translates that into the protein sequences responsible for most organismal function, is shared by all organisms. Eukaryotes share a complex cellular architecture, which allows regular meiotic sex. Several groups have elaborated mechanisms of cellular differentiation and for the reliable development of multicellular organisms. These have elaborated further into complex general-purpose organs such as the adaptive immune system and the brain that allow these words to be written and understood.

Such global adaptations have accumulated over geological timescales and are shared by many individuals and species. Although we know the history of their origin in some detail, we are very far from being able to estimate, from first principles, how much complexity could arise in the time available or how much can be maintained despite mutation and random drift. Understanding local adaptation—the diverse ways by which populations and species find different ways to make a living—seems more tractable because it happens over shorter times, and in repeated instances. However, local adaptation raises the same question about how quickly a population can adapt and also new questions about how it can do so despite gene flow and an unpredictably changing environment.

These are fundamental questions for evolutionary biology, but are of specific concern for some subfields. Most immediately, conservation of endangered species requires us to understand what currently limits their range, whether they will be able to adapt to changing conditions, and how we might assist that adaptation (Frankham et al., 2014; Parmesan et al., 2023). Speciation may begin with local adaptation and ultimately require that each species efficiently exploits a distinct niche; there is a continuum between balanced polymorphism within a well-mixed population, adaptation to local conditions with no additional reproductive isolation, and a set of distinct species that may nevertheless occasionally exchange genes.

Evolutionary genetics is currently devoted to trying to make sense of the abundance of sequence data (e.g., Aeschbacher et al., 2017; Elyashiv et al., 2016; Jones et al., 2012). This allows us to infer past population structure (i.e., demography and gene flow) and to detect signs of adaptation (i.e., selective sweeps) and degeneration (e.g., in small populations or at the range edge; Peischl et al., 2015). However, many of the key quantities require direct (and laborious) measurements of

Received November 18, 2023; revised April 02, 2024; accepted May 08, 2024

traits and demography. It may be that polygenic adaptation leaves no detectable trace in genome sequences.

Most obviously, the question of what limits a species' range lies at the confluence of evolution and ecology: how does the genetic composition of a population affect its abundance and spatial distribution, and how does population structure affect adaptation (Antonovics, 1976; Roughgarden, 1979)? Genetic models usually assume constant numbers, whilst ecological models assume genetic homogeneity. Yet, here, we must specify how genotype frequencies influence population dynamics and, beyond that, how populations diverge into different niches when their range is limited by biological competitors.

Most population genetic models are deliberately simple—aiming to build a basis for intuition rather than represent reality. When we consider the limit to a species' range, it is important to be aware of when simple models can be extrapolated and when they may mislead. Most theories are based on a single deme that receives migrants, or its extension to an island model with limited gene flow but no explicit spatial structure (e.g., Gomulkiewicz et al., 1999; Pease et al., 1989). Then, local adaptation is only possible if selection is stronger than migration. Yet, in a spatial continuum, a population can always adapt to local conditions, provided that they extend over a sufficient scale, of at least $\sim 1/\sqrt{s}$ dispersal ranges (Nagylaki, 1975; Slatkin, 1973). Thus, local adaptation seems much easier over an extended spatial range and may involve more weakly selected alleles. Similarly, most population genetic models assume a single gene, in which, again, strong selection seems necessary for local adaptation. Yet, the mean of a quantitative trait can readily adapt to a local optimum, provided there is sufficient genetic variance, even if that variance is based on alleles of very small effect (Barton, 1999; Polechová, 2018; Tufto, 2000; Yeaman, 2022).

We usually consider adaptation to a single trait. This may be reasonable if there is a single environmental gradient, which defines the relevant trait. Yet, in reality, organisms may be faced with a multitude of challenges from the environment. Finally, we usually model a single, evolving population or a set of fully isolated species. How does a low level of gene flow influence adaptation? The usual situation may be that of a set of ill-defined populations with varying (and evolving) degrees of interconnection.

In the following, I elaborate on this summary argument and try to bring together some of the diverse theory concerning the limits to local and global adaptation. Rather than reviewing the extensive theoretical and empirical literature, this article is primarily an argument about how we can understand the tension between global and local adaptation and the implications of this tension for range limits and for the role of the biological species.

Limits to adaptation

Physical, organismal, and population genetic constraints

Ultimately, there are physical limits to the environments in which an organism can survive and reproduce. Yet, these are mediated by the particular organism that we consider. We are surprised by the extraordinary range of temperatures that some bacteria can endure, and the energy sources that they can exploit—a range unavailable to any eukaryote. In principle, any extant organism could evolve into any other, since they are connected via a chain of more or less fit ancestors. Yet, this would require on enormous time, and a sequence of appropriate environments. In practice, we consider the limits to adaptation of a particular organism, and consider only the range of phenotypes that are accessible to it— a range that is not at all easy to define.

Such constraints are sometimes discussed in terms of "trade-offs", but that term can be misleading. For example, it is impossible for an organism to maximise both survival and reproduction: if we imagine the set of possible life histories to be bounded by a convex curve, then a population that maximises fitness will evolve along that boundary, and the two fitness components will be "traded off". However, mutations may tend to reduce both components, leading to a positive genetic correlation between them (Partridge & Barton, 1993, Figure 2). Thus, it seems better to think in terms of the constraints on what can evolve, rather than *necessary* trade-offs between fitnesses components, or fitnesses in different environments.

We may make more progress in considering limits to adaptation based on the fundamental population genetic processes, which apply to all organisms, and which can be quantified. This will at least identify the key parameters. Traditional population genetic arguments are primarily based on the idea of "load": the loss of fitness due to evolving by natural selection, rather than by some ideal process that immediately fixes the best genotype (Felsenstein, 1971). This loss is constrained by the reproductive capacity of the organism, and so sets limits to adaptation. Thus, the loss of mean fitness during the fixation of a newly beneficial allele limits the substitution rate (Haldane, 1957; Kimura, 1961), an argument used as evidence for neutrality (Kimura, 1983); the segregation of unfit homozygotes limits the number of balanced polymorphisms (Crow, 1958); and the loss of fitness due to random drift around an optimum limits the number of traits that can be optimised (Barton, 2017). Yet, these arguments can be circumvented by assuming a certain kind of epistasis (Kondrashov, 1988): if unfit genotypes can be eliminated together, then less reproductive excess is needed. It is unclear whether such negative epistasis is widespread, or why it should evolve, but nevertheless, traditional load arguments are not decisive.

Limits to accumulation of information

We can make the limits to the power of natural selection precise by asking how much information it can accumulate. This idea traces back to Kimura (1961), who saw that the fixation of one particular amino-acid in the population increases information, in the sense defined by Shannon (1948) and Crow (2001); this information increase equals $\log(1/p_0)$, where the initial frequency p_0 is the probability that the mutation would be fixed by chance. Kimura (1961) related this to the substitution load, and following Haldane (1957), estimated that the mammalian genome could have accumulated at most $\sim 10^8$ bits of information since the Cambrian, roughly corresponding to the actual number of amino-acid substitutions. We can now see that this argument is flawed, because on the one hand, much information must be associated with non-coding regulatory evolution, and on the other, because Haldane's "cost of selection" can be greatly reduced by certain kinds of epistasis. For example, truncation selection (in some sense, the most extreme form of epistasis) is more efficient.

