Genetic Drift Widens the Expected Cline but Narrows the Expected Cline Width

Jitka Polechová*1 and Nick Barton*†

*Institute for Science and Technology (IST Austria), Klosterneuburg A-3400, Austria and †University of Edinburgh, Institute for Evolutionary Biology, Edinburgh EH9 3J, Scotland, United Kingdom

ABSTRACT Random genetic drift shifts clines in space, alters their width, and distorts their shape. Such random fluctuations complicate inferences from cline width and position. Notably, the effect of genetic drift on the expected shape of the cline is opposite to the naive (but quite common) misinterpretation of classic results on the expected cline. While random drift on average broadens the overall cline in expected allele frequency, it narrows the width of any particular cline. The opposing effects arise because locally, drift drives alleles to fixation—but fluctuations in position widen the expected cline. The effect of genetic drift can be predicted from standardized variance in allele frequencies, averaged across the habitat: $\langle F \rangle$. A cline maintained by spatially varying selection (step change) is expected to be narrower by a factor of $\sqrt{1-\langle F \rangle}$ relative to the cline in the absence of drift. The expected cline is broader by the inverse of this factor. In a tension zone maintained by underdominance, the expected cline width is narrower by about $1-\langle F \rangle$ relative to the width in the absence of drift. Individual clines can differ substantially from the expectation, and we give quantitative predictions for the variance in cline position and width. The predictions apply to clines in almost one-dimensional circumstances such as hybrid zones in rivers, deep valleys, or along a coast line and give a guide to what patterns to expect in two dimensions.

WITHIN species, substantial genetic diversity may be maintained by the interaction between selection and gene flow. This may be manifest in gradients in heritable traits or in allele frequency (Endler 1977). Clines reflect the degradation by gene flow of adaptation to the local environment or genetic background, and they may be used to estimate the strength of selection experienced by natural populations. Often, it is assumed that the effects of drift on the realized clines can be neglected.

Random genetic drift will shift clines from side to side, alter their width, and distort their shape (Felsenstein 1975; Slatkin and Maruyama 1975; Nagylaki 1978). Such random fluctuations complicate inferences from cline width and position, but may themselves be used to infer rates of gene flow, drift, and selection. More important, drift is expected to reduce local diversity and so to make selection less effective, thereby interfering with adaptation to local conditions.

The common view is that genetic drift makes clines slightly shallower, because selection becomes less effective. This comes from a “generalization” of the effect of genetic drift on the “expected cline” as derived by Slatkin and Maruyama (1975) (see, e.g., Alleaume-Benharira et al. 2006). However, the characteristics of the expected cline do not necessarily give an adequate idea of the patterns we expect to see in specific clines. Here, we define the expected cline as the average allele frequency across independent realizations, taken at the same time, and so it includes the inevitable widening caused by shifts in position (see Figure 1). Even if each cline were exactly the same width, shifts in position would widen the expected allele frequency. This wobbling in position (already emphasized by Felsenstein 1975) will make the expected cline appear wider than any single realization. In contrast, because the allele frequencies tend to fix through genetic drift, an individual cline may in fact tend to be steeper. Indeed, Hallatschek and Korolev (2009, p. 108103-2) show that even with no selection, cline width stays finite, even though the expected cline (including fluctuation in position) gets ever broader.

Cline width is not a uniquely defined term (Endler 1977, pp. 32 and 61). Most often, one speaks of the inverse of the maximum slope of a cline, here denoted $w_{\text{max}}$. However, this...
width is inherently very noisy in a stochastic system: $w_{\text{max}}$ can approach infinity just due to random effects over small scales. Two other approximations to the cline width are commonly used:

i. $w_{\log}$, a linear regression on $\log(p/q)$. However, $\log(p/q)$ requires arbitrary cutoffs as one needs to discard fixed allele frequencies, which yet carry information: $w_{\log}$ can strongly depend on the chosen cutoff when fluctuations are significant.

ii. $w_{\text{ML}}$, the maximum-likelihood estimate for the shape of the cline and its slope at $p = \frac{1}{2}$, which does not require any cutoff, but depends on the model that is fitted (see Appendix A). With substantial fluctuations, though, clines can be distorted enough that cline “width” (however defined) does not have any consistent meaning.

Rather than estimating the strength of selection, we may be interested in the effects of selection: How much diversity is maintained across space? For this question, a different scheme (see Karlin and Richter-Dyn 1976; Barton and Gale 1993), the effect of genetic drift can be qualitatively different, depending on the form of selection. When selection depends on space, and is stable over time, the cline will wobble around a fixed position, the extent of wobbling being given by a balance between the selection on cline position and random drift. However, in a tension zone, for example maintained by position-independent hybrid incompatibility, the location is arbitrary, so the displacement of cline centers will diverge with (the square root of) time (Barton 1979). The expected cline then has a cline-like equilibrium only if we discount fluctuations in position first (i.e., by aligning the clines, so that they have the same position). With underdominance, the expected cline ultimately collapses to a uniform frequency of 0.5, even though the underlying clines stay sharp.

The effects of genetic drift on the expected cline can be calculated for discrete time and space or by using the spatial diffusion approximation (Fisher 1937). A continuous version of Equation 1 of Slatkin and Maruyama (1975) gives

$$\frac{\partial}{\partial t} \langle p(x, t) \rangle = \frac{\sigma^2}{2} \frac{\partial^2}{\partial x^2} \langle p(x, t) \rangle + s(x)(\langle p(x, t) \rangle \langle q(x, t) \rangle - \text{var}(p(x, t))).$$

where $\sigma$ is the standard deviation of dispersal distance, and the selective advantage $s(p, x)$ (to the first order) of an allele varies across space and may depend on allele frequency $(p, q = 1 - p)$. Whereas one cannot solve the spatial diffusion equation with selection and drift, because the fluctuations due to genetic drift $\xi(x, t)$ depend in general on both time and space, the equation for expected change in the allele frequency in a one-dimensional habitat can be solved analytically when selection depends only on location, $s(x)$.

Taking the expectation (denoted $\langle \rangle$) of the equation above (as in Slatkin and Maruyama 1975) gives

$$\frac{\partial}{\partial t} \langle p(x, t) \rangle = \frac{\sigma^2}{2} \frac{\partial^2}{\partial x^2} \langle p(x, t) \rangle + s(x)(\langle p(x, t) \rangle \langle q(x, t) \rangle - \text{var}(p(x, t))).$$

The variance term comes from $\text{cov}(p, q) = \langle pq \rangle - \langle p \rangle \langle q \rangle = -\text{var}(p)$, and the drift term disappears as the expected change due to genetic drift is zero, $\langle \xi(x, t) \rangle = 0$. Because the variance in allele frequency at any given time and space is always positive (and driven by genetic drift), genetic drift clearly weakens the effect of selection on the expected cline and hence broadens it.

