Is the sky the limit? On the expansion threshold of a species' range

Jitka Polechová[1,2*]

***For correspondence:** jitka@univie.ac.at (JP)

[1]University of Vienna, Department of Mathematics; [2]Institute for Science and Technology (IST Austria), Klosterneuburg, Austria

Abstract Gene flow across environments has conflicting effects: while it increases the genetic variation necessary for adaptation and counters the loss of genetic diversity due to genetic drift, it may also swamp adaptation to local conditions. This interplay is crucial for the dynamics of a species' range expansion. I show that it can be understood based on two dimensionless parameters: i) the fitness cost of dispersal – a measure of environmental heterogeneity – and ii) the strength of genetic drift – a measure of reduction of genetic diversity. Together, these two parameters define an *expansion threshold*: adaptation fails when the the population's neighbourhood size is so small that genetic drift reduces diversity below the level requisite for adaptation to the environmental heterogeneity. This threshold provides a novel, theoretically justified and testable prediction for formation of a range margin and a collapse of a species' range in two-dimensional habitats.

Introduction

Species' borders are not just determined by the limits of their ecological niche (***Griggs 1914***; ***Hargreaves et al. 2013***). A species' edge is typically sharper than would be implied by the environment (both biotic and abiotic). Moreover, although species' ranges are inherently dynamic, it is puzzling that they typically expand rather slowly. The usual – but tautological – explanation is that lack of genetic variation at the range margin prevents further expansion (***Antonovics 1976***). Indeed, a species' range edge is often associated with lower genetic variation (***Gaston 2003***; ***Kawecki 2008***; ***Eckert et al. 2008***; ***Sexton et al. 2009***; ***Cahill and Levinton 2016***; ***Takahashi et al. 2016***). Yet, why does selection for new variants near the edge of the range not increase adaptive genetic variance, thereby enabling it to continuously expand (***Antonovics 1976***; ***Barton 2001***)? ***Haldane*** (***1957***) proposed a general explanation: even if environmental conditions vary smoothly, "swamping" by gene flow from central to marginal habitats will cause more severe maladaptation in marginal habitats, further reducing their population density. This would lead to a sharp edge to a species' range, even if genetic variance at the range margin is large. However, the consequences of dispersal and gene flow for evolution of a species' range continue to be debated (***Nosil and Crespi 2004***; ***Sexton et al. 2011***): a number of studies suggest that intermediate dispersal may be optimal (***Gomulkiewicz et al. 1999***; ***Holt and Gomulkiewicz 1997***; ***Alleaume-Benharira et al. 2006***; ***Uecker et al. 2014***). Gene flow across heterogeneous environments can be beneficial, because the increase of genetic variance allows the population to adapt in response to selection (***Barton 2001***). Current theory identifies that local population dynamics, dispersal, and evolution of niche-limiting traits (including their variance), and both genetic and demographic stochasticity are all important for species' range

dynamics (***Kirkpatrick and Barton 1997***; ***Gomulkiewicz et al. 1999***; ***Holt and Gomulkiewicz 1997***; ***Barton 2001***; ***Alleaume-Benharira et al. 2006***; ***Chevin and Lande 2010***; ***Polechová and Barton 2015***; ***Gilbert and Whitlock 2017***). Yet, these core aspects have not been incorporated into a single study that would provide testable predictions for range limits in two-dimensional habitats.

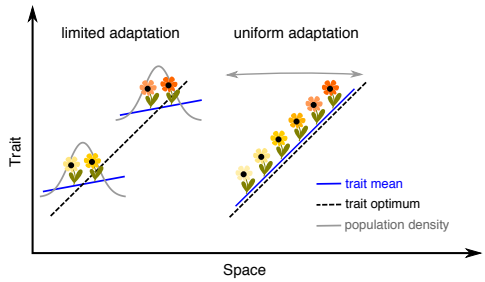


Figure 1. Two modes of adaptation. Assuming that genetic variance is fixed, deterministic theory ***Kirkpatrick and Barton*** (***1997***) predicts that there are two modes of adaptation to an environmental gradient. When the effective environmental gradient B is steep relative to the standing load due to genetic variance A, clinal adaptation fails, and the population only matches the optimum at the very centre of its range (*limited adaptation*). Conversely, when the standing load is large relative to the effective environmental gradient ($A > B^2/2$), a population adapts continuously, gradually expanding its range (*uniform adaptation*). Black dashed lines depict the trait optimum, blue lines depict the trait mean. Population density is shown in gray: it has a sharp margin for *limited adaptation* but is steadily expanding under *uniform adaptation*.

As ***Haldane*** (***1957***) previously pointed out, it is important to consider population- and evolutionary dynamics across a species' range jointly, as their effects interact. Due to maladaptation, both the carrying capacity of the habitat and the population growth rate is likely to decrease – such selection is called *hard* (***Christiansen 1975***). Classic deterministic theory (***Kirkpatrick and Barton 1997***) shows that when genetic variance is fixed, there are two stable regimes of adaptation to a spatially varying optimum (see ***Figure 1***): i) a *limited adaptation*, where a population is only adapted to a single optimum or becomes a patchy conglomerate of discrete phenotypes or ii) continuous or *uniform* adaptation, which is stable when the genetic variance, measured in terms of its cost in fitness (*segregation load*) is large relative to the cost of dispersal across environments (*dispersal load*). Under *uniform adaptation*, a species' range gradually expands – a stable boundary only forms when the genetic variance is too small to allow continuous adaptation to the spatially variable environment and hence the *limited adaptation* is stable. However, when genetic variance can evolve, such a limit no longer exists in infinitely large populations: continuous adaptation is always possible until the combined genetic and dispersal loads are so large that the population density drops to zero (***Barton 2001***). Deterministic theory thus predicts that a sharp and stable boundary to a species' range does not form when the environment changes smoothly. Yet, in finite populations, genetic drift reduces local genetic variance (***Wright 1931***), potentially qualitatively changing the dynamics. Indeed, it has been shown that for one-dimensional habitats (such as rivers), a sharp range margin arises when genetic drift relative to selection per locus becomes strong relative to the fitness cost of dispersal across environments (***Polechová and Barton 2015***). However, most species live in two-dimensional habitats. There, the mixing due to dispersal is more efficient and the effect of genetic drift changes qualitatively, becoming only weakly dependent on selection (***Barton et al. 2002***). Is there still an intrinsic threshold to range expansion in finite populations when dispersal and gene flow occur over two-dimensional space, rather than along a line? If so, what is its biological interpretation?

Results

I study the problem of intrinsic limits to adaptation in a two dimensional habitat. Throughout, I assume that the species' niche is limited by stabilising selection on a composite phenotypic trait. This optimum varies across one dimension of the two-dimensional habitat – such as temperature and humidity with altitude. Demography and evolution are considered together. Selection is *hard*: both the rate of density-dependent population growth and the attainable equilibrium density

decrease with increasing maladaptation. Both trait mean and genetic variance can freely evolve *via* change in allele frequencies. The populations are finite and both genetic and demographic stochasticity are included. The model is first outlined at a population level, in terms of coupled stochastic differential equations. While it is not possible to obtain analytical solutions to this model, this formalisation allows us to identify the effective dimensionless parameters which describe the dynamics. Next, individual based simulations are used to determine the driving relationship between the key parameters, and test its robustness. The details are described in ***Methods: Model***.