Nevertheless, the measure of information introduced by Kimura (1961) extends to give a remarkably general constraint on how effectively selection can establish a specific

(and *a priori* improbable) set of genotype or phenotypes. The measure generalises to:

$$D = \mathbb{E}\left[\log\left[\frac{\psi_S}{\psi_\mathcal{N}}\right]\right]. \quad (1)$$

which is the expected log ratio between the probability of the population state under selection, ψ_S, and the neutral probability with no selection, $\psi_\mathcal{N}$. In the simple case considered by Kimura (1961), where the state of the population is described by the allele frequency, $\psi_S = 1$, and $\psi_\mathcal{N} = p_0$. Hledík et al. (2022) show that quite generally, the gain in information per generation can be no greater than the product of population size and variance in fitness:

$$\Delta \mathrm{D} \leq 2N\,\mathrm{var}(W)/\log(2). \quad (2)$$

Intuitively, the information encoded in the genome by natural selection comes from the number of events in which individuals may or may not reproduce. (Note that the bound is far larger than Kimura's, by a factor of order population size, N)

This is a measure of improbability, and not necessarily of adaptedness: the genotype that is established must have had relatively fit ancestors, but the information measure makes no claim about its current fitness. One can make a connection through the "free fitness" F, introduced by Iwasa (1988), which is the difference between the mean fitness, $\mathbb{E}[W]$, and the information, D, divided by $2N$:

$$F[\psi] = \mathbb{E}[W] - \frac{1}{2N}D. \quad (3)$$

The free fitness increases through the combined effects of mutation, selection and drift, and is maximised at the stationary distribution. This implies that for a given mean fitness, evolution minimises D, or in other words, tends towards a state where fit genotypes are encoded by the largest number of genotypes—a phenomenon termed "survival of the flattest" (Wilke et al., 2001). Moreover, the changing distribution of allele frequencies is, at any time, close to that which minimises D, conditional on the constrained variables (Bod'ová et al., 2016).

What genetic basis for fitness variation is most efficient, approaching the upper bound for information accumulation? In the simplest case, where individual alleles are uniformly selected, this bound is approached when allelic effects on fitness are infinitesimally small (Hledík et al., 2022). The intuition here is simple: if an allele is certain to be fixed (because $Nsp_0 >> 1$), then reproductive capacity would be "wasted" by selecting it more strongly. Thus, slightly increasing the probability of fixation of very many alleles is a more efficient way of establishing improbable genotypes. The argument here is close to that of Robertson (1960), who showed that in the infinitesimal limit, the ultimate response to selection is maximised when genetic variance is lost primarily by random drift. This argument extends to allow epistasis, in which case the ultimate change in trait mean is determined by the components of trait variance (Paixão & Barton, 2016).

The conjecture is that for a given population size and fitness variance, selection accumulates information most rapidly when it is polygenic. This has only been demonstrated for simple cases, and remains an open question. Note that the argument for the efficiency of small effects is distinct from Fisher's (1930) geometric model, where he argued that small changes minimise pleiotropic side-effects. It contrasts with Lynch's (2007) argument that subtle molecular interactions, which only slightly affect fitness, can only evolve in very large populations—as is suggested by single-locus theory that requires $Ns >> 1$ for selection to be effective, relative to drift. However, as is clear from the infinitesimal model (Barton et al., 2017), selection on polygenic traits can be effective even if $Ns << 1$ for individual alleles (Robertson, 1960; Charlesworth, 2013). Nevertheless, the bound on information gain outlined here (Equation (2)) implies that very large populations are required to assemble complex adaptations, and suggests that these adaptations may evolve most efficiently when they have a diffuse genetic basis (Barton, 2022).

Limits to local adaptation in discrete demes

Species must maintain the adaptations that have evolved over their entire ancestry, and must continually adapt to their changing physical and biotic environment. In order to expand their range, they may also adapt to the diverse conditions that they encounter. Such local adaptations are limited by population size and by variance in reproductive capacity, just as for global adaptations, as discussed above. Population subdivision has rather little effect on global adaptation, provided that the number of migrants exchanged between demes is sufficient (Whitlock, 2002); indeed, the fixation probability of a globally favoured allele is invariant to subdivision, provided that gene flow is conservative (Maruyama, 1974; Nagylaki, 1982). Constraints on local adaptation are more severe, since selection acts only within a small fragment of the population. In addition, there is an extra constraint from gene flow, which (although needed to provide genetic variation) tends to swamp local adaptation.

In the simplest case, where a fraction m of a deme is replaced by incoming migrants in each generation, a locally favoured allele can be maintained, provided that its selective advantage, s, is greater than the migration rate ($s > m$). (If there is emigration to other demes, then the allele is more likely to be maintained, as long as selection against it elsewhere is not too strong). However, the presence of locally deleterious alleles reduces mean fitness by $2m$, independent of selection (provided $s > m$); this "migration load" is analogous to the mutation load (Bolnick & Nosil, 2007; Haldane, 1937); it may be more severe, because migration rates are typically much higher than mutation rates, and because this load accumulates with the number of local adaptations that are maintained despite immigration. Assuming multiplicative effects on fitness, and linkage equilibrium, mean fitness is decreased by a factor $\exp(-2nm)$, where there are n locally adapted loci. This may not imperil the population, provided that there is rapid growth from low density, such that the fittest genotype would increase at a rate $r > 2nm$. Moreover, as with other load arguments, the constraint can be circumvented by negative epistasis (e.g., truncation selection that removes the least fit genotypes), or by linkage (see "Reproductive isolation" section) .

The single-locus theory suggests that polygenic traits cannot sustain local adaptation, because each locus would be too weakly selected to resist swamping, and because each would contribute to the migration load. However, that is incorrect: slight shifts in allele frequency at many loci can sustain adaptive changes in the trait mean. Suppose that an additive trait has genetic variance V_g, and is subject to stabilising selection V_s, such that fitness decreases with distance from the optimum (more precisely, fitness is proportional to

$\exp\left[-(z-z_{\text{opt}})^2/(2V_s)\right]\Big)$. Then, the trait mean at equilibrium will be a compromise between the local optimum, z_{opt}, and the mean in the rest of the metapopulation, $\bar{z}_S$:

$$\bar{z} = \frac{m\bar{z}_S + (V_g/V_s)\, z_{\text{opt}}}{m + (V_g/V_s)}. \tag{4}$$

(Hendry et al., 2001; Yeaman, 2015). There is now no threshold beyond which local adaptation is swamped, but rather, a balance between local and global adaptation that depends on the relative values of (V_g/V_s) (twice the loss of mean fitness due to genetic variance), versus m (the fraction of migrants). Distinct local adaptation can be maintained by slight shifts in allele frequency at very many loci, rather than near-fixation of alternative alleles. However, this comes with the cost of polymorphism, which causes variance around the local optimum, reducing fitness by $V_g/(2V_s)$.

Random drift further degrades local adaptations: even if selection is quite strong ($Ns >> 1$), as migration approaches the critical threshold, random fluctuations reduce genetic variance and make loss of the locally favoured allele likely (Yeaman & Otto, 2011). For polygenic traits, there is no sharp migration threshold, but random drift reduces genetic variance, which shifts the mean further from its optimum (Equation (4)).

Local adaptation over an extended range

Most recent eco-evolutionary theory has modelled discrete demes and has emphasised the difficulty in adapting to local conditions despite gene flow. Yet, most species extend over a broad geographic range. Then, gene flow can be modelled as diffusion of genes at a rate σ, the variance of distance between parent and offspring along some axis. This is as much an approximation as is dividing species into discrete demes; in reality, collective movement through large-scale range changes are important, especially in the longer term. However, contrasting diffusion in one or two dimensions against island models with no explicit spatial structure at least somewhat expands our imagination, even though it does not capture the more drastic movements that must eventually dominate.