An interesting insight follows directly from the paragraphs above. We can define the standardized local variance in allele frequency $p(x, t)$ as $F(x, t) = \text{var}(p(x, t))/\langle p(x, t) \rangle \langle q(x, t) \rangle$ and its weighted average [with $\langle p(x, t) \rangle \langle q(x, t) \rangle$] as $F(t) = \int_{-\infty}^{\infty} \text{var}(p(x, t)) dx/\int_{-\infty}^{\infty} \langle p(x, t) \rangle \langle q(x, t) \rangle dx$. The parameter $F(x, t)$ differs from what is typically measured as local $F_{ST}(x)$ along a cline, as $F(x, t)$ includes shifts in the position of the cline. For $F_{ST}$ along a cline, clines are typically aligned in space first, so that their centers match.
Clines maintained by selection across different environments reach a stationary distribution as the effect of fluctuations stabilizes; so we have \( F(t) = \mathcal{F} \). For spatially independent selection, as in tension zones, the position of clines will keep diverging in space over time—the formulas below are then valid for specific time \( t \) provided that the history of selection and mixing is the same for all replicates. Keeping that in mind, we write the random variable \( p(x, t) \) as \( p \) from now on.

From the definition of \( w_{pq} \) and \( \mathcal{F} \), the expected cline width (which includes fluctuations in position) is \( w_{pq}(p) = 4 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} q(x) (1 - \mathcal{F}(p,q) dx = (1 - \mathcal{F}) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} q(x) (1 - \mathcal{F}) dp \). The expected cline width \( w_{pq}(p) \) is thus narrower than the width of the expected cline \( w_{pq}(p) \) by a factor of \( (1 - \mathcal{F}) \).

Moreover, because \( (pq) = (1 - F)(p,q) \), when selection depends on the location only \( (s = s(x)) \), the selection term \( s(x)(pq) \) in Equation 2 can be replaced by \( s(x)(1 - F)(p,q) \). Now, \( F \) reaches an equilibrium over time but it still depends on space, so the equation for the equilibrium cline cannot be easily solved directly. Typically, \( F \) varies little along a cline (see Felsenstein 1975; Nagylaki and Lucier 1980), and hence we can take again the expected \( F \) across space, \( \mathcal{F} = F(x) \) as a good approximation. After the stationary distribution has been reached, the width of the expected cline must be close to \( w_{pq}(p) = w_{pq}(p_0) / \sqrt{1 - \mathcal{F}} \), as width of a cline in allele frequency \( p_0 \) in the absence of genetic drift, \( w_{pq}(p_0) \), is proportional to the characteristic length \( l = \sigma / \sqrt{2s} \). Therefore, the expected cline width must be narrower by the same factor, relative to the width of a cline in the absence of genetic drift: \( w_{pq}(p) = w_{pq}(p_0) \sqrt{1 - \mathcal{F}} \). If we measure \( F \) along the cline, the effect should be bounded between the minimum and maximum \( F \): \( \sqrt{1 - F_{\text{min}}} \) and \( \sqrt{1 - F_{\text{max}}} \).

In the case of tension zones, where clines wobble freely in space, the expected cline steadily widens with time due to fluctuations in position. Hence, the local standardized variance \( F \) increases indefinitely with time. Aligning the cline centers in space, we obtain the local standardized variance on aligned clines, \( F_A \), which reaches a stationary state. In experimental work on the clines, \( F_A \) is commonly referred to as \( F_{\text{ST}} \), and it is always smaller than \( F \) as it is missing the components due to shifts in position. The first part of the above argument still holds: the expected width of the aligned cline, \( w(p_A) \), should be narrower by a factor of \( (1 - (F_A)) \) than the width of the expectation of aligned clines, \( w(p_A) \), where \( F_A \) is the average standardized variance for the aligned clines. When the shifts in position of a cline are discounted, \( w(p_A) \) should not be wider than \( w_{pq}(p_0) \) because locally, drift tends to fix allele frequencies.

Nagylaki (1978) showed that in a one-dimensional habitat, the effects of drift on the expectation of the allele frequency and its variance \( (F) \) are proportional to \( (N_\text{\text{}} / \sqrt{s}) \), where \( s \) is the strength of selection, \( N_\text{\text{}} \) is the neighborhood size in one dimension \( [N = 2 / \pi \sigma \rho \text{ (Wright 1943)}] \), \( \sigma \) is the standard deviation of the dispersal distance, and \( \rho \) is the local (diploid) effective population density. The effective population density is the effective population size \( N_\text{\text{}} \) per distance; \( N_\text{\text{}} \) gives how much time it takes for a neutral locus to coalesce to a common ancestor. Throughout the article, we refer to this dimensionless parameter \( 1 / (N_\text{\text{}} \sqrt{s}) \) as the strength of drift.

### Model and Results

First, we outline a stepping-stone model for two forms of selection: (i) spatially dependent (like a step change of selection along a habitat) and (ii) free of spatial context—“tension zones” (here due to frequency-dependent selection that also gives an approximation for underdominance). Second, we relate the simulation results to our basic predictions on the expected cline width, outlined in the Introduction, and extend the predictions to variance in cline displacement and cline width.

#### Stepping-stone model

We take the simplest model, discrete in time and space: a one-dimensional array of demes, spaced at \( \delta x = 1 \), each deme with \( N \) haploid individuals. Generation time is \( \delta t = 1 \). Migration is to neighboring demes (stepping stone): \( N p(x, t') = Np(x, t) + Nm/2 (p(x + 1, t) + p(x - 1, t) - 2p(x, t)) \). Selection then takes two forms: (i) step change in selection, one allele has fitness of \( 1 - s \) in one-half of the habitat and \( 1 + s \) in the other, (relative to 1 of the other allele), and (ii) a linear frequency-dependent selection, where the fitness of the allele with frequency \( p \) is \( 1 - q \) (and vice versa)—as would be the marginal fitness with “underdominance” in a diploid model. With underdominance, allele frequency in the next generation after migration and selection is given by \( p(x, t + 1) = p(x, t) ((1 - sq(x, t')) / (1 - 2sp(x, t')) q(x, t')) \); for a step change it is \( p(x, t + 1) = p(x, t') ((1 + s(x)p(x, t')) / (1 + s(x) (p(x, t') - q(x, t')))) \).