The dynamics of the evolution of a species' range, as formalised by this model, are well described by three dimensionless parameters, which give a full description of the system. Details are given in the ***Methods: Rescaling***. The first dimensionless parameter carries over from the phenotypic model of ***Kirkpatrick and Barton*** (***1997***): the effective environmental gradient B measures the fitness cost of maladaptation incurred by dispersal across a heterogenous environment. The second parameter is the neighbourhood size of the population, $\mathcal{N}$, which can be understood as the number of diploid individuals within one generation's dispersal range. Originally, neighbourhood size was defined by ***Wright*** (***1946***) as the size of the single panmictic diploid population which would give the same probability of identity by descent in the previous generation. The inverse of neighbourhood size $1/\mathcal{N}$ hence describes the local increase of homozygosity due to genetic drift. The third dimensionless parameter is the ratio s/r^* of the strength of selection s per locus relative to the strength of density dependence, r^*. Detailed description of the parameters can be found in ***Methods: Table 1***.

In order to see how the rescaled parameters capture the evolution of a species' range, I simulated evolving populations adapting to a linear gradient in the optimum. Depending on the parameters, the population either expands, gradually extending its phenotypic range by consecutive sweeps of loci advantageous at the edges or the species' range contracts or disintegrates as adaptation fails. ***Figure 2*** shows the range dynamics of 750 evolved populations, each based on different parametrizations, in relation to the first two compound dimensionless parameters, i) the effective environmental gradient B and ii) the inverse of neighbourhood size $1/\mathcal{N}$ which describes the effect of genetic drift on the allele frequencies. The individual-based model uses unscaled (naïve) model parameters and includes mutation (see ***Methods: Individual-based Model*** and ***Table 1***).

Demographic stochasticity alone is sufficient to drive the species' range to collapse where its local neighbourhood size $\mathcal{N}$ is smaller than about one. For larger neighbourhood sizes, the two parameters B and $\mathcal{N}$ give a clear separation between expanding populations when the neighbourhood size $\mathcal{N}$ is large relative to the effective environmental gradient B (shown in blue, ***Figure 2***) and the rest, where adaptation is failing. The separation gives an *expansion threshold*, estimated at $\mathcal{N} \approx 6.3\, B + 0.56$ (red line). Above the *expansion threshold*, populations are predicted to expand (see ***Figure 3***); below it, adaptation fails abruptly. If conditions change uniformly across space (as in these simulation runs, with constant gradient and carrying capacity), this means that adaptation fails everywhere – a species' range then either collapses from the margins (***Figure 2***, red hues) and/or disintegrates (***Figure 2***, open circles), forming a fragmented metapopulation. The subpopulations forming this metapopulation have only a very narrow phenotypic range and maintain locally only minimal adaptive variance. They correspond to the *limited adaptation* regime identified for a phenotypic model with genetic variance as a parameter (***Kirkpatrick and Barton 1997***). In contrast to one-dimensional habitats (***Polechová and Barton 2015***), these patchy metapopulations are stabilised by dispersal from surrounding subpopulations in the two-dimensional habitat and can thus persist for a long time. An example of such a metapopulation is given in ***Figure 4***.

Interestingly, the third dimensionless parameter s/r^* has no detectable effect. In other words, whilst the *expansion threshold* reflects the total fitness cost of dispersal in a heterogeneous environment, it appears independent of the strength of selection per locus (s). This is both because

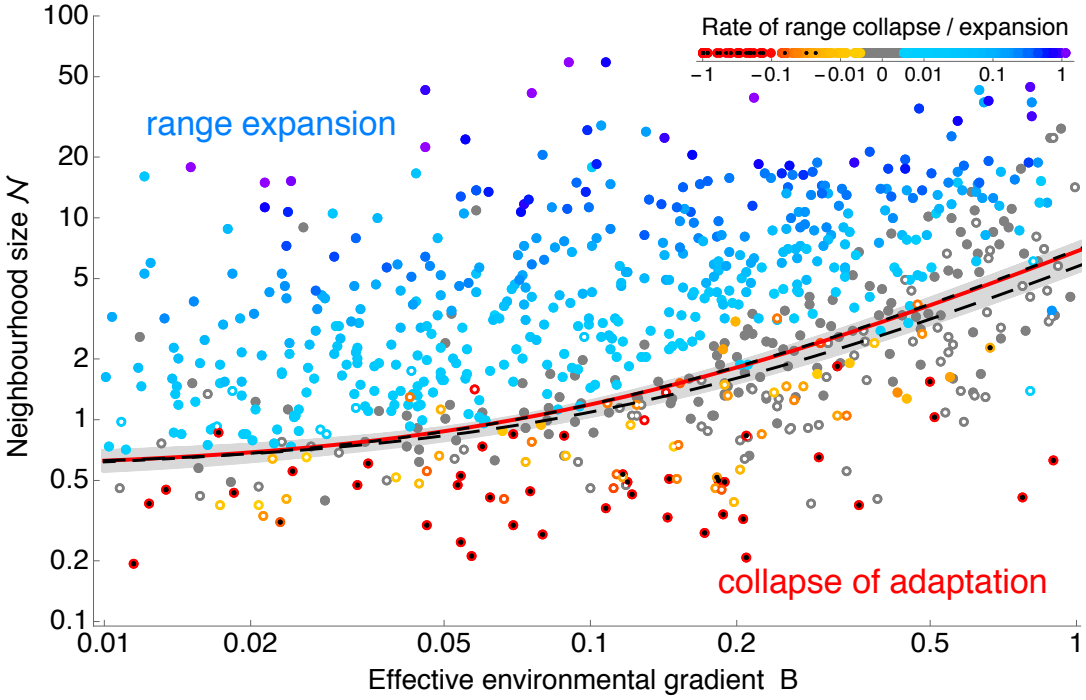


Figure 2. Two dimensionless parameters, the neighbourhood size $\mathcal{N}$ and the effective environmental gradient B give a clear prediction whether a species' range can expand (blue hues). The red line shows the fitted boundary between expanding populations (in blue) and collapsing ranges (red hues): populations expand above the *expansion threshold*, when $\mathcal{N} \gtrapprox 6.3\, B + 0.56$. The grey region gives 95% bootstrap confidence intervals, whilst the dashed lines depict the predicted *expansion threshold* for weak selection ($s/r^* < 0.005$, indistinguishable from the whole data set) and for strong selection $s/r^* > 0.005$ (lower dashed line). Stagnant populations, changing by less than five demes per 1000 generations, are shown in grey. Open circles denote populations where continuous adaptation has collapsed and the population consists of many discrete phenotypes adapted to a single optimum each (*limited adaptation*, ***Figure 4***), whilst local genetic variance is very small (defined as less than 10% of the predicted adaptive variance in the absence of genetic drift). Simulations were run for 5000 generations, starting from a population adapted to a linearly changing optimum in the central part of the available habitat. Populations that went extinct are marked with a black dot. The top corner legend gives the colour coding for the rate of range collapse and expansion in units of demes per generation; rates of collapse are capped at -1. The *expansion threshold* is fitted as a step function changing linearly along B: all blue dots are assigned a value of 1; all red dots and open circles are assigned a value of 0. The coefficient of determination is $R^2 = 0.964$.