Clines that involve a balance between selection and dispersal necessarily span a scale $\sim \sigma/\sqrt{2s}$ (Fisher, 1937; Haldane, 1948; Slatkin, 1973). Thus, if selection favours an allele over at least this scale, it can be maintained despite gene flow from elsewhere (the critical scale of course depends on the strength of selection for and against the allele in different places; Nagylaki, 1975). This result derives from dimensional arguments, and is robust, applying over a wide range of particular models. The same argument extends to clines in a quantitative trait, maintained by spatial variation in the optimum: then, the typical scale is $\sigma\sqrt{V_s/V_g}$ (Barton, 1999). Laroche and Lenormand (2023) connect these two regimes, extending Nagylaki's (1975) approach to a model where the trait optimum changes abruptly at a habitat boundary.

Because the spatial scale is inversely proportional only to the square root of selection, modest selection can maintain divergence over just a few tens of dispersal ranges. This makes it easier to envisage local adaptation than in island models with extensive gene flow. More convincing than the theoretical argument is the common observation of narrow clines, sometimes only a few tens of metres across (e.g., in *Littorina*, Johannesson et al., 1995; *Agrostis*, Caisse & Antonovics, 1978), and, in most cases, typically much narrower than the species' range.

Eco-evolutionary models

The joint distribution of population size, allele frequencies, and trait mean

In order to understand how local adaptation influences population size, we need to combine ecology with evolution. Although this relation, mediated by the dependence of absolute fitness on both density and genotype, has long been understood (e.g., Fisher, 1930, chapter 2; Haldane, 1956), explicit models are relatively recent (see Roughgarden (1979) and Pease et al. (1989) for the earliest models, and Débarre et al. (2013) for a recent review). Seminal works include Holt and Gomulkiewicz (1997) and Kawecki et al. (1997), who ask when a locally favoured allele can establish; Holt et al. (2003) extend this to a quantitative trait. Lynch et al. (1995) simulate how accumulation of deleterious mutations can lead to a population collapse. Ronce and Kirkpatrick (2001) analyse the positive feedback between maladaptation and reduced population size, which exacerbates collapse of local populations, assuming a trait with constant genetic variance. Tufto (2000, 2001) considers the same question, but allows for the inflation of genetic variance by linkage disequilibrium, under the infinitesimal model. These analyses are primarily deterministic, and so do not include stochastic limits to the efficiency of selection in small populations.

Szep et al. (2021) included the effects of random genetic drift and demographic stochasticity by extending Wright's (1937) formula for the stationary distribution of allele frequencies in the island model to include fluctuations in local deme size, N. This requires that selection be independent of density, and that alleles are combined randomly, in *linkage equilibrium*, so that we need only follow the joint distribution of local population size and allele frequencies. The log fitness (i.e., the growth rate in continuous time) is $r(N)+r_g(p)$, where $r(N)$ is the log fitness of the fittest genotype, which decreases with N, and $r_g(p) < 0$ is the genetic component of fitness, which decreases with the frequency of locally deleterious alleles. The stationary distribution is then:

$$\psi[N,p] \sim \psi_D[N]\psi_{\mathcal{N}}[p]e^{2N r_g[p]} \text{ where } \psi_{\mathcal{N}}[p] = \prod_{i=1}^{L} p_i^{2M\bar{p}_i-1} q_i^{2M\bar{q}_i-1}, \tag{5}$$

where p is the vector of allele frequencies p_i and q_i across multiple biallelic loci; $\bar{p}_i$ and $\bar{q}_i$ are the corresponding frequencies in the migrant pool. Here, $\psi_D[N]$ is the distribution of N for a (haploid) population fixed for the fittest genotype; $\psi_{\mathcal{N}}[p]$is the neutral distribution of allele frequencies, and M is the number of migrants entering per generation (assumed independent of the local deme size, N). The migrant pool consists of the whole metapopulation, assumed to consist of very many demes. These two distributions are coupled together by $\exp(2N r_g[p])$. If we consider the distribution of allele frequencies, this pulls the population towards fit genotypes, a pull which is more effective in large populations. If, on the other hand, we consider population size, increased mean fitness r_g, in turn, increases N (sometimes referred to as "hard selection"). Selection (represented by the sum of selection over loci r_g) may vary between demes.

When density regulation is strong, population size is insensitive to the allele frequencies, and in the limit is termed "soft selection". Then, under the assumptions of this model, loci evolve independently, and each can establish locally favoured alleles provided that their selective advantage is greater than the proportion of immigrants ($m = M/N$). However, if the total migration load is strong enough to reduce population size significantly ("hard selection"), then local adaptation will collapse if alleles are, in aggregate, too strongly selected. Thus, there is an intermediate selection strength (per locus) that is best able to sustain local adaptation—strong enough to resist swamping but not so strong that the net migration load ($\sim L\,m$, with L loci) extinguishes the population (Figure 6 in Szep et al., 2021). Such positive feedbacks, due to demographic effects, can make it difficult for a metapopulation to extend its range by adapting to rare habitats, despite gene flow from the better-adapted bulk of the population (cf. Lynch et al., 1995; Ronce & Kirkpatrick, 2001).

A similar model can be formulated for the simpler case of a single island that receives migrants from the mainland. We follow the mean of a quantitative trait under stabilising selection, assuming a constant genetic variance, and logistic density regulation, such that the baseline growth rate is $r_0(1 - N/K)$:

$$\begin{aligned}\psi\left[N,\bar{z}\right] &\sim \psi_D[N]\psi_{\mathcal{N}}\left[\bar{z}\right] e^{2N r_g\left[\bar{z}\right]} \text{ where } \psi_D\left[\bar{z}\right] \\ &\sim N^{2M-1}\exp\left[-\frac{r_0}{2K}(N-K)^2\right], \\ \psi_{\mathcal{N}}\left[\bar{z}\right] &\sim \exp\left[-\frac{M}{V_g}\left(z_s-\bar{z}\right)^2\right], r_g = -\bar{z}^2/\left(2V_s\right), \qquad (6)\end{aligned}$$

Here, the baseline demography ψ_D leads to a Gaussian distribution around the carrying capacity K, combined with a probability mass N^{2M-1}, which is concentrated near extinction if $M < 0.5$. The neutral trait distribution $\psi_{\mathcal{N}}$ is a Gaussian around the source value, z_S. The population will ultimately go extinct when $M << 1$, through demographic stochasticity, and increases gradually with migration pressure (upper curve in Figure 1, left). With immigration from a distinct source population, there is an optimal number of migrants, large enough to allow recovery from extinction, but not so large as to swamp local adaptation (lower curve in Figure 1, left; cf. Gomulkiewicz et al., 1999, Uecker et al., 2014). In this regime, the population fluctuates between near extinction and a well-adapted state (Figure 1, right). Note that, in contrast to the case where adaptation relies on discrete loci (Szep et al., 2021), there is no critical migration threshold. Moreover, the trait must be strongly selected, and strongly divergent, for the migration load to make extinction likely. (In Figure 1, stabilising selection reduces population growth rate by twice the baseline rate ($V_g/\left(2V_s\right) = 2r_0, z_s = 4\sqrt{V_g}$).

The continent-island scheme illustrated in Figure 1 is very much a "toy model", since it neglects the evolution of the genetic variance. We shall now discuss how the consequences of an evolving variance are, in some ways, better understood in continuous space.

Adaptation across continuous space

The apparent ease with which populations can adapt across a spatially continuous cline makes it hard to understand why species' ranges often end abruptly along an apparently continuous environmental gradient. Such sharp thresholds may be due to a positive feedback between adaptation and population size, as discussed above for a model of discrete demes: an influx of genes from better-adapted and hence denser regions causes maladaptation, lower density, and a stronger asymmetry in gene flow (Haldane, 1956). Kirkpatrick and Barton (1997) showed (assuming fixed genetic variance) that whilst a species can in principle adapt to an arbitrarily steep linear gradient in trait optimum, there is a critical steepness $B_{\text{crit}} = \sqrt{2A}$, above which a sharp range boundary forms, and the species is restricted to a small region (Barton, 2001; Polechová et al., 2009). Here, $B = b\sigma/\left(r\sqrt{2V_s}\right)$ is the loss of fitness due to moving one dispersal range, σ, along a gradient, b, in trait optimum, relative to the rate of return to equilibrium density, r, and $A = V_g/\left(2V_s r\right)$ is the loss of fitness due to genetic variance around the optimum, again relative to r.