Dropping terms of order \( \delta t^2 \), \( \delta t^3 \) and smaller and expanding \( p(x \pm \delta x) \) in series to the second order gives the diffusion equation (Equation 1, where \( \sigma^2 / 2 \approx 8 \delta x \sqrt{m/2} \)).

Throughout, migration rate is \( m = 0.5 \). Haploid population numbers in each deme are typically \( N = 60 \), so that the density of the corresponding diploid population is \( \rho = 30 \). (In test runs, other values of migration rates and population size have also been used.) The strength of selection \( s \) varies between 0.001 and 0.2. Simulations start at deterministic equilibrium and are then run for 1000 generations till the extent of fluctuations due to genetic drift equilibrates even for the weakest selection that we use (with the exception of the variance in the shift in position, which equilibrates slower for step change in selection weaker than 0.002). Allele frequencies are stored every 50 generations till generation 1000. We ran 500 replicates for each parameter combination. The length of the habitat is set up so that it is \( >10 \) times the cline width in the absence of drift. This is necessary to ensure that the only effect of the boundary conditions on the cline is to keep the equilibrium polymorphic
(i.e., the allele frequencies are fixed to 0 and 1 at the margins).

**Results and predictions**

We focus on the effects of drift on the variance in position of the center of the cline and on changes in the width of the cline, defined as the cumulative heterozygosity across all sites $i$: $w_{pq} = \sum p_i q_i \delta x_i$, where $\delta x_i$ are the distances between neighboring sites. In the simulation, $\delta x = 1$ and hence $w_{pq} = \sum p_i q_i$. Finally, we give estimates for the variance in cline width $w_{pq}$.

For a deterministic cline maintained by positive frequency-dependent selection (underdominance), $w_{pq} = w_{\text{max}}$; for step change in selection, $w_{\text{max}} = 3 \int_0^\infty p_i q_i d\delta x = (3/4)w_{pq}$. The distribution of the widths $w_{pq}$ and corresponding cline examples for the strength of drift $1/(N\sqrt{s}) = 0.2$ are shown in Figure 2. Most clines get narrower due to genetic drift, but a few can be much wider than the deterministic expectation. In tension zones with strong genetic drift, several serial clines can co-exist across a range—although for detectable selection, such clines are very rare and typically transient. On average, cline width under genetic drift gets slightly narrower.

**Variance in position of a cline**

Before assessing the joint predictions based on the average standardized variance in allele frequencies, $\langle F \rangle$, we turn to another effect of genetic drift on a cline: the placement of a cline in space. With underdominance, a cline will tend to wobble away from its prior location, as its position is arbitrary. The variance in allele frequency due to shifts in position of a cline grows linearly with time as expected for a random walk, and so does the variance in the cline displacement: $\text{var}(\Delta x) = (6/5) \cdot \sigma t/(2p)$, as illustrated in Figure 3A (for derivation, see Appendix B). In scaled units, variance in the shift in position is $\text{var}(\Delta X) = (6/5) \cdot \sqrt{2\pi T/(N\sqrt{s})}$, where $\Delta X = x/l$ scales with the characteristic length $l = \sigma/\sqrt{2s}$ and the strength of drift is given by $N\sqrt{s}$ where $N = 2\sqrt{\pi \sigma}$ is the neighborhood size and time scales with selection, $T = ts$.

With a step change, the variance in cline displacement reaches an equilibrium within a few tens to a few hundreds of generations, for selection ranging between 0.2 and 0.001. The eigenfunctions driving the fluctuations are more difficult to find but the variance in the cline displacement should increase proportionally with $w^2/(N\sqrt{s})$. Figure 3B shows a good agreement between the prediction and the simulation, unless selection is very weak (then $T > 1000$ is required to reach the equilibrium).

**Cline width and its variance**

As outlined in the Introduction, the prediction is that the expected cline width $\langle w_{pq}(p) \rangle$ under genetic drift, migration, and selection should be narrower than the cline width under selection and migration only. For a step change, $\langle w_{pq}(p) \rangle$...
clines maintained by step change in selection is proportional to \( w^2/(N\sqrt{s}) \) (dashed line). Unlike for the underdominance, where we get an exact approximation as we know the eigenfunction driving the shift in a position, for step change the proportionality constant (of \( \sim 1 \)) comes from the fit to the simulations. Parameters: \( 2p \rightarrow N = 60, \sigma^2 \rightarrow m = 0.5 \), and for \( A, s = 0.005 \). At time 0, the population is at the deterministic equilibrium.

should be narrower by a factor of \( \sqrt{1-(\bar{F})} \), whereas the width of the expected cline \( w_{pq}(\bar{p}) \) increases with drift by a factor of \( 1/\sqrt{1-(\bar{F})} \) compared to \( w_{pq}(p_0) \). When the cline is maintained by frequency-dependent selection or hybrid incompatibility (underdominance), the expected cline has a cline like-equilibrium only after the clines have been aligned in space. We estimate the position of a cline center as \( n_d = \sum_{i=1}^{N} p_i + 0.5 \), where \( n_d \) is the number of demes. This gives a consistent estimate of a position even for highly distorted clines, unlike the MLE (Appendix A). The expectation of the width of a cline after alignment, \( \langle w_{pq}(\bar{p}) \rangle \), should be narrower by about a factor of \( 1-(\bar{F}) \) than the width of the expectation of aligned clines \( w_{pq}(\bar{p}_a) \). The simulations show that \( w_{pq}(\bar{p}_a) \) is close to the width without drift, \( w_{pq}(p_0) \). Figure 4 shows a good agreement between the predictions and simulations.

Hallatschek and Korolev (2009, p. 108103-2) showed that even with no selection, cline width stays finite due to genetic drift—as long as variance is maintained across the range. (Imagine a secondary contact of broad-ranging neutral polymorphisms.) For neutral loci, \( \langle w_{pq} \rangle \) is slowly approaching \( 8\sigma^2p \) as time goes to infinity (Hallatschek and Korolev 2009, p. 108103-2)—see Supporting Information, Figure S1. This cline width is still wider than the range of our Figure 4 with selection and drift. Note that, in a finite habitat, neutral diversity will be eventually lost, not by flattening of the cline but rather as the neutral cline wanders to the edge of the habitat.