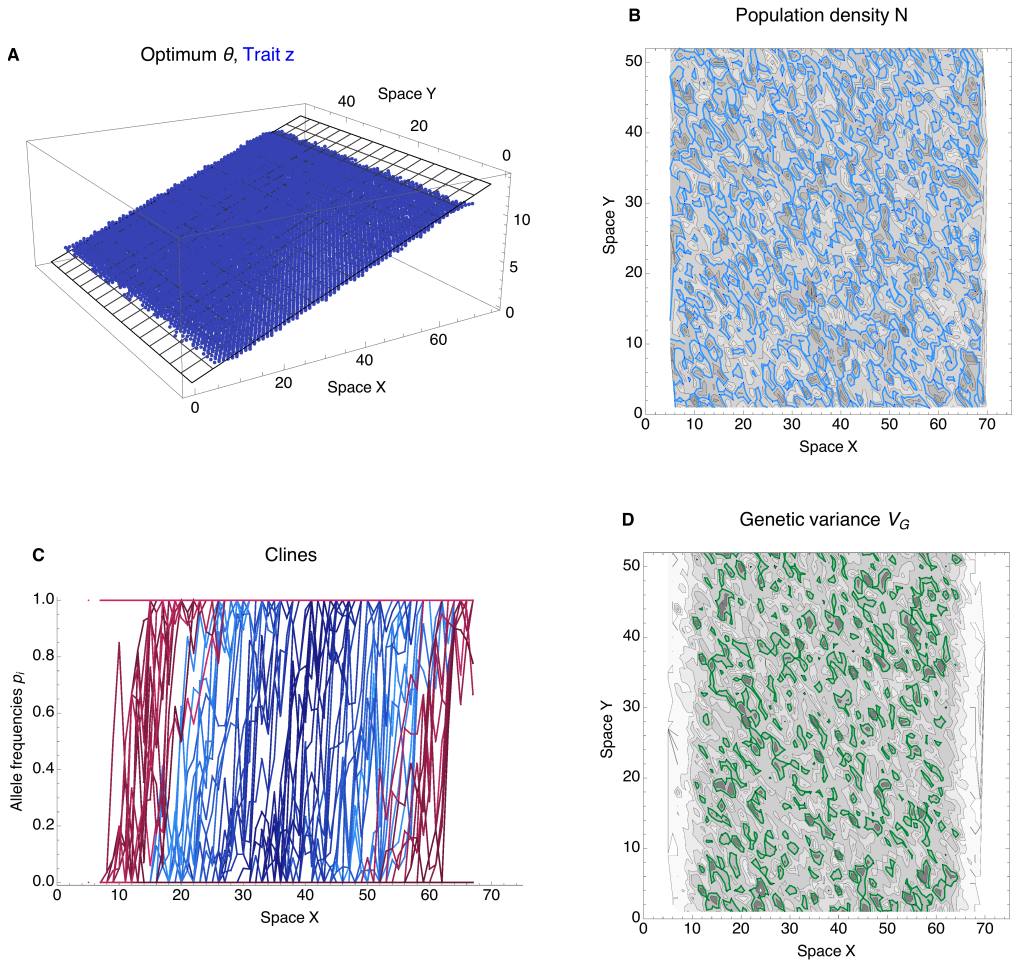


Figure 3. Uniform adaptation: above the *expansion threshold*, the population expands gradually through the available habitat. (**A**) Trait (in blue) is closely matching the environmental gradient (grey) along the X axis. (**B**) Population steadily expands, whilst population size varies across space with mean of $N = 19$ (the prediction at ($\hat{N} = 20$) and standard deviation about 5.8. (**C**) Population starts from the centre (blue clines) and as it expands new clines (red hues) contribute to the adaptation – as each allele frequency changes from 0 to 1, the trait value increases by α. At each location, multiple clines contribute to the trait (and variance); clines are shown at a middle transect ($Y = 25$). (**D**) Genetic variance changes continuously across space, with median slightly lower than the deterministic prediction (green contour). Deterministic predictions are based on the study by ***Barton*** (***2001***), and are explained in ***Methods***, along with the specification of the unscaled parameters. The population was evolving for 2000 generations, starting from a population adapted to the central part of the available habitat.

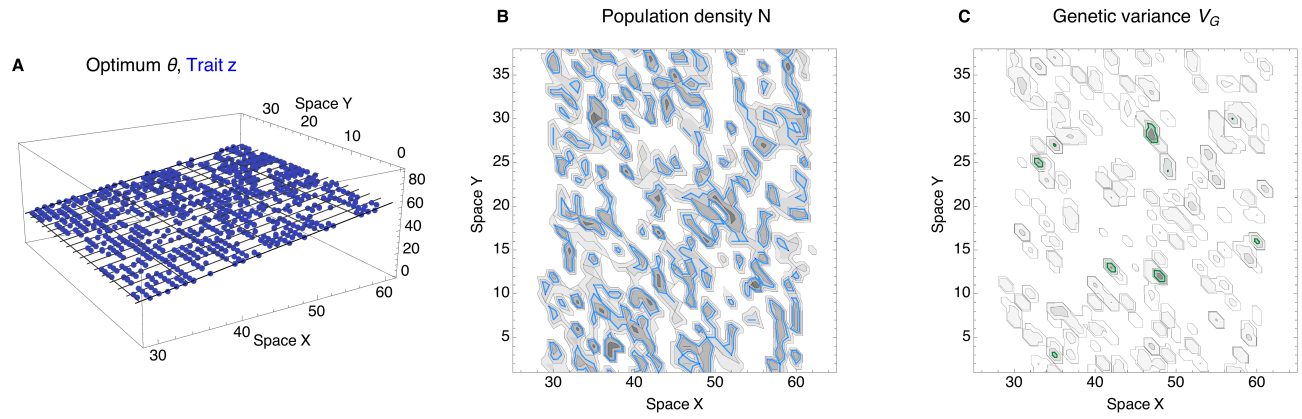


Figure 4. A metapopulation can form when the population is below the *expansion threshold* throughout its range. The population fragments rapidly (within tens of generations) to small patches of tens to a few hundred of individuals, whilst losing local adaptive variation. In two-dimensional habitats, such a metapopulation with *limited adaptation* can persist for long times. Nevertheless, the population very slowly contracts, eventually forming a narrow band adapted to a single optimum. (**A**) The distribution of phenotypes across space is fragmented. (**B**) The sub-populations are transient, although they are stabilised by dispersal across space, especially along the neutral direction with no change in the optimum (Y). Locally, the population density may be higher than under *uniform adaptation* (blue contours). (**C**) The adaptive genetic variance stays about an order of magnitude lower than would be maintained by gene flow under *uniform adaptation* (green contours): typically, only a few clines in allele frequencies contribute to the genetic variance within a sub-population. The parameterization is detailed in the ***Methods: Individual-based model***; predicted neighbourhood size is $\hat{\mathcal{N}} = 2.7$, effective environmental gradient is $B = 0.48$. Shown here after 5000 generations – the population collapses to a narrow band (at $X = 45$ and which appears persistent) after a further 20.000 generations.

the effect of genetic drift is nearly independent of selection in two-dimensional habitats (***Barton et al. 2002***), and because we are asking whether the species' range expands or adaptation collapses, rather than studying the rate of range expansion quantitively. The dashed lines in ***Figure 2*** compare the estimated expansion threshold for small and large s/r^*.

In nature, conditions are unlikely to change uniformly. Thus, I investigate whether adaptation fails near the *expansion threshold* as conditions change across space. For example, we can imagine that the population starts well adapted in the central part of the available habitat, and as it expands, conditions become progressively poorer (see ***Figure 5-Figure supplement 1***); i.e. the effective environmental gradient B gets steeper. As the expanding population approaches the *expansion threshold*, adaptive genetic variance increasingly decreases below the predicted value (***Barton 2001***), which would be maintained by gene flow in the absence of genetic drift (***Figure 5A***, gray dashed line). This is a result of an increased frequency of demes under *limited adaptation*, leading to higher rates of extinctions and re-colonizations, which reduce both adaptive and neutral diversity (see ***Figure 5B***). Range expansion then ceases at the *expansion threshold* as the genetic variance drops to the critical value where only *limited adaptation* is stable (***Kirkpatrick and Barton 1997***) assuming genetic variance is fixed (***Figure 5A***, dotted line). This is because although populations can persist with *limited adaptation* (***Figure 4***), the transient amount of genetic variance maintained under *limited adaptation* is almost never consistent with range expansion (see ***Figure 2***, open circles).