This result assumes a constant genetic variance and neglects random drift. Gene flow across a spatial gradient itself generates genetic variance, and in a deterministic analysis, allows adaptation to an arbitrarily steep gradient (provided that variation around the optimum does not itself cause too much load; Barton, 2001). Polechová (2018) showed that in two dimensions, the limit is ultimately set by the neighbourhood size,

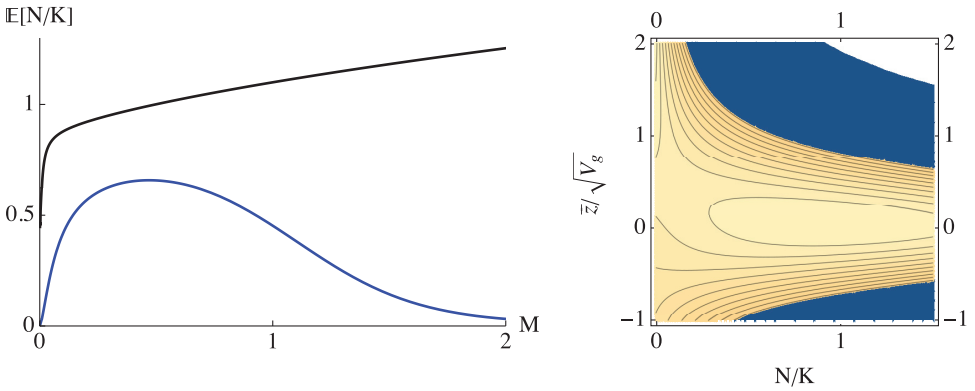

Figure 1. Left: Mean population size, $\mathbb{E}[N/K]$ as a function of the number of migrants per generation, M; the product of growth rate and carrying capacity is $r_0K = 10$. The black curve shows the neutral case, whilst the blue curve includes immigration from a source population with mean $z_s = 4\sqrt{V_g}$; the strength of stabilising selection is $V_g/\left(2\,V_s\right) = 2r_0$. Right: the joint distribution of $N/K, \bar{z}/\sqrt{V_g}$ for the selected case (blue curve at left). Contours show probability density decreasing by factors of 10; thus, the population is likely to have either a broad distribution around $N \sim K, \bar{z} \sim 0$, or near extinction with $\bar{z} \sim z_s$ (upper left).

$\mathcal{N} = 4\pi\rho\sigma^2$ (Wright, 1943), a dimensionless quantity proportional to the number of individuals in a dispersal range. A species can adapt to a continuous spatial gradient provided that genetic drift is weak enough that clines can form, approximately when $\mathcal{N} > 6.3B + 0.56$ (Polechová, 2018).

These models only consider stabilising selection towards a trait optimum that changes through space or time. Yet, organisms must maintain the global adaptations that have accumulated over their whole evolutionary history in the face of deleterious mutations. Such mutations tend to accumulate at range margins, due to small population size and range expansion, and so may contribute to population collapse (Henry et al., 2015; Peischl et al., 2015; Willi et al., 2018). These ideas can be connected by extending stabilising selection models to include a large number of traits, orthogonal to the environmental gradient, which represent global adaptations.

Reproductive isolation

These arguments allow for interaction between genes in their effect on fitness (i.e., epistasis, as with stabilising selection), but assume that alleles combine at random (i.e., linkage equilibrium). However, if migration and net selection are together strong enough to substantially reduce mean fitness, and to influence demography, then this assumption fails, even with outcrossing and free recombination. Sets of maladapted alleles introgress together, and are eliminated together. This reduces the effective migration rate, causing partial reproductive isolation, which facilitates adaptation to local conditions. Most pairs of loci are unlinked, and so linkage disequilibrium dissipates quickly. Therefore, one can define an effective migration rate as the fraction of incoming alleles that recombine onto the new genetic background (Barton & Bengtsson, 1986). In the simplest case, with no linkage and repeated backcrossing, gene flow is reduced in proportion to the product of mean fitness in successive backcross generations:

$$\frac{m_e}{m} = \left(\bar{W}_0 \bar{W}_1 \bar{W}_2 \ldots\right), \tag{7}$$

(Westram et al., 2022, Equation (2)). Sachdeva (2022) used this approximation to extend the model of Szep et al. (2021; Equation (5)) to strong (but soft) selection, and found it to be remarkably accurate in describing how the barrier to gene flow due to genome-wide selection increases the critical migration rate, below which locally adapted populations can resist swamping (Figure 2 in Sachdeva, 2022). Such genetic barriers to gene flow not only allow local adaptations to be assembled by reducing gene flow, but may also allow enough migration to maintain the genetic variance necessary for global adaptation.

Associations amongst incoming alleles cause selection on each to act on the other, reducing the effective rate of gene flow. There will be selection for various mechanisms that further reduce deleterious gene flow—that is, for reinforcement of reproductive isolation (Butlin & Smadja, 2018; Dobzhansky, 1940). Lower recombination will be favoured, to mitigate the migration load. In particular, inversions that bring together sets of alleles favourable in particular environments may establish (Charlesworth & Barton, 2018; Kirkpatrick & Barton, 2006; Kirkpatrick & Barrett, 2015). In effect, inversions form a separate gene pool for part of the genome, allowing complex adaptations to particular habitats to be assembled (analogous to the evolution of sex chromosomes). A particularly striking example is the species pair, *Drosophila pseudoobscura* and *D. persimilis*. These have overlapping ranges, and rarely hybridise. Yet, most of their differences are due to a few inversions, the rest of their genome being well-mixed (Noor et al., 2001; Korunes et al., 2021). However, one should note that inversions also link together allelic combinations that may become maladaptive under new conditions (Roesti et al., 2022): there is an advantage to recombination in changing environments.

Signatures of local adaptation

Discrete demes

At the phenotypic level, sharp clines and host races give some of the best examples of selection and incipient speciation; some have a simple genetic basis and so have served as classic examples in ecological genetics (e.g., *Biston betularia*, Cook et al., 1986; *Heliconius* (Jiggins & Lamas, 2016); inversions in *Drosophila* (Lewontin, 1981), and many more). At the genomic level, scans for excess F_{st} (Roux et al., 2016) and associations with environment (Coop et al., 2010) can reveal candidates for local adaptation. These have confirmed known loci (e.g., in *Heliconius*, Reed et al., 2011; *Antirrhinum*, Tavares et al., 2018) and located specific genes (e.g., sticklebacks, Jones et al., 2012; deer mice, Kingsley et al., 2009), but such indirect methods do not estimate selection strength, and may miss the genetic basis of most polygenic adaptation (Yeaman, 2015, 2022). Here, we briefly consider what "signature" of local adaptation we may hope to see in sampled genomes.

The short-term response to selection is captured remarkably well by the *infinitesimal model*. This assumes that traits are influenced by an enormous number of alleles of very small effect, such that the genetic variance (and covariance) evolve almost neutrally, and so can be predicted from the population structure, without complications from selection (Barton et al., 2017); thus, local adaptation would leave no trace. Over longer timescales, this extreme model cannot hold: it would imply that the genetic variance is maintained in a balance between mutation and drift. Even if this were the case for most sequence variation, neutrality seems unlikely for alleles that influence selected traits (Johnson & Barton, 2005; Sella & Barton, 2019). Moreover, distributions of allele frequency imply that deleterious mutations are strongly selected (i.e., $N_e s >> 1$, where N_e is the global effective population size; Charlesworth, 2015). Nevertheless, it may still be that variance in any particular trait is maintained as a side-effect of alleles selected for other reasons (Barton, 1990; Kondrashov & Turelli, 1992). In that case, the effects of local adaptation mediated by that trait would leave no clear trace.