The expected standardized variance \( \langle F \rangle \) is obtained directly from the simulations. However, for step change in selection, one can get an exact expression for the standardized variance at the edge of the cline (Luskin and Nagylaki 1979): \( F_{\text{u}} = 1/(\rho \sqrt{6}s) = 2\sqrt{\pi}/(N\sqrt{s}) \). Figure S2 shows that the agreement between the two is very good. For weak selection, one can obtain an approximation for equilibrium standardized variance \( F \) comparing the linear terms driving the growth of a perturbation of a model with selection to a neutral model where variance is maintained by either mutation (Wright 1943; Malécot 1969) or long-distance migration (Kimura and Weiss 1964) (see Felsenstein 1975).

Still, the effect of random genetic drift on the expected cline shape is small unless selection is very weak and/or neighborhood size is small (say \( 1/(N\sqrt{s}) > 0.1 \)). However, individual clines can still differ substantially from the dotted line shows the estimate using the standardized variance \( F_{\text{u}} = 1/(\rho \sigma \sqrt{6}s) \) at the edge of the cline. The agreement with the expected \( \langle F \rangle \) is close (see Figure S2), making the predictions with \( F_{\text{u}} \) and \( \langle F \rangle \) indistinguishable. With underdominance (A), the position of the cline diverges. After spatial alignment of the clines, the mean cline width \( \langle \rangle \) is narrower by a factor of \( 1-(\bar{F}) \) than the width of the mean (aligned) cline \( \langle \rangle \) and the latter is close to cline width without drift, \( w_{pq}(\bar{p}_a) \sim w_{pq}(p_0) \). Parameter \( F_{\text{u}} \) gives the standardized variance in allele frequencies after spatial alignment. The selection coefficient \( s \) ranges from 0.001 to 0.2, haploid deme size is \( N = 60 \), and the number of replicates is 500. Generation time is 1000; the equilibrium in the extent of fluctuations is reached by generation 1000 even for \( s = 0.001 \) (visual check), at time zero the cline is at deterministic equilibrium.
expectation: variance in the cline shape due to genetic drift is large (see Figure 2) even when drift is not particularly strong. The variance in cline width \( w_{pq} \) rises with the strength of drift approximately as \( \text{var}(w_{pq}) = 4.21(\sigma/(2\sqrt{2}s^2)) = 4.21(\sqrt{2}\pi F/(N\sqrt{s})) \) (see Appendix C and Figure S3). Importantly, the standard deviation in cline width is substantial even though the effect on the expected cline width is small (Figure 5). The fit of the prediction for the variance in cline width can be seen in Figure 5.

**Discussion**

Random genetic drift shifts clines from side to side, alters their width, and distorts their shape. Such random fluctuations complicate inferences from cline width and position but may themselves be used to infer rates of gene flow, drift, and selection. We explain and quantify the effects of drift on a cline shape and position for (i) tension zones represented by underdominance and (ii) spatially dependent selection represented by a step change in selection across a habitat. Notably, as a cline wobbles about its position, the expected cline widens, whereas the expected width of a cline decreases. Individual clines become narrower because locally, drift divests allele frequencies to fixation. Our predictions apply directly to organisms that live in a habitat that is adequately described by one dimension (like a river, a deep valley, or a coast line) and give a guide to what patterns to expect in two dimensions.

With a step change in selection, while random drift broadens the overall cline in expected allele frequency by a factor \( 1/\sqrt{1-F_s} \), it is expected to narrow the width of any particular cline by the inverse of this factor. With underdominance, the expected cline width is narrower due to drift by about \( 1-F_s \), relative to the width in the absence of drift. Note that by “cline width” we mean the sum of heterozygote frequencies across space, \( w_{pq} = 4\sum p_i q_i \delta x_i \), where \( \delta x_i \) is the distance between sampling sites \( i \) along a cline. The parameter \( F_s \) is the standardized variance of fluctuations in the cline in one-dimensional space, averaged across the habitat weighted by the product of expected allele frequencies, \( p \). \( F_s \) is calculated after aligning the clines in space. The standardized variance \( \text{var}(p) \) of allele frequencies, \( F_s = \text{var}(p)/(\langle p \rangle(1-\langle p \rangle)) \) varies across space, and in general depends on a complicated way on selection, drift, and gene flow. For a step change in selection, \( F_s \) is close to \( 2\sqrt{\pi/6}/(N\sqrt{s}) \) (Nagyalki 1978), where \( N \) is the neighborhood size. \( F_s \) differs from what is typically denoted by \( F_{ST} \) along a cline, as \( F_s \) includes fluctuations in the position of the cline, whereas for \( F_{ST} \), clines are typically aligned first.

Still, it should be stressed that the expected shape of the cline changes appreciably only if drift is strong: i.e., \( 1/(N\sqrt{s}) \) is large. In this study, we kept neighborhood size \( N \) quite low, mostly to 75, and varied the strength of selection \( s \). Similarly small neighborhood sizes are often found in plants (Crawford 1984; Rieseberg et al. 2004) and in some reptiles; commonly studied vertebrates and insects tend to have generally much larger neighborhood sizes (see Wright 1978, pp. 61–76). However, as genetic drift weakens, the standard deviation in cline width decreases much slower than the expected change in cline width. The variance in cline width for the model of underdominance/frequency-dependent selection rises with drift as \( \sim 10.6 \cdot F_s^2/(N\sqrt{s}) \). Even when genetic drift is not particularly strong, the width and position of individual clines can differ considerably from the expectation, potentially biasing our inference about selection and gene flow.

In two dimensions, the effect of drift, i.e., the inverse of the neighborhood size, depends only weakly on selection, unlike in one dimension. Dimensional arguments and linear analysis of weak fluctuations show that when selection depends on the habitat, \( F_s \) is close to that for neutral alleles in two dimensions (Barton et al. 2002) and in the island model (Whitlock 2002). Thus, \( F_s \) can be predicted from the classical theory and estimated from neutral markers, giving us a practical way to find the effect of drift on the width of selected clines. This argument is similar to Whitlock’s (2003) analysis of fixation probability, where he shows that for a wide range of population structures, selection and hence fixation probability are in effect reduced by a factor \( (1-F_s) \), \( F_s \) being close to its neutral value.

The effect of genetic drift can be strong in mosaic hybrid zones: patchy interfaces between genetically distinct populations. Mosaic hybrid zones have received much attention in recent years (Klingenberg et al. 2000; Ross and Harrison 2002; Bierne et al. 2003; Vines et al. 2003). In some cases, local patches are correlated with local habitat, implying that the mosaic structure is maintained by strong and spatially heterogeneous selection (Ross and Harrison 2002). In other cases, as when the habitat is patchily distributed but the selective pressure is intrinsic, random drift provides a more parsimonious explanation: one of the most robust results in population genetics is that selection cannot maintain spatial structure over scales shorter than \( \sim \alpha/\sqrt{2s} \) (Slatkin 1973). Mosaic hybrid zones that are caused by strong random drift
provide a direct qualitative illustration of our central point, that while drift broadens the overall cline, it narrows local interfaces (possibly as in *Solenopsis* (Shoemaker et al. 1996)).