In a large population, the ability to adapt to heterogeneous environments is independent of dispersal: this is because both the local genetic variance (measured by *segregation load*), which enables adaptation to spatially variable environments, and the perceived steepness of the environmental gradient (*dispersal load*), increase quadratically with gene flow (***Barton 2001***). Yet, population density declines gradually due to these combined fitness costs – and this effect gets increasingly detrimental as the population size decreases. In small populations, dispersal has a net beneficial effect because the neighbourhood size $\mathcal{N}$ increases faster with dispersal than the effect of swamping by gene

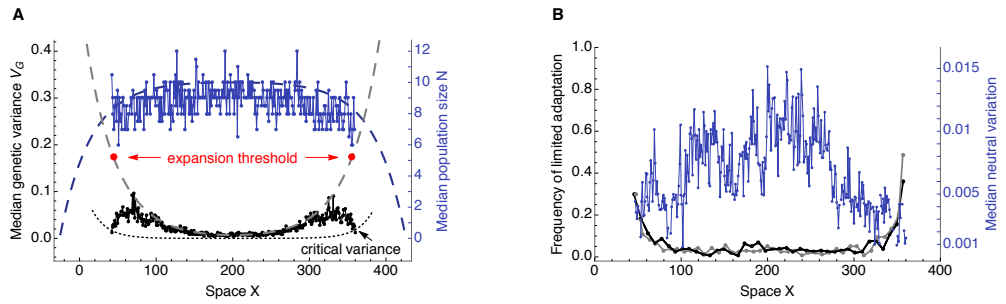


Figure 5. On a steepening environmental gradient, a sharp range margin forms near the *expansion threshold*. As the effective environmental gradient steepens away from the central location, adaptive genetic variance must increase correspondingly in order to maintain *uniform adaptation*. (**A**) The blue dots depict median population density (over the neutral spatial axis Y), whilst the expected density $\hat{N}$, is given by the blue dashed line. Genetic variance (black dots) increases due to gene flow across the phenotypic gradient – the deterministic expectation is given by the gray dashed line. (See ***Methods: Model*** for details.) Yet as the environmental gradient steepens, genetic variance fails to increase fast enough and near the *expansion threshold*, adaptation fails. The dotted line gives the corresponding critical genetic variance, below which only *limited adaptation* is expected in a phenotypic model with a fixed genetic variance ($A \approx B^2/2$, where A is the standing genetic load; ***Kirkpatrick and Barton 1997***). (**B**) As the environmental gradient steepens, the frequency of *limited adaptation* within the metapopulation increases (black and gray) and hence neutral variation decreases (blue). The black line gives the proportion of demes with *limited adaptation* after 50.000 generations, when the range margin appears stable; grey gives the proportion after 40.000 generations (depicted is an average over a sliding window of 15 demes).

Figure 5–Figure supplement 1. Shows the population trait mean, variance and population density in two-dimensional space, and also lists all the parameters.

flow (B): the drift-reducing effect of dispersal overpowers its maladaptive effect. Hence, near the *expansion threshold*, an increase of dispersal is favourable. This effect is demonstrated in ***Figure 6***. It gives a general, scale-free prediction of when "intermediate" dispersal is beneficial for adaptation across a species' range.

Discussion

Here I have shown that adaptation fails when positive feedback between genetic drift, maladaptation and population size reduces adaptive genetic variance to levels which are incompatible with continuous adaptation. The revealed *expansion threshold* differs qualitatively from the limit to adaptation identified previously (***Polechová and Barton 2015***) for a population living along a one-dimensional habitat because dispersal in two dimensions weakens the loss of diversity due to genetic drift more effectively, such that it becomes independent of selection. The *expansion threshold* implies that populations with very small neighbourhood sizes ($\mathcal{N} \lessapprox 1$), which suffer a severe reduction in neutral heterozygosity, will be prone to collapse based on demographic stochasticity alone. However, even in the absence of demographic stochasticity, genetic drift reduces the adaptive genetic variance required to sustain adaptation to a heterogeneous environment. The *expansion threshold* describes when this reduction due to genetic drift is incompatible with continuous adaptation, predicting a collapse of a species' range. If the *expansion threshold* is reached at an existing range margin as the species expands through its habitat, a sharp and stable range margin forms. If there is a drop below the *expansion threshold* throughout the species' range, as after a sudden drop in carrying capacity, adaptation abruptly collapses throughout a species' range. The result is either extinction, or a fragmented metapopulation consisting of a conglomerate of sub-populations, each adapted to a single phenotypic optimum. It follows that near a range margin, we expect increased range fragmentation, and a decrease in adaptive genetic variance. The threshold gives a theoretical base to the controversial issue of the importance of evolution (genetics)

and ecology (demography) for assessing vulnerability of a species (***Spielman et al. 2004***; ***Lande 1988***). The predicted sharp species' range edge is in agreement with the reported lack of evidence for *abundant centre* of a species' range, which although commonly assumed in macroecological theory, has little support in data (***Sagarin and Gaines 2002***; ***Dallas et al. 2017***).

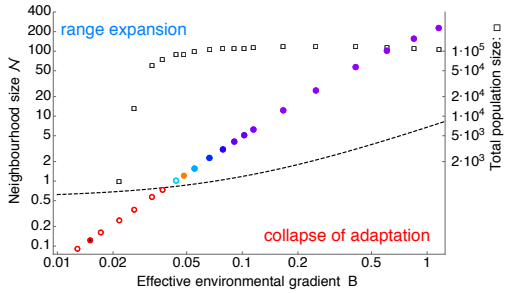


Figure 6. Dispersal aids adaptation in small populations. As dispersal σ increases, neighbourhood size $\mathcal{N} = 2\pi N\sigma^2$ increases faster that the effective environmental gradient $B = b\sigma/(r^*\sqrt{2V_s})$ – and population gets above the *expansion threshold* (dashed line). In red are collapsing populations (black centre indicates extinction), in blue are expanding populations. The hue indicates the rate of expansion or collapse (from light to dark blue) and from orange to red, respectively (the rate of range change is not significantly different from zero for first three simulations above the *expansion threshold*). Open circles indicate *limited adaptation*, where range is fragmented and each subpopulation is only matching a single optimum whilst its genetic variance is very small. Local population size stays fairly constant (around $N = 3.6$), there is only minor decrease (towards $N = 3.4$) as the dispersal load (B^2) and standing load due to genetic variance increase with higher dispersal. In contrast, total population size □ changes abruptly as population when population sustains continuous adaptation (above the *expansion threshold*).

Importantly, these predictive scale-free parameters can be measured in wild populations: the effective environmental gradient B can be measured as fitness loss associated with transplant experiments across *vs.* along an environmental gradient. The environmental gradient can include the presence of competitors (***Case and Taper 2000***). The neighbourhood size $\mathcal{N}$ can be estimated from neutral allele frequencies (***Slatkin and Barton 1989***; ***Rousset 1997***). Estimates of neighbourhood size are fairly robust distribution of dispersal distances (***Charlesworth et al. 2003***), yet near the *expansion threshold*, both the noisiness of the statistics and the homozygosity will increase due to local extinctions and recolonizations (***Whitlock and McCauley 1999***). An alternative estimate of neighbourhood size can be also obtained from mark-recapture studies, by measuring population density and dispersal (as an approximation for gene flow) independently (***Slatkin and Barton 1989***). In general, while the numerical constants may change when natural populations deviate in their biology from our model assumptions, the scale-free parameters identified in this study remain core drivers of the intrinsic dynamics within a species' range. Because the *expansion threshold* is free of any locus- or trait- specific measure, the result it is readily applicable to a species' range- and niche- dynamics which includes adaptation in multiple traits, assuming any pre-existing trade-offs can gradually evolve.