Even if adaptive selection on individual alleles is strong relative to drift $(N_e s >> 1)$, and even if phenotypic clines are clear and well-behaved, the underlying allele frequencies may show complex patterns. Figure 2 shows divergence in allele frequency for the simple case of two large populations that have adapted to local optima that differ by 4.7 genetic standard deviations. The trait is influenced by alleles with a distribution of effects; alleles with small effects shift only slightly, whilst those of large effect cannot diverge, because they would increase the variance around the optimum and so reduce fitness too much. Thus, only loci with intermediate effect will diverge, and could be detected in genome scans (Figure 2, red); yet, in this example, they contribute only 28.3% of the divergence, with most divergence being due to alleles of small effect. In general, how much the different classes of loci contribute,

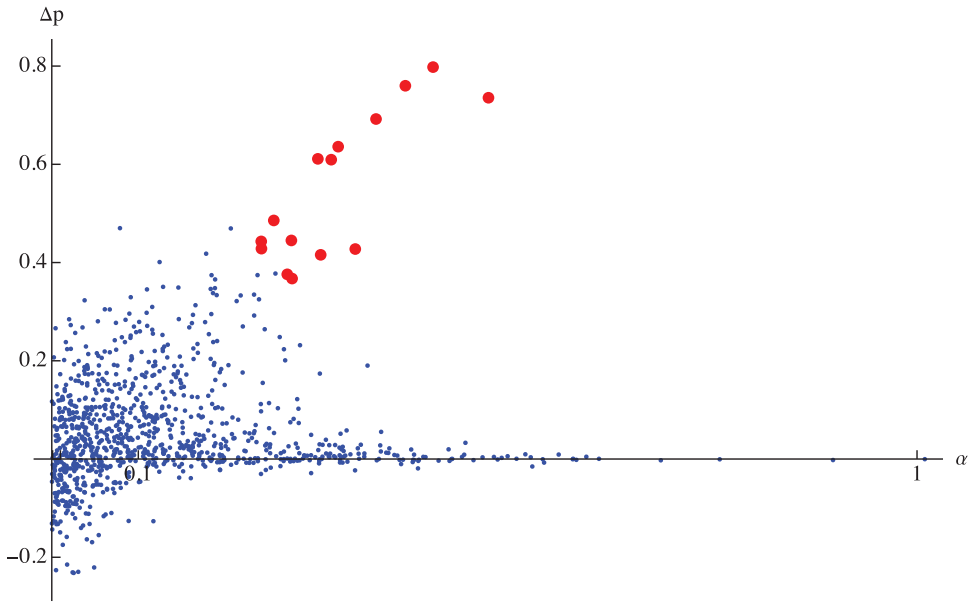

Figure 2. Divergence in allele frequency between two demes, selected for different optima ($z_{opt} = \pm 5$), plotted against the effects, α, of the $n = 1,000$ loci . The 15 loci that each contribute more than 1% of the divergence (i.e., $\alpha\Delta p > 0.1$) are highlighted by large red discs, and have effects $0.24 < \alpha < 0.50$. In total, these contribute 28.3% of the divergence, and 16.6% of the genetic variance. Eight hundred and thirty-nine smaller effect loci ($\alpha < 0.24$) contribute 58.4% of the divergence, and 62% of the genetic variance, whilst 20 larger effect loci ($\alpha > 0.50$) contribute 0.03% of the divergence, and 1.2% of the genetic variance. The remaining 11% of divergence comes from 126 loci with $0.24 < \alpha < 0.50$ but lesser divergence. Effects are drawn from an exponentially distribution with mean $4/\sqrt{n}$; the genetic standard deviation is $\sqrt{V_g} = 2.14$, and so the optima differ by $\sim 4.7\sqrt{V_g}$. However, the maximum possible range of the trait is much larger ($\pm\sum\alpha/2 \sim \pm 64.7$). Each deme contained 10^4 haploid individuals, with symmetric migration $m = 0.001$, and symmetric mutation $\mu = 2.5 \times 10^{-5}$. Stabilising selection had strength $V_s = 20$. For efficiency, the simulation follows allele frequencies, and so neglects linkage disequilibria. It was run for 5,000 generations with $z_{opt} = 0$ in both demes, and then for a further 5,000 generations with $z_{opt} = \pm 5$. The mean equilibrated to within 0.15 of the optimum by ∼100 generations, whilst the genetic variance within each deme increased from ∼0.73 (before divergence) to ∼ 1.3 after divergence.

either to genetic variance or to divergence, will depend in a complex way on the distribution of effects and on the size of the shift. Patterns would be yet more obscure in smaller populations, where drift is stronger.

Continuous space

In continuous space, species can in principle adapt to an environmental gradient via clines at the underlying loci. However, even if selection is strong enough on each allele that random drift can be ignored, the allelic clines will not follow a gradual phenotypic cline. Instead, there will be a series of sharp clines, which only in aggregate generate a smooth gradient in phenotype (Barton, 1999; Figure 3). Under stabilising selection, each allele experiences selection to reduce the variance around the optimum, which is equivalent to selection against heterozygotes $s = \alpha^2/(2V_s)$, where α is the allelic effect. Therefore, each cline has width $\sim \sigma/\sqrt{2s}$, which may be much narrower than the phenotypic gradient. If alleles affect multiple traits (as is inevitable), responding to multiple environmental heterogeneities, then their spatial pattern may have no clear interpretation (Lotterhos, 2023). Moreover, even if selection is stronger than drift, there can still be considerable random fluctuation in allele frequency. Selection strongly constrains the trait mean to follow its optimum, but that can be achieved by summing a variety of obscure underlying patterns, each of which fluctuates unpredictably (Figure 3). Moreover, clines for environment-independent incompatibilities can become coupled to clines due to local adaptation, obscuring the relation between these different processes (Bierne et al., 2011).

Discussion

A species must maintain the complex adaptations that enable it to survive and reproduce across some range of environments. We have seen that accumulation of the information required for such global adaptations is fundamentally limited by total population size, and may be most efficient when very many alleles of small effect are brought together by recombination. The abundance of genetic diversity (Lewontin, 1974), and the almost universal prevalence of sexual reproduction (at least across eukaryotes) also indicate the crucial role of polygenic adaptation. Yet, species encounter environments that change continually, in space and time, not least through the evolution of competing species, and must to some degree rely on transient and local adaptations to maintain their range. Although simple models of single loci and discrete demes suggest that such local adaptation is sensitive to gene flow, local adaptation can nevertheless evolve when it is based on subtle shifts in allele frequency, across a spatially extensive habitat. Moreover, once local adaptations are sufficiently strong, they confer partial reproductive isolation, often facilitated by chromosomal inversions.