So far we have focused on the analysis of a single locus. Often, clines in different loci cluster together; such hybrid zones may be produced by secondary contact, by a common response to an environmental gradient, and by the accumulation of incompatibilities in parapatry. Within such hybrid zones, strong linkage disequilibria can be generated, with an excess of combinations of genes that flow in from either side. Thus, the effective selection on each locus is augmented by selection on all the others, steepening the clines and reducing gene flow (Barton and Hewitt 1985). It is not clear how random drift will affect the width of such a complex hybrid zone: on the one hand, fluctuations at different loci will be correlated as a result of linkage disequilibria, increasing their effect, but on the other hand, strong effective selection will dampen fluctuations.

The clines in a hybrid zone may differ in position and width because they respond differently to the environment and because epistasis may favor some recombinants over others. Indeed, if the ancestral genotypes are fitter than those genotypes produced by random hybridization, we expect them to increase within the hybrid zone, causing clines in different loci to move apart: in effect, reconstructing the sequence of fit ancestral genotypes through which the populations diverged (Searle 1986; Virdee and Hewitt 1992). However, position and width may also differ by chance, and so we need to test whether any differences are significantly greater than expected under the null hypothesis of symmetrical selection and random drift.

To fit cline shape and position consistently for an unknown selection regime, a “concordance” method was developed by Szymura and Barton (1986) and used to analyze the hybrid zones in *Bombina* and, most recently, used for the hybrid zone in *Mus* by Macholán et al. (2011). This method is similar to the one in Gomper and Buerkle (2009), which uses diploid genotype frequencies. In concordance analysis, the allele frequency of each locus is fitted to a cubic polynomial model

\[ p = \bar{p} + \bar{q} \alpha + \beta (\bar{p} - \bar{q}), \]

where approximately, the coefficient \( \alpha \) is twice the shift in the position of the cline and \( \beta \) gives the decrease in cline width, both taken relative to the average. Residual variance gives an estimate of the net effects of both sampling error and the true variance in allele frequency, represented by \( F_{ST} \) (assumed constant across the cline). The concordance method, however, does not account for correlations in fluctuations between nearby locations and so underestimates the cumulative effects of drift over many generations. With support from simulations, the methods described could be developed to make a proper test of the cause of discordance in shape and position between clines, as we intend for the future.

Another complication arising when studying natural populations is the sampling of the cline by the researcher: insufficient sampling, especially at the central part of the cline, can bias the inference. This issue has been addressed by Dušková et al. (2011). Here, we assume the sampling is sufficient not to introduce further bias.

In general, random drift narrows individual clines, reducing net heterozygosity, and so impeding the response to selection. Drift can be especially important at the edge of a species’ range, where gene flow may prevent local adaptation, causing a collapse of the marginal population. Deterministic models in which the genetic variance is free to evolve show that, surprisingly, populations can adapt to arbitrarily steep gradients in environment by inflating their genetic variance (Barton 2001). Locally, random drift reduces genetic variance and may lead to a sharp threshold gradient beyond which adaptation fails, as is seen in deterministic models with fixed variance (Kirkpatrick and Barton 1997). The results here provide a way toward analyzing this problem.

**Acknowledgments**

We thank Simon Aeschbacher, Ferran Palero, and Roger Butlin and his group for their comments on the results and this manuscript. N.B. is supported by the Energy Regulatory Commission. Advanced Grant “Information and Evolution” 250152 and the Engineering and Physical Sciences Research Council.

**Literature Cited**


Appendix A: MLE for Cline Width and Position for Underdominance

Maximum-likelihood estimation of the shape of the cline, including the position ($x_0$), for frequency-dependent selection $s(p - q)pq$ and underdominance (to the first order of $s$), where the cline shape at deterministic equilibrium is $p_0 = 1/(1 + \exp[-4(x - x_0)/\omega])$.

The log probability we would observe $p(x, t)$ given $p_0(x)$, given that we assume that the deviations from the deterministic equilibrium $p(x, t) - p_0(x)$, is drawn from binomial distribution, taking locations as independent. This means we assume that the variance of $p(x, t) - p_0(x)$ is taken as proportional to $p_0(x)q_0(x)$. This is of course not strictly correct, but gives a good approximation, as the standardized variance of allele frequencies, $F = \text{var}(p(x, t))/\langle p(x, t)q(x, t) \rangle$, and “normalized” variance, where $\langle p(x, t) \rangle$ is replaced by $p_0(x)$, vary little across space.

The logistic fit then requires maximization of

$$-\int (p \log(p_0) + q \log(q_0)) dx = \int (\beta x - \log(\gamma)) q + \log(1 + \gamma \exp[-\beta x]) dx,$$

(A1)

where $\beta = 4/w_{\text{MLE}}, w_{\text{MLE}} = (\partial p_{\text{MLE}}/\partial x)^{-1}$ at $p = 1/\gamma$, and $\gamma = \exp[\beta x_0]$.

We define the position of the cline as $\int_{-\infty}^{x} p dx = \hat{x}$. It is convenient to express the integral as $\int_{-\infty}^{x} p dx - \int_{-\infty}^{x} q dx$ to avoid infinities. The maximum-likelihood estimate for the width of the cline is $w_{\text{MLE}} = 4/\beta = 4/\pi\sqrt{6A - 3B^2 + 3\hat{x}^2}$; the position of the cline $x_0 = B$, where $A = \int_{-\infty}^{x} q dx - \int_{-\infty}^{x} p dx$ and $B = x^* - \int_{-\infty}^{x} q dx - \int_{-\infty}^{x} p dx$.

Appendix B: Variance in Cline Position—Underdominance

We consider a model of a linear frequency-dependent selection, where the fitness of the allele with frequency $p$ is $1 - qs$ (and vice versa), as would be the marginal fitness with “underdominance” of selection strength $s$ in a diploid model. To the first order of $s$, allele frequency $p = p(x, t)$ changes according to
where \( \sigma \) is the standard deviation of dispersal distance, and \( x \) denotes space, \( t \) denotes time, and \( \zeta(x, t) \) represents random genetic drift. As usual, \( q(x, t) = 1 - p(x, t) \). In the following, we drop the \( (x, t) \) unless we want to specifically highlight the dependency.