Methods

Model

I model evolution of a species' range in two-dimensional habitat, where both population dynamics and evolution (in many additive loci) is considered jointly. The coupling is *via* the mean fitness $\bar{r}(\bar{z}, N)$, which gives the growth rate of the population, and decreases with increasing maladaptation: $\bar{r}(\bar{z}, N) = r_e(N) + \bar{r}_g(\bar{z})$. The ecological component of growth rate, r_e, can take various forms: here the regulation is logistic, so that fitness declines linearly with density N: $r_e = r_m(1 - N/K)$, where r_m is the maximum per capita growth rate in the limit of the local population density $N \to 0$. The carrying capacity K (for a perfectly adapted phenotype) is assumed uniform across space. The second term, $r_g(\bar{z}) \leq 0$, is the reduction in growth rate due to deviation from the optimum. Selection is stabilising: the optimum θ changes smoothly with one spatial dimension (x): for any individual, the

drop in fitness due to maladaptation is $r_g(z) = -(z-\theta)^2/(2V_s)$. A population with mean phenotype $\bar{z}$ has its fitness reduced by $\bar{r}_g(\bar{z}) = -(\bar{z}-\theta)^2/(2V_s) - V_P/(2V_s)$, where $V_P = V_G + V_E$ is the phenotypic variance. The phenotype z is determined by many di-allelic loci with allelic effects α_i; the model is haploid, hence $\bar{z} = \sum_i \alpha_i p_i$, where p_i is the allele frequency at locus i. Because the loss of fitness due to environmental variance V_E can be included in $r_m^* = r_m - V_E/(2V_s)$; V_E is a redundant parameter. The selection is *hard*: both the mean fitness (growth rate) and the attainable equilibrium density $\hat{N} = Kr^*/r_m = K(1-V_G/(2r_mV_s))$ decrease with maladaptation. Expected genetic variance maintained by gene flow in the absence of genetic drift is $V_G = b\sigma V_s$ (***Barton 2001***); the contribution due to mutation is small, at mutation-section balance $V_{G,\mu/s} = 2\mu V_s n_l$.

Individual-based simulations

Discrete-time individual based simulations are set to correspond to the model with continuous time and space. The space is a two-dimensional lattice with spacing between demes of $\delta x = 1$. Every generation, each individual mates with a partner drawn from the same deme, with probability proportional to its fitness, to produce a number of offspring drawn from a Poisson distribution with mean of $Exp[r(z,N)]$. The effective diploid population density N_e hence equals half of the haploid population density N, and $\mathcal{N} = 4\pi N_e\sigma^2 = 2\pi N\sigma^2$. The life-cycle is selection → mutation → recombination → migration. Generations are non-overlapping and selfing is allowed at no cost. The genome is haploid with unlinked loci (the probability of recombination between any two loci is $1/2$). The allelic effects α_i of the loci combine in an additive fashion; the allelic effects are uniform throughout this study, $\alpha_i \equiv \alpha$. Mutation is set to $\mu = 10^{-6}$, independently of the number of loci. Migration is diffusive with a Gaussian dispersal kernel. The tails of the dispersal kernel need to be truncated: truncation is set to two standard deviations of the dispersal kernel throughout, and dispersal probabilities and variance are adjusted so that the discretised dispersal kernel sums to 1 (***Polechová and Barton 2005***, p. 1209). Simulations were run at the computer cluster of IST Austria using *Mathematica 9 (Wolfram)* ; the code used for simulations is available at request.

Parameters: There are in total 10 parameters in the individual-based model but only 7 are used to describe the model dynamics in continuous time. These are listed in ***Table 1*** (bottom). They are the environmental gradient $b = [0.012, 2]$, dispersal distance $\sigma = [0.1, 1.3]$, carrying capacity for a well adapted phenotype $K = [3, 31]$, width of stabilising selection $V_s = [0.005, 6]$, the maximum intrinsic rate of increase $r_m = [0.2, 2]$ and the mutation rate μ, fixed to $\mu \equiv 10^{-6}$. The $[x, y]$ interval gives the parameter range used in the 750 randomly sampled runs, with their distributions described in ***Figure 7***. The number of genes and demes are not included in the continuous time description (and hence the rescaling) because it assumes that space is not limiting, and that all loci have equivalent effect with no statistical associations among them. In the individual-based model, the habitat width is set to be wide enough to be effectively two-dimensional under diffusive dispersal for thousands of generations (***Slatkin 1993***) – 100 dispersal distances σ along the neutral direction, and at least 10 cline (deterministic) widths wide along the gradient. The number of genes contributing to the adaptation across the species' range is $n_l = [5, 3000]$, with the number of locally polymorphic genes between 1 and 300. Since mutation rate is fixed at $\mu = 10^{-6}$, the genomic mutation rate has a wide range, $U = [5 \cdot 10^{-6}, 3 \cdot 10^{-3}]$, with median of $U = 10^{-4}$.

Parameters for ***Figure 3*** are: $b = 0.18$, $\sigma = 0.52$, $V_s = 0.23$, $K = 26.7$, $r_{m_0} = 1$ and $\alpha = 0.14$, $s = 0.04$, $\mu = 10^{-6}$, 97 genes. Median genetic variance is at $V_G = 0.031$, deterministic prediction $V_G = b\sigma\sqrt{V_s} = 0.45$ (***Barton 2001***). In ***Figure 4***, the parameters are: $b = 1$, $\sigma = 0.4$, $V_s = 0.4$, $K = 4$, $r_m = 1.2$ and $\alpha = 0.1$, $s = 0.015$, $\mu = 10^{-6}$, 874 genes. Median genetic variance within patches is around 0.02, whilst the maximum contribution by a single cline $1/4\alpha^2 = 0.0026$; in contrast, variance maintained by gene flow under *uniform adaptation* (***Barton 2001***) would be $V_G = b\sigma\sqrt{V_s} \doteq 0.25$. Parameters for ***Figure 6*** are $b_0 = 0.3$, $\sigma = [0.05, 3]$, $V_s = 1$, $K = 4$, $r_{m_0} = 1$ and $\alpha = 0.05$, $s = 0.1$, $\mu = 10^{-6}$,

1000 genes, 1000 demes along X, 200 demes along Y. These populations were evolving for 500 generations.

Table 1. Three scale-free parameters: $\mathbf{B}$, $\mathcal{N}$ and $\mathbf{s/r^*}$ (top) describe the system. Middle section gives informative derived parameters. The bottom section gives seven parameters of the model before rescaling; where the seventh parameter, mutation rate μ, can be neglected because variance maintained by mutation-selection balance, $V_{G,\mu/s} = 2\mu V_s n_l$, is typically much smaller than variance generated by gene flow across environments, $V_G = b\sigma V_s$.

param.	dim.	
B	–	effective environmental gradient $B = b\sigma/(r^*\sqrt{2V_s})$
$\mathcal{N}$	–	neighbourhood size $\mathcal{N} = 4\pi N_e \sigma^2 = 2\pi N \sigma^2$
s/r^*	–	strength of selection per locus relative to the strength of density-dependence
s	$1/T$	selection per locus: $s \equiv \alpha^2/(2V_s)$
r^*	$1/T$	rate of return to equilibrium pop. size: $r^* \equiv -N\,\partial\bar{r}/\partial N\|_{N\to\hat{N}} = r_m - V_G/(2V_s)$
b	Z/D	gradient in the environmental optimum
V_s	Z^2T	variance of stabilising selection
σ	$D/\sqrt{T}$	dispersal per generation
K	T/D^2	max. carrying capacity (haploid) K = N when all phenotypes are perfectly adapted
r_m	$1/T$	max. intrinsic rate of increase
α	Z	allelic effect
μ	$1/T$	mutation rate, $\mu \equiv 10^{-6}$