These arguments suggest that the boundaries of biological species should be drawn more widely than is usual. The frequent observation of extensive gene flow between distinct "species" suggests that the biological species may be a much larger entity than we usually conceive, allowing local and

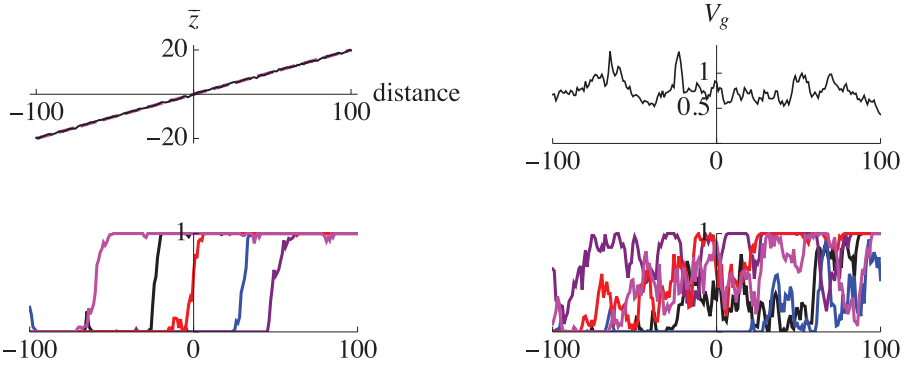

Figure 3. Clines in a polygenic trait, and underlying allele frequencies. Top left: Trait mean (blue), which almost coincides with the optimum (red); top right: genetic variance; bottom left: allele frequencies at the five largest effect loci (γ = 0.71 to 1.42); bottom right: at the five smallest effect loci (γ ∼0.20). In each case, only four of the five loci diverge, and so show visible clines. The simulation began with uniform allele frequencies, and ran for 1,000 generations; the mean reached the optimum within 100 generations, whilst the variance settled within ∼300 generations. There are 201 demes (labelled –100 to 100), each with 100 haploid individuals, and with migration 0.05 to each of the 2 neighbours. Other parameters as in Figure 2, except that only the 200 largest effect loci were followed, to save computation.

global adaptation to be to some degree reconciled (see Barraclough, 2024): a biological species can draw on a wide pool of adaptive variation, even if there is substantial local adaptation and reproductive isolation. This view spans a broad taxonomic, temporal and spatial range. Practical concern over conservation of endangered populations are necessarily short term; nevertheless, we should bear in mind that whilst isolated populations of a thousand or so individuals may survive immediate threats (Frankham et al., 2014; Lande & Barrowclough, 2010), in the longer term much larger populations, and wider gene exchange, may be essential (see Parmesan et al., 2023).

These arguments suggest that the boundaries of biological species should be drawn more widely than is usual. The frequent observation of extensive gene flow between distinct "species" suggests that the biological species may be a much larger entity than we usually conceive, and a biological species can draw on a wide pool of adaptive variation, even if there is substantial local adaptation and reproductive isolation (see Barraclough, 2024). Consequently, local adaptation may become easier when the species' boundary is blurred.

Over recent years, there has been a profusion of methods that attempt to infer adaptation from sampled sequences; these include estimates of overall adaptation (inbreeding load; e.g., Willi et al., 2018) and local adaptation (from associations between SNP and location, or SNP and environment; Roux et al., 2016; Coop et al., 2010). Such methods aim for practical application in conservation (e.g., by assisted migration; Aitken & Whitlock, 2013). However, the theoretical considerations reviewed here suggest that it will be difficult to relate genomic data to environmental heterogeneity, especially if we seek quantitative estimates of selection strength and the distribution of allelic effects; we may miss much polygenic adaptation altogether. It will be crucial to test inferences from genetic data against those few systems where we know the actual genetic basis of local adaptation.

How can we better understand the limits to a species' range? The broad argument has been that we need to see this question in a wider context, as part of understanding how selection generates and maintains complex adaptations. Two specific questions seem especially important. First, very few sequence-based methods estimate the strength of selection, and so although we have found many "signals of selection", there have been few attempts to estimate the net strength of selection, or its distribution across loci. The dependence of sequence diversity on recombination rate and substitutions can give such estimates (e.g., Elyashiv et al., 2016), and can be extended to include divergence between populations (Aeschbacher et al., 2017); another promising route uses correlations in allele frequency fluctuations to estimate the heritable variance in fitness (Buffalo & Coop, 2020; Robertson, 1961). Second, we have very limited evidence as to the effects of a population's genotypic composition on its abundance (Sexton et al., 2009), and still less on how species interactions modulate this (Louthan et al., 2015). Here, it seems unlikely that indirect sequence-based methods can be informative: a better understanding will require sustained field experiments.

Acknowledgements

This work was supported by a grant from the ERC, 101055327, "HaplotypeStructure". I thank Himani Sachdeva, Michal Hledik, Jitka Polechova, and the reviewers for their helpful comments.

Conflicts of interest

None declared.

Data availability

All code used to generate the figures is in the accompanying *Mathematica* notebook.

Appendix

Joint distribution of N, $\bar{z}$

Suppose that a trait with mean $\bar{z}$ and genetic variance V_g is subject to stabilising selection towards an optimum at $z_{\rm opt}$; the population growth rate is $r_0 - \lambda N - \left(\bar{z} - z_{\rm opt}\right)^2/2V_s$, where V_s is the strength of stabilising selection. M migrants enter per generation, from a source with trait mean $\bar{z}_s$. The genetic

variance is assumed to be constant, and the trait distribution Gaussian; this is an approximation, since even under the infinitesimal model, the variance will be dissipated by inbreeding and changed by migration. If the mean were at the optimum, and demographic fluctuations negligible, then the equilibrium population size would be $K = r_0/\lambda$. Then:

$$\Delta\bar{z} = \frac{M}{N}\left(z_s - \bar{z}\right) - \frac{V_g}{V_s}\left(\bar{z} - z_{\text{opt}}\right) + \zeta_z$$

$$\Delta \text{N} = M + N\left(r_0 - \lambda N - \frac{\left(\bar{z} - z_{\text{opt}}\right)^2}{2V_s}\right) + \zeta_N. \qquad (8)$$

where demographic fluctuations ζ_N have variance N, and fluctuations in the mean have variance V_g/N. The joint stationary distribution is then:

$$\psi\left[N, \bar{z}\right] \sim N^{2M-1} \times \exp\left[-\frac{M}{V_g}\left(z_s - \bar{z}\right)^2 - \frac{N}{V_s}\left(z_{\text{opt}} - \bar{z}\right)^2 - \frac{r_0}{2K}(N-K)^2\right]. \qquad (9)$$

Integrating over z, we find the marginal distribution of N:

$$\psi[N] \sim \frac{N^{2M-1}}{\sqrt{M + 2Nr_0\mathcal{L}}} \exp\left[-\frac{1}{2}\frac{r_0}{K}\left((N-K)^2 + \frac{4MNK\, z_s^2\mathcal{L}}{M + 2Nr_0\mathcal{L}}\right)\right]$$

$$\text{where } \mathcal{L} = \frac{V_g}{2r_0V_s}. \qquad (10)$$

We can ask where the least and most likely value of N lies for given z, and conversely for $\bar{z}$:

$$z_{\max} = \frac{Mz_s}{M + 2Nr_0\mathcal{L}} N_{\min/\max}$$

$$= \frac{1}{2} - z^2\mathcal{L} \pm \sqrt{\frac{(2M-1)}{r_0K} + \frac{1}{4}\left(1 - 2z^2\mathcal{L}\right)^2}. \qquad (11)$$

References

Aeschbacher, S., Selby, J. P., Willis, J. H., & Coop, G. (2017). Population-genomic inference of the strength and timing of selection against gene flow. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, 7061–7066.

Aitken, S. N., & Whitlock, M. C. (2013). Assisted gene flow to facilitate local adaptation to climate change. *Annual Review of Ecology, Evolution, and Systematics*, *44*, 367–388.

Antonovics, J. (1976). The nature of limits to natural selection. *Annals Missouri Botanical Garden*, *63*, 224–247.

Barraclough, T. (2024). Does selection favour the maintenance of porous species boundaries? *Journal of Evolutionary Biology*, *37*(6), 616–627.

Barton, N. H. (1990). Pleiotropic models of quantitative variation. *Genetics*, *124*, 773–782.

Barton, N. H. (1999). Clines in polygenic traits. *Genetics Research*, *74*, 223–236.

Barton, N. H. (2001). Adaptation at the edge of a species' range. In J. Silvertown, & J. Antonovics (Eds.), *Integrating ecology and evolution in a spatial context* (pp. 365–392). Blackwell.