The local change in the allele frequency away from the equilibrium without genetic drift \( (p_0) \) is \( z(x, t) = p(x, t) - p_0(x) \) and \( z(x, t) = (\partial p_0(x)/\partial x) \Delta x \), where \( \partial p/\partial x = pq/2 \) and \( l = \sigma/\sqrt{2\pi} \) is the characteristic length. Fluctuations due to shifts in the position are given by \( \partial z = e_0(x) \Delta \zeta \). Here \( e_0(x) = \sqrt{6/\pi} \cdot p_0(x) q_0(x) \) is the eigenfunction describing the growth of a perturbation to the equilibrium solution due to shifts in the position and is again proportional to the gradient in genotype frequencies (see Barton 1979, pp. 349–350). Taking the square and the expectation of \( \tilde{\zeta}(t) \tilde{\zeta}(x) = (\partial p/\partial x) \Delta x \) and assuming that relative to the variance in position, \( p(x, t) q(x, t) \) is close to \( p_0(x) q_0(x) \), the variance in position is \( \text{var}(\Delta x) = \Delta x^2 = 6l(\Delta \zeta_0(t)) \), where the transformed variance due to random genetic drift \( \tilde{\Delta} \zeta_0(t) = t \int_{-\infty}^{\infty} \left( p(x) q(x) / 2p \right) e_0^*(x)^2 \). The complete eigenfunction expansion is explained in File S1. Assuming the variance in position is much bigger than changes in the expected cline shape, the variance in position of a cline maintained by frequency-dependent selection (with allelic fitness the same as the marginal finesses with underdominance) should be close to \( \text{var}(\Delta x) = (6/5) \cdot \lambda \text{ft} / (2p) \) — as illustrated in Figure 3A.

### Appendix C: Variance in Cline Width: Underdominance

The change in cline width \( w_{pq} = 4 \int_{-\infty}^{\infty} p(x) q(x) dx \) due to fluctuations \( z(x, t) \) away from the equilibrium allele frequency in the absence of genetic drift, \( p_0(x) \), is \( \Delta w_{pq}(t) = -4 \int_{-\infty}^{\infty} p_0(x) - q_0(x) \Delta x dx \). Assuming that equilibrium allele frequency is close to the deterministic solution, the perturbation of allele frequency due to genetic drift can be decomposed as \( z(x, t) = \sum_{i,j} \lambda_i \lambda_j \xi_i(x) \xi_j(x) \), where \( \xi_i(x) \) are the eigenfunctions of the linear operator for the diffusion equation with selection (see File S1). As the variance in cline width is much larger than the change to the mean shape of the cline, variance in the cline width can be taken as the expected squared deviation for the equilibrium, \( \text{var}(w_{pq}) = \int_0^\infty \Delta w_{pq}^2 \) :

\[
\text{var}(w_{pq}) = \text{var}(w_{pq}|e_1^2) + 2 \text{cov}(w_{pq}|e_1^2) + \text{var}(w_{pq}|e_2^2) = A_1(x)^2 \langle \tilde{\zeta}_1 \rangle^2 + 2A_1(x) \int_{-\infty}^{\infty} A_n(x) \langle \tilde{\zeta}_1 \tilde{\zeta}_n \rangle dx + \int_{-\infty}^{\infty} A_n(x) A_n^*(y) \langle \tilde{\zeta}_n \tilde{\zeta}_n \rangle dy db.
\]

Above, the first term gives the contribution to the variance in cline width \( w_{pq} \) due to the second eigenfunction, \( e_1 = \sqrt{3/4} (\beta q_0(x) - p_0(x) q_0(x)) \). The last term gives the contribution to the variance due to all other components, and the middle term gives the covariance between them. \( A_1(x) = \int_{-\infty}^{\infty} p_0(x) - q_0(x) \xi_1(x) dx \) and \( \langle \tilde{\zeta}_1 \tilde{\zeta}_n \rangle \) are the transformed components of the fluctuations, \( \langle \tilde{\zeta}_n \tilde{\zeta}_n \rangle = -(1 / 2p(\lambda_i + \lambda_j)) \int_{-\infty}^{\infty} p_0(x) (1 - p_0(x)) \xi_i(x) \xi_j(x) dx \) as time goes to infinity (see File S1). The \( \lambda_i \)'s are the eigenvalues of the system, \( \lambda_1 = -3/4, \lambda_2 = -(1 + \alpha^2) \), and \( \xi_1(x) = \sqrt{3\pi}(p_0 - q_0) / (1 + \alpha^2) \exp(i\alpha x) \). The components of the fluctuation \( \zeta(x, t) \) due to the first eigenfunction \( e_1 \) can be dropped as the width \( w_{pq} \) is independent of the position of the cline.

In total, the variance in cline width rises with the strength of drift approximately as \( \text{var}(w_{pq}) \approx 4.21 (\sigma/2\sqrt{2} \sqrt{3/2}) = 4.21 (\sqrt{2\pi} \sqrt{\ell^2} / (N \sqrt{\sigma})) \). The estimate based solely on the second eigenfunction is nearly twice as large, \( \text{var}(w_{pq}|e_2^2) = 4\pi^2/5 (\sqrt{2\pi} \sqrt{\ell^2} / (N \sqrt{\sigma})) = 8 (\sqrt{2\pi} \sqrt{\ell^2} / (N \sqrt{\sigma})) \) as the covariance with the other components of the fluctuation is negative, \( \text{cov}(w_{pq}|e_1, e_2) = -4.71 (\sqrt{2\pi} \sqrt{\ell^2} / (N \sqrt{\sigma})) \).