Continuous model

For any given additive genetic variance V_G (assuming a Gaussian distribution of breeding values), the trait mean $\bar{z}$ satisfies:

$$\frac{\partial\bar{z}}{\partial t} = \frac{\sigma^2}{2}\left(\frac{\partial^2\bar{z}}{\partial x^2} + \frac{\partial^2\bar{z}}{\partial y^2}\right) + \sigma^2\left(\frac{\partial^2\ln(N)}{\partial x}\frac{\partial\bar{z}}{\partial x} + \frac{\partial^2\ln(N)}{\partial y}\frac{\partial\bar{z}}{\partial y}\right) + V_G\frac{\partial\bar{r}}{\partial\bar{z}} + \zeta \tag{1}$$

The first term gives the change in the trait mean due to migration with mean displacement of σ, the second term describes the effect of the asymmetric flow from areas of higher density and the third term gives the change due to selection (***Pease et al. 1989***, Eq. 2). The last term ζ gives the fluctuations in the trait variance due to genetic drift: $\zeta = \sqrt{V_{G,LE}/N}\,dW_\zeta(x,y,t)$, where dW_* represents white noise in space and time (***Nagylaki 1978***; ***Barton et al. 2002***). $V_{G,LE} = \sum_i \alpha_i^2 p_i q_i$ denotes genetic variance assuming linkage equilibrium.

The trait mean is $\bar{z} = \sum_i \alpha_i p_i$ for a haploid model,where p_i is the i-th allele frequency, $q_i = 1 - p_i$ and α_i is the effect of the allele on the trait – the change of the trait mean $\bar{z}$ as frequency of locus i changes from 0 to 1. For both haploid and diploid models, the allele frequencies p_i change as:

$$\frac{\partial p_i}{\partial t} = \frac{\sigma^2}{2}\left(\frac{\partial^2 p_i}{\partial x^2} + \frac{\partial^2 p_i}{\partial y^2}\right) + \sigma^2\left(\frac{\partial p_i}{\partial x}\frac{\partial\ln(N)}{\partial x} + \frac{\partial p_i}{\partial y}\frac{\partial\ln(N)}{\partial y}\right) + p_i q_i\frac{\partial\bar{r}}{\partial p_i} - \mu(p_i - q_i) + \epsilon \tag{2}$$

The expected change of allele frequency due to a gradient in fitness and local heterozygosity is $p_i q_i \frac{\partial\bar{r}}{\partial p_i} = s_i\, p_i q_i (p_i - q_i - 2\Delta_i)$, where selection at locus i is $s_i \equiv \alpha_i^2/(2V_s)$ and $\Delta_i = (\bar{z} - bx)/\alpha_i$ (***Barton 2001***, Appendix 3). Here, the fourth term describes the change due to (symmetric) mutation at rate

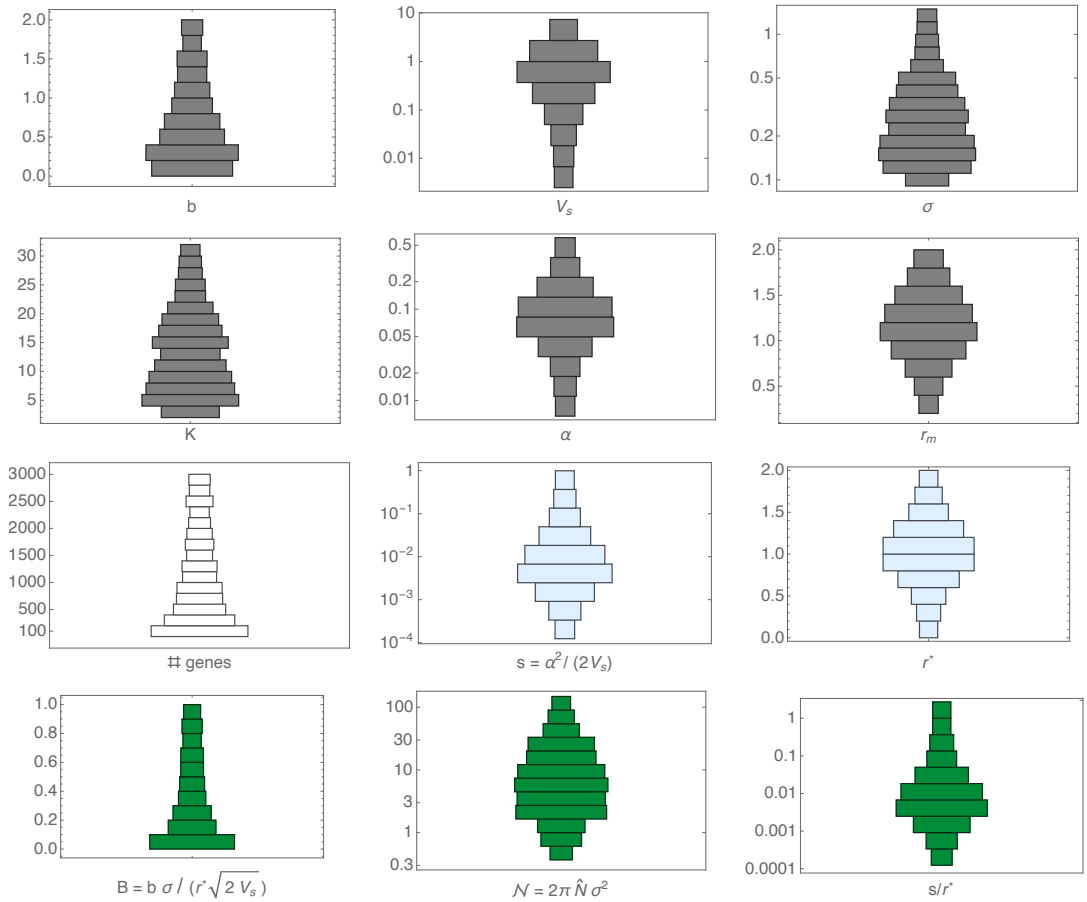


Figure 7. Distribution of parameters used in the "random set", as described in *Table 1*. *Grey:* all unscaled parameters used in the continuous time and space model. The ranges for the unscaled parameters are: environmental gradient $b = [0.012, 2]$, dispersal distance $\sigma = [0.1, 1.3]$, carrying capacity for a well adapted phenotype $K = [3, 31]$, width of stabilising selection $V_s = [0.005, 6]$, the maximum intrinsic rate of increase $r_m = [0.2, 2]$. *White:* the number of genes (an extra parameter in the individual-based model). The number of genes contributing to the adaptation across the species' range is $n_l = [5, 3000]$, with the number of locally polymorphic genes between 1 and 300. *Light blue:* composite parameters: the strength of selection, $s = \alpha^2/(2V_s) = [0.0002, 0.5]$, with median of $s = 0.006$. The effect size $\alpha = [0.01, 0.4]$. *Dark green:* scale-free parameters: the effective environmental gradient $B = [0.01, 1]$, with median $B = 0.14$; the neighbourhood size $\mathcal{N} = [0.2, 60]$, with median 2.8; strength of selection relative to the strength of density dependence, $s/r^* = [0.00015, 1.34]$, with median $s/r^* = 0.006$. Not pictured is the uniform mutation rate per locus ($\mu = 10^{-6}$), the number of demes along the two spatial directions, which was set to min. 10 deterministic cline widths along the spatial axis with environmental gradient (Space X), whereas along the neutral spatial axis (Space Y) the habitat width was 100σ.

μ. The last term ϵ describes genetic drift (***Barton et al. 2002***, Eq. 7): $\epsilon = \sqrt{\frac{p_i q_i}{N}}\, dW_\epsilon(x, y, t)$; where N is the haploid population density.