Barton, N. H. (2017). How does epistasis influence the response to selection?. *Heredity*, *118*, 96–109.

Barton, N. H. (2022). The "new synthesis". *Proceedings of the National Academy of Sciences of the United States of America*, *119*, e2122147119.

Barton, N. H., & Bengtsson, B. O. (1986). The barrier to genetic exchange between hybridising populations. *Heredity*, *57*, 357–376.

Barton, N. H., Etheridge, A. M., & Véber, A. (2017). The infinitesimal model: Definition, derivation, and implications. *Theoretical Population Biology*, *118*, 50–73.

Bierne, N., Welch, J., Loire, E., Bonhomme, F., & David, P. (2011). The coupling hypothesis: Why genome scans may fail to map local adaptation genes. *Molecular Ecology*, *20*, 2044–2072.

Bod'ová, K., Tkačik, G., & Barton, N. H. (2016). A general approximation for the dynamics of quantitative traits. *Genetics*, *202*, 1523–1548.

Bolnick, D. I., & Nosil, P. (2007). Natural selection in populations subject to a migration load. *Evolution*, *61*, 2229–2243.

Buffalo, V., & Coop, G. (2020). Estimating the genome-wide contribution of selection to temporal allele frequency change. *Proceedings of the National Academy of Sciences of the United States of America*, *117*, 20672–20680.

Butlin, R. K., & Smadja, C. M. (2018). Coupling, reinforcement, and speciation. *The American Naturalist*, *191*, 155–172.

Caisse, M., & Antonovics, J. (1978). Evolution in closely adjacent plant populations. *Heredity*, *40*, 371–384.

Charlesworth, B. (2013). Stabilizing selection, purifying selection, and mutational bias in finite populations. *Genetics*, *194*, 955–971.

Charlesworth, B. (2015). Causes of natural variation in fitness: Evidence from studies of *Drosophila* populations. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 1662–1669.

Charlesworth, B., & Barton, N. H. (2018). The spread of an inversion with migration and selection. *Genetics*, *208*, 377–382.

Cook, L. M., Mani, G. S., & Varley, M. E. (1986). Postindustrial melanism in the peppered moth. *Science*, *231*, 611–613.

Coop, G., Witonsky, D., Di Rienzo, A., & Pritchard, J. K. (2010). Using environmental correlations to identify loci underlying local adaptation. *Genetics*, *185*, 1411–1423.

Crow, J. F. (1958). Some possibilities for measuring selection intensities in man. *Human Biology*, *30*, 1–13.

Crow, J. F. (2001). Shannon's brief foray into genetics. *Genetics*, *159*, 915–917.

Débarre, F., Ronce, O., & Gandon, S. (2013). Quantifying the effects of migration and mutation on adaptation and demography in spatially heterogeneous environments. *Journal of Evolutionary Biology*, *26*, 1185–1202.

Dobzhansky, T. (1940). Speciation as a stage in evolutionary divergence. *The American Naturalist*, *74*, 312–321.

Elyashiv, E., Sattath, S., Hu, T. T., Strutsovsky, A., McVicker, G., Andolfatto, P., Coop, G., & Sella, G. (2016). A genomic map of the effects of linked selection in *Drosophila*. *PLoS Genetics*, *12*, e1006130.

Felsenstein, J. (1971). On the biological significance of the cost of gene substitution. *The American Naturalist*, *105*, 1–11.

Fisher, R. A. (1930). *The genetical theory of natural selection*. Oxford University Press.

Fisher, R. A. (1937). The wave of advance of advantageous genes. *Annals of Eugenics*, *7*, 355–369.

Frankham, R., Bradshaw, C. J., & Brook, B. W. (2014). Genetics in conservation management: Revised recommendations for the 50/500 rules, Red List criteria and population viability analyses. *Biological Conservation*, *170*, 56–63.

Gomulkiewicz, R., Holt, R. D., & Barfield, M. (1999). The effects of density dependence and immigration on local adaptation and niche evolution in a black-hole sink environment. *Theoretical Population Biology*, *55*, 283–296.

Haldane J. B. S. (1937). The effect of variation on fitness. *The American Naturalist*, *71*, 337–349.

Haldane, J. B. S. (1948). The theory of a cline. *Journal of Genetics*, *48*, 277–284.

Haldane, J. B. S. (1956). The relation between density regulation and natural selection. *Proceedings of the Royal Society B: Biological Sciences*, *145*, 306–308.

Haldane, J. B. S. (1957). The cost of natural selection. *Journal of Genetics*, *55*, 511–524.

Hendry, A. P., Day, T., & Taylor, E. B. (2001). Population mixing and the adaptive divergence of quantitative traits in discrete populations: A theoretical framework for empirical tests. *Evolution*, *55*, 459–466.

Henry, R. C., Bartoń, K. A., & Travis, J. M. (2015). Mutation accumulation and the formation of range limits. *Biology Letters*, *11*. https://doi.org/10.1098/rsbl.2014.0871

Hledík, M., Barton, N., & Tkačik, G. (2022). Accumulation and maintenance of information in evolution. *Proceedings of the Royal Society B: Biological Sciences*, *119*, e2123152119.

Holt, R. D., Gomulkiewicz, R., & Barfield, M. (2003). The phenomenology of niche evolution via quantitative traits in a "black-hole" sink. *Proceedings of the Royal Society B: Biological Sciences*, *270*, 215–224. https://doi.org/10.1098/rspb.2002.2219

Holt, R. D., & Gomulkiewicz, R. (1997). How does immigration influence local adaptation? A re-examination of a familiar paradigm. *The American Naturalist*, *149*, 563–572.

Iwasa, Y. (1988). Free fitness that always increases in evolution. *Journal of Theoretical Biology*, *135*, 265–281.

Jiggins, C. D., & Lamas, G. (2016). *The ecology and evolution of Heliconius butterflies*. Oxford University Press.

Johannesson, K., Rolán-Alvarez, E., & Ekendahl, A. (1995). Incipient reproductive isolation between two sympatric morphs of the intertidal snail *Littorina saxatilis*. *Evolution*, *49*, 1180–1190.

Johnson, T., & Barton, N. (2005). Theoretical models of selection and mutation on quantitative traits. *Philosophical Transactions of the Royal Society B*, *360*, 1411–1425.

Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M. C., White, S., & Birney, E. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, *484*, 55–61.

Kawecki, T. J., Barton, N. H., & Fry, J. D. (1997). Mutational collapse of fitness in marginal habitats and the evolution of ecological specialisation. *Journal of Evolutionary Biology*, *10*, 407–430.

Kimura, M. (1961). Natural selection as the process of accumulating genetic information in adaptive evolution. *Genetics Research*, *2*, 127–140.

Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge University Press.

Kingsley, E. P., Manceau, M., Wiley, C. D., & Hoekstra, H. E. (2009). Melanism in *Peromyscus* is caused by independent mutations in *Agouti*. PloS One, *4*, e6435.

Kirkpatrick, M., & Barrett, B. (2015). Chromosome inversions, adaptive cassettes and the evolution of species' ranges. *Molecular Ecology*, *24*, 2046–2055. https://doi.org/10.1111/mec.13074

Kirkpatrick, M., & Barton, N. H. (1997). Evolution of a species' range. *The American Naturalist*, *150*, 1–23.

Kirkpatrick, M., & Barton, N. H. (2006). Chromosome inversions, local adaptation, and speciation. *Genetics*, *173*, 419–434.

Kondrashov, A. S. (1988). Deleterious mutations and the evolution of sexual reproduction. *Nature*, *336*, 435–440.

Kondrashov, A. S., & Turelli, M. (1992). Deleterious mutations, apparent stabilizing selection and the maintenance of quantitative variation. *Genetics*, *132*, 603–618.