All variance components are shown in Figure S3.
Supporting Information
http://www.genetics.org/content/suppl/2011/06/24/genetics.111.129817.DC1

Genetic Drift Widens the Expected Cline but Narrows the Expected Cline Width

Jitka Polechová and Nick Barton
Figure S1: Even with no selection, the cline width $w_{pq}$ stays finite, $\langle w_{pq} \rangle \to 8\sigma^2 \rho$ (black dot) as time $t$ goes to infinity. The black dotted line shows the prediction, $\langle w_{pq} \rangle = 8\sigma^2 \rho \left(1 - \exp\left(-\frac{1}{4\sigma^2 \rho^2}\right) \operatorname{erfc}\left(\frac{4\rho}{4\sigma^2}\right)\right)$, where $\operatorname{erfc}$ is the complementary error function. The replicates are shown in coloured lines, their average cline width $w_{pq}$ in a black dashed line. The formulae are due to Hallatschek and Korolev (2009, p. 2, where $I(t) \equiv \langle w_{pq} \rangle / 2$). At time zero, we assume the beginning of a secondary contact – i.e. the allele frequency changes abruptly from 0 to 1 in the middle of the habitat. Over the time of 10 000 generations, the cline moves about in a random walk but does not get to the margins of the habitat range of 500 demes. Parameters: migration rate $\sigma^2 \to m = 1/2$, haploid deme size $2\rho \to N = 30$, deme spacing $\delta x = 1$. 
Figure S2: The standardized variance at the margin of the cline, and the expected one over the range, are close (notice the logarithmic scale). The dotted line shows the $F_{\infty} = 1/(\rho \sigma \sqrt{6s})$, as derived by Luskin and Nagylaki (1979). The dots are the measured expected $\langle F \rangle$ for the model with step change. Selection coefficient $s$ ranges from 0.001 to 0.2, haploid deme size is $N = 60$; number of replicates is 100. Generations time is 1000; the equilibrium in the extent of fluctuations is reached by generation 1000 even for $s = 0.001$ (visual check); at time zero the cline is at deterministic equilibrium.
Figure S3: A: Variance in cline width (●) increases with the strength of drift as $w_{pq} = 4.21 \frac{\sqrt{2\pi l^2}}{N\sqrt{s}} = 4.21 \frac{\sqrt{m}}{2p\sqrt{2s^{3/2}}}$ (middle dashed line). The prediction is a sum of components of variance in cline width due to the second eigenfunction, representing changes in the slope of the cline (○, top dashed line: $w_{pq|e_1} = 4\pi^2/3 \frac{\sqrt{2\pi l^2}}{N\sqrt{s}}$), variance due to fluctuations at higher frequencies (□, bottom dashed line: $w_{pq|e_\alpha} = 1.02 \frac{\sqrt{2\pi l^2}}{N\sqrt{s}}$) and (B:) negative covariance between the component due to second eigenfunction and higher order terms, $w_{pq|e_1, e_\alpha} = 4.71 \frac{\sqrt{2\pi l^2}}{N\sqrt{s}}$. Parameters as for Figure 5.
SUPPORTING INFORMATION

Eigenfunction expansion for growth of a perturbation to a cline

\[
\frac{\partial p(x, t)}{\partial t} = \frac{\sigma^2}{2} \frac{\partial^2 p(x, t)}{\partial x^2} + s(p, x)p(x, t)q(x, t) + \zeta(x, t) \quad (S1)
\]

where \( \sigma \) is the standard deviation of dispersal distance, and the selective advantage \( s(p, x) \) (to the first order) of an allele varies across space and may depend on allele frequency \( p \). Whereas one cannot solve the spatial diffusion equation with selection and drift, because the fluctuations due to genetic drift \( \zeta(x, t) \) depend in general on both time and space, the equation for expected change in the allele frequency in a one dimensional habitat can be solved analytically when selection depends only on location, \( s(x) \).

We consider a model of a linear frequency-dependent selection, where the fitness of the allele with frequency \( p \) is \( 1 - q s \) (and vice versa) – as would be the marginal fitness with “underdominance” of selection strength \( s \) in a diploid model. To the first order of \( s \), allele frequency \( p \) changes according to

\[
\frac{\partial p}{\partial t} = \frac{\sigma^2}{2} \frac{\partial^2 p}{\partial x^2} + spq(p - q) \quad (S2)
\]

where \( \sigma \) is the standard deviation of dispersal distance, and \( x \) denotes space, \( t \) time. As usual, \( q(x, t) = 1 - p(x, t) \) - and we drop the \((x, t)\) unless we want to specifically highlight the dependency.

Scaling time with selection \( T = st \) and distance relative to the dispersal, \( X = x \sqrt{\frac{2s}{\sigma^2}} \) gives

\[
\frac{\partial p}{\partial T} = \frac{\partial^2 p}{\partial X^2} + pq(p - q). \quad \text{The equilibrium solution is } p_0 = \frac{1}{1 + Exp[-(X - X_0)]} \quad (\text{and the mirror image, } 1 - p_0). \quad \text{The position of the cline, } X_0, \text{ is arbitrary.}
\]

Including random genetic drift \( \epsilon(X, T) \), a small perturbation to the equilibrium of Eq. S2, \( Z \), grows (to the first order) as

\[
\frac{\partial Z}{\partial T} = \frac{\partial^2 Z}{\partial X^2} - (1 - 6p_0q_0)Z + \epsilon(X, T) \quad (S3)
\]

We look for a solution of the equation above in a form of

\[
Z(X, T) = \sum_{i=1}^{\infty} \tilde{Z}_i(T)e_i(X)
\]

- i.e. decomposing \( Z \) with its time-dependent components \( \tilde{Z}_i(T) \) and space-dependent
components \(e_i(X)\). We choose the \(e_i(X)\) so that they are the eigenfunctions of the corresponding homogeneous system (with \(\epsilon(X, T) = 0\)): \(\frac{\partial^2}{\partial X^2} - (1 - 6p_0q_0)\) - hence, \(Le_i = \lambda_i e_i\), where \(\lambda_i\) are eigenvalues and \(e_i\) the eigenfunctions of \(L\).

Normalizing so that \(\int e_i(X)e_j^*(X)dX = \delta_{ij}\), the eigenfunctions and respective eigenvalues are:

\[
e_0(X) = \sqrt{6p_0q_0} \quad \lambda_0 = 0 \quad (S4)
\]
\[
e_1(X) = (p_0 - q_0)\sqrt{6p_0q_0} \quad \lambda_1 = -3/4 \quad (S5)
\]
\[
e(\alpha, X) = \frac{3\alpha i(p_0 - q_0) + 2\alpha^2 - (1 - 6p_0q_0)}{\sqrt{2\pi(1 + \alpha^2)(1 + 4\alpha^2)}} \exp(\alpha X) \quad \lambda(\alpha) = -1 - \alpha^2 \quad (S6)
\]

Apart from the first two, the eigenfunctions form a continuum along \(\alpha\). The transformed perturbation components are: \(\tilde{Z}_0(T) = \int_{-\infty}^{\infty} Z(X, T)e_0^*(X)dX\), \(\tilde{Z}_1(T) = \int_{-\infty}^{\infty} Z(X, T)e_1^*(X)dX\), \(\tilde{\epsilon}(T) = \int_{-\infty}^{\infty} \epsilon(\alpha, X)dX\).