Population dynamics reflect diffusive migration in two-dimensional habitat, growth due to the mean Malthusian fitness $\bar{r}$, and stochastic fluctuations. The number of offspring follows a Poisson distribution with mean and variance of N, fluctuations in population numbers are given by ***Lande et al.*** (***2003***): $\xi = \sqrt{N}\, dW_\xi(x, y, t)$:

$$\frac{\partial N}{\partial t} = \frac{\sigma^2}{2}\left(\frac{\partial^2 N}{\partial x^2} + \frac{\partial^2 N}{\partial y^2}\right) + \bar{r}N + \xi \tag{3}$$

Continuous model: Rescaling

The model can be simplified by rescaling (***Pease et al. 1989***; ***Barton 2001***) time t relative to the strength of density dependence r^*, distance x relative to dispersal σ, trait z relative to strength of stabilising selection $1/(2V_s)$ and local population size N relative to equilibrium population size with perfect adaptation: $\hat{N} = Kr^*/r_m,\ T = r^*t,\ X = x\sqrt{\frac{2r^*}{\sigma^2}},\ Z = \frac{z}{\sqrt{r^*V_s}},\ \tilde{N} = N/\hat{N}$. Note that near the equilibrium of a well-adapted population, $\tilde{N} \approx 1$.

The rescaled equations for evolution of allele frequencies and for demographic dynamics are:

$$\frac{\partial \tilde{N}}{\partial T} = \frac{\partial^2 \tilde{N}}{\partial X^2} + \frac{\partial^2 \tilde{N}}{\partial Y^2} + \overline{R}\tilde{N} + \sqrt{\frac{2\tilde{N}}{\hat{N}\sigma^2}}\, dW_{\tilde{\xi}(X,Y,T)} \tag{4}$$

$$\begin{aligned}\frac{\partial p_i}{\partial T} &= \frac{\partial^2 p_i}{\partial X^2} + \frac{\partial^2 p_i}{\partial Y^2} + 2\left(\frac{\partial p_i}{\partial X}\frac{\partial \ln(\tilde{N})}{\partial X} + \frac{\partial p_i}{\partial Y}\frac{\partial \ln(\tilde{N})}{\partial Y}\right) + \\ &+ \frac{s}{r^*}\left(p_i q_i - 2\frac{\overline{Z} - BX}{\alpha^*}\right) - \frac{\mu}{r^*}(p_i - q_i) + \sqrt{\frac{pq}{\tilde{N}\hat{N}\sigma^2}}\, dW_{\tilde{\epsilon}(X,Y,T)}\end{aligned} \tag{5}$$

where $\overline{R} \equiv \bar{r}/r^* = 1 - \tilde{N} - (BX - Z)^2/2$.

The rescaled equations 4 and 5 show that four parameters fully describe the system. The first two are i) the effective environmental gradient, $B \equiv b\sigma/(r^*\sqrt{2V_s})$ and ii) the strength of genetic drift $1/\hat{\mathcal{N}} = 1/(2\pi\hat{N}\sigma^2)$. The third parameter is the strength of selection relative to the strength density dependence, s/r^*; the scaled effect of a single substitution α^* also scales with s/r^*: $\alpha^* \equiv \alpha/\sqrt{r^*V_s} = \sqrt{2s/r^*}$. The effect of this third parameter s/r^* is expected to be small, because typically, $s \lll r^*$. Therefore assuming throughout that s is uniform across loci is a reasonably justified simplification. The fourth parameter, μ/r^*, will typically be very small, and will be neglected throughout. ***Table 1*** (top) summarises the full set that describes the system.

Acknowledgments

I would like to thank my friends and colleagues for their suggestions and comments: Andrea Betancourt, David Field, Jarrod Hadfield, Joachim Hermisson, Sebastian Matuszewski (who also created the flower art), Glenn Marion, Richard Nichols, Tiago Paixão and Derek Setter. Special thanks to Nick Barton and to the computer cluster team of IST Austria for their support long after I have left IST Austria. The work has been funded by Elise Richter Programme, Austrian Science Fund (FWF); project number V 458-B25.

Citations

Alleaume-Benharira M, Pen I, Ronce O. Geographical patterns of adaptation within a species' range: interactions between drift and gene flow. Journal of Evolutionary Biology. 2006; 19(1):203–215.

Antonovics J. The nature of limits to natural selection. Annals of the Missouri Botanical Garden. 1976; 63(2):224–247.

Barton NH. Adaptation at the edge of a species' range. In: Silvertown J, Antonovics J, editors. *Integrating ecology and evolution in a spatial context*, vol. 14 Blackwell; 2001.p. 365–392.

Barton NH, Depaulis F, Etheridge AM. Neutral evolution in spatially continuous populations. Theoretical Population Biology. 2002; 61(1):31–48.

Cahill AE, Levinton JS. Genetic differentiation and reduced genetic diversity at the northern range edge of two species with different dispersal modes. Molecular ecology. 2016; 25(2):515–526.

Case TJ, Taper ML. Interspecific competition, environmental gradients, gene flow, and the coevolution of species' borders. The American Naturalist. 2000; 155(5):583–605.

Charlesworth B, Charlesworth D, Barton NH. The effects of genetic and geographic structure on neutral variation. Annual Review of Ecology, Evolution, and Systematics. 2003; 34(1):99–125.

Chevin L, Lande R. When do adaptive plasticity and genetic evolution prevent extinction of a density-regulated population? Evolution;. 2010; 64(4):1143–1150.

Christiansen FB. Hard and soft selection in a subdivided population. American Naturalist. 1975; 109(965):11–16.

Dallas T, Decker RR, Hastings A. Species are not most abundant in the centre of their geographic range or climatic niche. Ecology Letters. 2017; .

Eckert C, Samis K, Lougheed S. Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. Molecular ecology. 2008; 17(5):1170–1188.

Gaston KJ. The structure and dynamics of geographic ranges. Oxford University Press; 2003.

Gilbert KJ, Whitlock MC. The genetics of adaptation to discrete heterogeneous environments: frequent mutation or large-effect alleles can allow range expansion. Journal of Evolutionary Biology. 2017; .

Gomulkiewicz R, Holt RD, Barfield M. The effects of density dependence and immigration on local adaptation and niche evolution in a black-hole sink environment. Theoretical Population Biology. 1999; 55(3):283–296.

Griggs RF. Observations on the behaviour of some species at the edges of their ranges. Bulletin of the Torrey Botanical Club. 1914; 41:25–49.

Haldane JBS. The cost of natural selection. Journal of Genetics. 1957; 55(3):511–524.

Hargreaves AL, Samis KE, Eckert CG. Are species' range limits simply niche limits writ large? A review of transplant experiments beyond the range. The American Naturalist. 2013; 183(2):157–173.

Holt RD, Gomulkiewicz R. How does immigration influence local adaptation? A reexamination of a familiar paradigm. The American Naturalist. 1997; 149(3):563–572.

Kawecki TJ. Adaptation to marginal habitats. Annual Review of Ecology, Evolution, and Systematics. 2008; 39:321–342.

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

Lande R. Genetics and demography in biological conservation. Science. 1988; 241(4872):1455–1460.

Lande R, Engen S, Sæther BE. Stochastic population dynamics in ecology and conservation. Oxford University Press; 2003.

Nagylaki T. Random genetic drift in a cline. Proceedings of the National Academy of Sciences of the United States of America. 1978; 75(1):423–426.

Nosil P, Crespi B. Does gene flow constrain adaptive divergence or vice versa? A test using ecomorphology and sexual isolation in *Timema cristinae* walking-sticks. Evolution. 2004; 58(1):102–112.

Pease CM, Lande R, Bull JJ. A model of population growth, dispersal and evolution in a changing environment. Ecology. 1989; 70(6):1657–1664.

Polechová J, Barton NH. Speciation through competition: A critical review. Evolution. 2005; 59(6):1194–1210.