Korunes, K. L., Machado, C. A., & Noor, M. A. (2021). Inversions shape the divergence of *Drosophila pseudoobscura* and *Drosophila persimilis* on multiple timescales. *Evolution*, *75*, 1820–1834.

Lande, R. (1976). Natural selection and random genetic drift in phenotypic evolution. *Evolution*, *30*, 314–334.

Lande, R., & Barrowclough, G. F. (2010). Effective population size, genetic variation, and their use in population management. In M. E. Soulé (Ed.), *Viable populations for conservation* (pp. 87–124). Cambridge University Press.

Laroche, F., & Lenormand, T. (2023). The genetic architecture of local adaptation in a cline. *Peer Community Journal*, *3*, e20. https://doi.org/10.24072/pcjournal.245.

Lewontin, R. C. (1974). *The genetic basis of evolutionary change*. Columbia University Press.

Lewontin, R. C. (1981). The scientific work of T. Dobzhansky. In R. C. Lewontin, J. A. Moore, W. B. Provine, & B. Wallace (Eds.), *Dobzhansky's genetics of natural populations I-XLIII*. Columbia University Press.

Lynch, M. (2007). *The origins of genome architecture*. Sinauer Associates.

Lynch, M., Conery, J., & Burger, R. (1995). Mutation accumulation and the extinction of small populations. *The American Naturalist*, *146*, 489–518.

Maruyama, T. (1974). A simple proof that certain quantities are independent of the geographical structure of population. *Theoretical Population Biology* , *5*, 148–154.

Lotterhos, K. E. (2023). The paradox of adaptive trait clines with nonclinal patterns in the underlying genes. *Proceedings of the National Academy of Sciences of the United States of America*, *120*, e2220313120.

Louthan, A. M., Doak, D. F., & Angert, A. L. (2015). Where and when do species interactions set range limits? *Trends in Ecology & Evolution*, *30*, 780–792. https://doi.org/10.1016/j.tree.2015.09.011

Nagylaki, T. (1975). Conditions for the existence of clines. *Genetics*, *80*, 595–615.

Nagylaki, T. (1982). Geographical invariance in population genetics. *Journal of Theoretical Biology*, *99*, 159–172.

Noor, M. A. F., K. L. Grams, L. A. Bertucci, & J. Reiland. (2001). Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences of the United States of America*, *98*, 12084–12088.

Paixão, T., & Barton, N. H. (2016). The effect of gene interactions on the long-term response to selection. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 4422–4427.

Parmesan, C., Singer, M. C., Wee, B., & Mikheyev, S. (2023). The case for prioritizing ecology/behavior and hybridization over genomics/taxonomy and species' integrity in conservation under climate change. *Biological Conservation*, *281*, 109967.

Partridge, L., & Barton, N. H. (1993). Optimality, mutation and the evolution of ageing. *Nature*, *362*, 305–311.

Pease, C. M., Lande, R., & Bull, J. J. (1989). A model of population growth, dispersal and evolution in a changing environment. *Ecology*, *70*, 1657–1664.

Peischl, S., Kirkpatrick, M., & Excoffier, L. (2015). Expansion load and the evolutionary dynamics of a species range. *The American Naturalist*, *185*, E81–E93.

Polechová, J. (2018). Is the sky the limit? On the expansion threshold of a species' range. *PLoS Biology*, *16*, e2005372.

Polechová, J., N. H. Barton, & G. Marion. (2009). Species' range: Adaptation in space and time. *The American Naturalist*, *174*, E186–E204.

Reed, R. D., Papa, R., Martin, A., Hines, H. M., Counterman, B. A., Pardo-Diaz, C., Jiggins, C. D., Chamberlain, N. L., Kronforst, M. R., Chen, R., & Halder, G. (2011). *Optix* drives the repeated convergent evolution of butterfly wing pattern mimicry. *Science*, *333*, 1137–1141.

Robertson, A. (1960). A theory of limits in artificial selection. *Proceedings of the Royal Society of London*, *153*, 234-249.

Robertson, A. (1961). Inbreeding in artificial selection programmes. *Genetics Research*, *2*, 189–194.

Roesti, M., Gilbert, K. J., & Samuk, K. (2022). Chromosomal inversions can limit adaptation to new environments. *Molecular Ecology*, *31*, 4435–4439. https://doi.org/10.1111/mec.16609

Ronce, O., & Kirkpatrick, M. (2001). When sources become sinks: Migrational meltdown in heterogeneous habitats. *Evolution*, *55*, 1520–1531.

Roughgarden, J. (1979). *Theory of population genetics and evolutionary ecology: An introduction*. Macmillan.

Roux, C., Fraisse, C., Romiguier, J., Anciaux, Y., Galtier, N., & Bierne, N. (2016). Shedding light on the grey zone of speciation along a continuum of genomic divergence.*PLoS Biology*, *14*, e20 00234.

Sachdeva, H. (2019). Effect of partial selfing and polygenic selection on establishment in a new habitat. *Evolution*, *73*, 1729–1745.

Sachdeva, H. (2022). Reproductive isolation via polygenic local adaptation in sub-divided populations: Effect of linkage disequilibria and drift. *PLoS Genetics 18*, e1010297.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*, 379–423.

Sella, G., & Barton, N. H. (2019). Thinking about the evolution of complex traits in the era of genome-wide association studies. *Annual Review of Genomics and Human Genetics*, *20*, 461–493.

Sexton, J. P., McIntyre, J., Angert, A. L., & Rice, K. J. (2009). Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 415–436.

Slatkin, M. (1973). Gene flow and selection in a cline. *Genetics*, *75*, 733–756.

Szep, E., Sachdeva, S., & Barton, N. H. (2021). Polygenic local adaptation in metapopulations: A stochastic eco-evolutionary model. *Evolution* , *75*, 1030-1045.

Tavares, H., Whibley, A., Field, D. L., Bradley, D., Couchman, M., Copsey, L., Elleouet, J., Burrus, M., Andalo, C., Li, M., & Li, Q. (2018). Selection and gene flow shape genomic islands that control floral guides. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 11006–11011.

Tufto, J. (2000). Quantitative genetic models for the balance between migration and stabilizing selection. *Genetics Research*, *76*, 285–293.

Tufto, J. (2001). Effects of releasing maladapted individuals: A demographic-evolutionary model. *The American Naturalist*, *158*, 331–340.

Uecker, H., Otto, S. P., Hermisson, J. (2014). Evolutionary rescue in structured populations. *The American Naturalist*, *183*, E17–E35.

Yeaman, S. (2015). Local adaptation by alleles of small effect. *The American Naturalist*, *186*, S74–S89.

Yeaman, S. (2022). Evolution of polygenic traits under global vs local adaptation. *Genetics*, iyab134.

Yeaman, S., & Otto, S. P. (2011). Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift. *Evolution*, *65*, 2123–2129.

Westram, A. M., Stankowski, S., Surendranadh, P., & Barton, N. (2022). What is reproductive isolation? *Journal of Evolutionary Biology*, *35*, 1143–1164.

Whitlock, M. C. (2002). Selection, load and inbreeding depression in a large metapopulation. *Genetics*, *160*, 1191–1202.

Wilke, C. O., Wang, J. L., Ofria, C., Lenski, R. E., & Adami, C. (2001). Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature*, *412*, 331–333.

Willi, Y., Fracassetti, M., Zoller, S., & van Buskirk, J. (2018). Accumulation of mutational load at the edges of a species range. *Molecular Biology and Evolution*, 35, 781–791.

Wright, S. (1937). The distribution of gene frequencies in populations. *Proceedings of the National Academy of Sciences of the United States of America*, 23, 307–320.

Wright, S. (1943). Isolation by distance. *Genetics*, 28, 114.