Substituting to the above equation (Eq. S3) the eigenfunction transformation of \(Z, \epsilon\) and using \(Le_i = \lambda_i e_i\), gives (as \(\tilde{Z}\) is constant with respect to \(L\)):

\[
\frac{\partial}{\partial T} (\tilde{Z}_0(T)e_0(X) + \tilde{Z}_1(T)e_1(X) + \int_{-\infty}^{\infty} \tilde{Z}(\alpha, T)e(\alpha, X)d\alpha) = \lambda_0 \tilde{Z}_0(T)e_0(X) + \lambda_1 \tilde{Z}_1(T)e_1(X) + \int_{-\infty}^{\infty} \lambda(\alpha) \tilde{Z}(\alpha, T)e(\alpha, X)d\alpha + \tilde{\epsilon}_0(T)e_0(X) + \tilde{\epsilon}_1(T)e_1(X) + \int_{-\infty}^{\infty} \tilde{\epsilon}(\alpha, T)e(\alpha, X)d\alpha \quad (S7)
\]

As the eigenfunctions are orthogonal, we get

\[
\frac{\partial}{\partial T} (\tilde{Z}_0(T) + \tilde{Z}_1(T) + \int_{-\infty}^{\infty} \tilde{Z}(\alpha, T)d\alpha) = \lambda_0 \tilde{Z}_0(T) + \lambda_1 \tilde{Z}_1(T) + \int_{-\infty}^{\infty} \lambda(\alpha) \tilde{Z}(\alpha, T)d\alpha + \tilde{\epsilon}_0(T) + \tilde{\epsilon}_1(T) + \int_{-\infty}^{\infty} \tilde{\epsilon}(\alpha, T)d\alpha \quad (S8)
\]

There is a solution where all of the components with \(i = 0, 1\) and all values of \(\alpha\) sum to zero, i.e. \(\frac{\partial}{\partial T} \tilde{Z}_i(T) = \lambda_0 \tilde{Z}_i(T) + \tilde{\epsilon}_i(T)\) for all \(i\). Multiplying by \(\exp(-\lambda_iT)\), using \(\frac{\partial}{\partial T} (\tilde{Z}_i(T)\exp(-\lambda_iT)) = \exp(-\lambda_iT)(\frac{\partial}{\partial T} \tilde{Z}_i(T) - \lambda_i \tilde{Z}_i(T))\) and integrating \(\int_{-\infty}^{T} d\tau\) (after \(T \to \tau\)) gives the time-dependent components of perturbations:

\(\tilde{Z}_i(T) = \int_{-\infty}^{T} \tilde{\epsilon}_i(\tau) \exp(-\lambda_i(\tau - T))d\tau\).
And so for the covariance in the transformed fluctuations in allele frequencies,

$$\langle \tilde{Z}_i(T) \tilde{Z}_j(T') \rangle = \int_{-\infty}^{T} \int_{-\infty}^{T'} \langle \bar{e}_i(\tau) \bar{e}_j(\tau') \rangle \text{Exp}(-\lambda_i(\tau - T) - \lambda_j(\tau' - T')) d\tau d\tau' \quad (S9)$$

Where $$\langle \bar{e}_i(\tau) \bar{e}_j(\tau') \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle \epsilon(X, \tau) \epsilon(X', \tau') \rangle e_i^*(X) e_j^*(X') dX dX'$$ are the transformed noise components. We assume each deme is sampled independently, hence the fluctuation caused by random genetic drift can be described by white noise, uncorrelated in space and time: $$\langle \epsilon(X, \tau) \epsilon(X', \tau') \rangle = \frac{\langle p(X)(1-p)(X) \rangle}{2p} \delta(X-X') \delta(\tau-\tau').$$ Hence, at a given time $$T = T$$, the covariance between the transformed components (with at least one $$\lambda_i \neq 0$$) is

$$\langle \tilde{Z}_i(T) \tilde{Z}_j(T) \rangle = \frac{1 - \text{Exp}(T(\lambda_i + \lambda_j))}{2\rho(\lambda_i + \lambda_j)} \int_{-\infty}^{\infty} \langle p(X)(1-p)(X) \rangle e_i^*(X) e_j^*(X) dX \quad (S10)$$

For eigenvalues $$\lambda_i < 0$$, the exponential term goes to zero. For the variance in fluctuations due to the shifts in position, where $$\lambda_0 = 0$$, $$\langle \tilde{Z}_0(T)^2 \rangle = \frac{T}{2p} \int_{-\infty}^{\infty} \langle p(X)(1-p)(X) \rangle e_0^*(X)^2 dX$$ from Eq. 9 and below, integrating from 0 to $$T$$.

Transforming back, and denoting $$\Lambda_{i,j} = \int_{-\infty}^{\infty} \langle p(X)(1-p)(X) \rangle e_i^*(X) e_j^*(X) dX$$ gives the covariance in the fluctuations for the spatially aligned clines. (i.e., discounting the fluctuations in position, $$\langle Z_0(X, T) Z_0(X', T) \rangle$$). The covariance in fluctuations between position and width, $$\Lambda_{0,1}$$, is zero.

$$\langle Z(X, T) Z(X', T) \rangle = -\frac{1}{2p} \left( \frac{\Lambda_{1,1}}{2\lambda_1} e_1(X) e_1(X') + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\Lambda_{1,0'}}{\lambda_0 + \lambda(\alpha)} e_0(X) e(\alpha, X) e(\alpha', X') d\alpha d\alpha' + \int_{-\infty}^{\infty} \frac{\Lambda_{0,0'}}{\lambda_0 + \lambda(\alpha')} e_0(X') e(\alpha', X) d\alpha' + \int_{-\infty}^{\infty} \frac{\Lambda_{1,0}}{\lambda_1 + \lambda(\alpha)} e_1(X) e(\alpha', X') d\alpha + \int_{-\infty}^{\infty} \frac{\Lambda_{0,1}}{\lambda_1 + \lambda(\alpha')} e_1(X') e(\alpha', X) d\alpha \right)$$

Scalining back to the original variables, and normalizing with average local allele frequency, we obtain standardized variance across space for aligned clines, $$F_A$$:

$$F_A = \frac{\sqrt{\pi/2}}{N\sqrt{s}} \cdot \frac{\langle Z(X, T) Z(X, T) \rangle}{pq} \quad (S11)$$

The fluctuations scale proportionally to $$\frac{1}{N\sqrt{s}}$$, where $$N$$ is the neighbourhood size in one dimension, and $$\frac{\sqrt{\pi/2}}{N\sqrt{s}} = 4\rho ls$$, which is the “$$\beta$$” of Nagylaki (1978).