Polechová J, Barton N. Limits to adaptation along environmental gradients. Proceedings of the National Academy of Sciences of the United States of America. 2015; 112(20):6401–6406. doi: 10.1073/pnas.1421515112.

Rousset F. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics. 1997; 145(4):1219–1228.

Sagarin RD, Gaines SD. The 'abundant centre' distribution: to what extent is it a biogeographical rule? Ecology letters. 2002; 5(1):137–147.

Sexton JP, Strauss SY, Rice KJ. Gene flow increases fitness at the warm edge of a species' range. Proceedings of the National Academy of Sciences. 2011; 108(28):11704–11709.

Sexton J, McIntyre P, Angert A, Rice K. Evolution and ecology of species range limits. Annual Review of Ecology, Evolution, and Systematics. 2009; 40:415–436.

Slatkin M. Isolation by distance in equilibrium and non-equilibrium populations. Evolution. 1993; 47(1):264–279.

Slatkin M, Barton NH. A comparison of three indirect methods for estimating average levels of gene flow. Evolution. 1989; 43(7):1349–1368.

Spielman D, Brook BW, Frankham R. Most species are not driven to extinction before genetic factors impact them. Proceedings of the National Academy of Sciences of the United States of America. 2004; 101(42):15261–15264.

Takahashi Y, Suyama Y, Matsuki Y, Funayama R, Nakayama K, Kawata M. Lack of genetic variation prevents adaptation at the geographic range margin in a damselfly. Molecular ecology. 2016; 25(18):4450–4460.

Uecker H, Otto SP, Hermisson J. Evolutionary rescue in structured populations. The American Naturalist. 2014; 183(1):E17–E35.

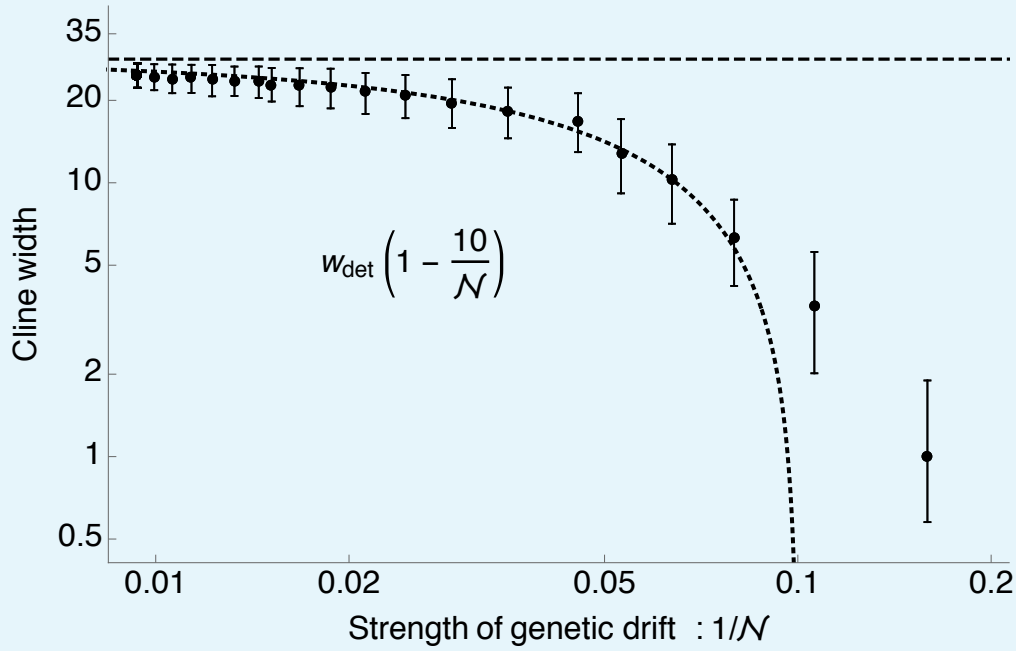
Whitlock MC, McCauley DE. Indirect measures of gene flow and migration: FST≠ 1/(4Nm+ 1). Heredity. 1999; 82(2):117–125.

Wright S. Evolution in Mendelian populations. Genetics. 1931; 16(2):97–159.

Wright S. Isolation by distance under diverse systems of mating. Genetics. 1946; 31(1):39.

Appendix 1

Effect on genetic drift on a cline width in two-dimensional habitats



Appendix 1 Figure 1. Cline width – and hence the local genetic variance – decrease with genetic drift. This is because the spacing between clines stays at α/b, and hence less clines contribute to genetic variance $V_G = \sum_i \alpha_i^2 p_i q_i$ when they are steeper. The numerical approximation $w_{\mathcal{N}} = w_{det}(1 - 10/\mathcal{N})$ works well for weak genetic drift but breaks down for very small neighbourhood sizes (strong genetic drift, right). The deterministic cline width is given by $w_{det} = \frac{4\sigma}{\sqrt{2s}}$. The figure represents a preliminary look at a complex problem that is a planned subject of a dedicated paper.

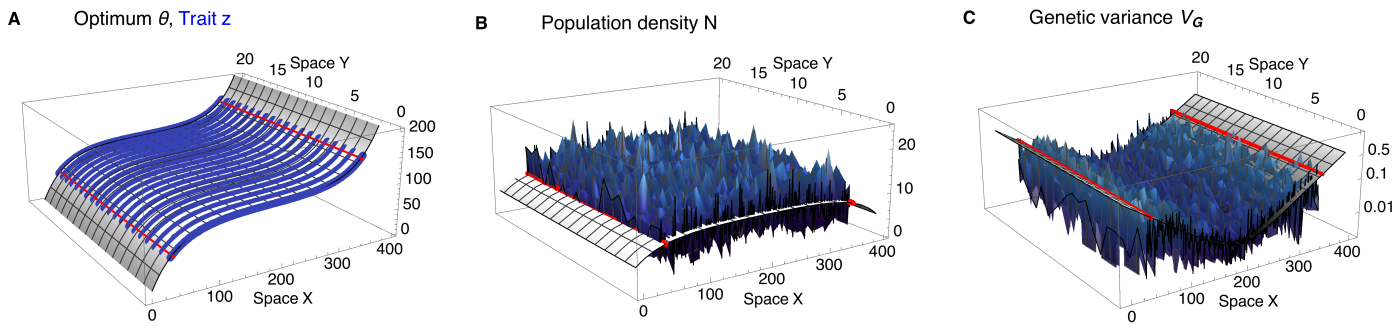


Figure 5–Figure supplement 1. With a steepening environmental gradient, a sharp range margin forms due to genetic drift. (**A**) The gradient in trait mean follows the steepening environmental gradient, changing across one of the spatial axes, X (gray mesh). The red lines show the *expansion threshold* $\mathcal{N} \approx 6.3\, B + 0.56$. (**B**) As the environmental gradient steepens, local population density gradually declines (the gray mesh gives the predicted density). (**C**) As the environmental gradient steepens, so does the predicted genetic variance maintained by gene flow in the absence of genetic drift (gray mesh; $V_G = b\,\sigma V_s$, ***Barton 2001***). Pass the threshold, genetic variance starts to abruptly fall off the prediction, and adaptation begins to fail. The median values (across the neutral Y-space) are shown in **Figure 5**. Parameters: $\sigma = 0.2$, $V_s = 0.5$, $K_0 = 10$, $r_{m_0} = 1$ and $\alpha = 0.1$, $s = 0.01$, $\mu = 10^{-6}$, 2000 genes. The environmental gradient steepens away from the centre with $b_0 = 0.05$ according to $b(x) = 1/40 \exp(-1/40(x - 200)) + 1/40 \exp(1/40(x - 200))$. The population was evolving for 50.000 generations, starting from a population adapted to the central part of the available habitat, about 100 demes wide.