doi:10.1111/evo.12532

# The temporal distribution of directional gradients under selection for an optimum

**Luis-Miguel Chevin[1,2] and Benjamin C. Haller[1]**

[1] *CEFE–CNRS, UMR 5175, 34293, Montpellier, France*

[2] *E-mail: luis-miguel.chevin@cefe.cnrs.fr*

Received May 21, 2014
Accepted September 18, 2014

**Temporal variation in phenotypic selection is often attributed to environmental change causing movements of the adaptive surface relating traits to fitness, but this connection is rarely established empirically. Fluctuating phenotypic selection can be measured by the variance and autocorrelation of directional selection gradients through time. However, the dynamics of these gradients depend not only on environmental changes altering the fitness surface, but also on evolution of the phenotypic distribution. Therefore, it is unclear to what extent variability in selection gradients can inform us about the underlying drivers of their fluctuations. To investigate this question, we derive the temporal distribution of directional gradients under selection for a phenotypic optimum that is either constant or fluctuates randomly in various ways in a finite population. Our analytical results, combined with population- and individual-based simulations, show that although some characteristic patterns can be distinguished, very different types of change in the optimum (including a constant optimum) can generate similar temporal distributions of selection gradients, making it difficult to infer the processes underlying apparent fluctuating selection. Analyzing changes in phenotype distributions together with changes in selection gradients should prove more useful for inferring the mechanisms underlying estimated fluctuating selection.**

**KEY WORDS:** **Adaptation, contemporary evolution, fluctuating selection, quantitative genetics, time-series analysis.**

Temporally variable natural selection is a topic of central importance in evolutionary genetics (reviewed by Felsenstein 1974; Hedrick et al. 1976; Hedrick 2006; Bell 2010). Its potential to maintain genetic polymorphism has been thoroughly investigated: Although early model concluded that no stable polymorphism exists in the simplest case of haploid density-independent (and frequency-independent) selection (Dempster 1955), further work showed that polymorphism can be maintained in more complex models that allow for genetic variation in the form of the density dependence function ("relative nonlinearity," Chesson 2000), stochastic (marginal) overdominance in diploids (Haldane and Jayakar 1963; Gillespie 1991), or "storage" effects in overlapping generations (Ellner and Hairston 1994; Svardal et al. 2011). Fluctuating selection also can drive the evolution of mechanisms that generate phenotypic variation, such as bet hedging (Gillespie 1974; Slatkin 1974; Bull 1987; Svardal et al. 2011), phenotypic plasticity (e.g., Gavrilets and Scheiner 1993; Lande 2009), or their combination (Scheiner 2014; J. Tufto, unpubl. ms.). And because individual fitness generally emerges from demographic vital rates (survival and fecundity), fluctuating selection has a close connection to environmental stochasticity in population growth rates (Lande 2007; Engen et al. 2011; Engen and Saether 2014), which is itself a strong determinant of extinction risk (Lande 1993; Lande et al. 2003). On the more technical side, stationary fluctuating selection may also explain why phenotypic evolution often involves rapid changes over a few generations, but stasis over geological time (Simpson 1944; Eldredge and Gould 1972; Gingerich 2001; Kinnison and Hendry 2001; Estes and Arnold 2007; Uyeda et al. 2011).

Several classic examples of natural selection in the wild exhibit temporal variation, from beak shape in Darwin's finches (Grant and Grant 2002) to banding patterns in Cepea snails (Cain et al. 1990), or spine number in three-spined sticklebacks (Reimchen and Nosil 2002). This variation is generally thought to be caused by a variable environment, although this connection is seldom demonstrated and quantified (Wade and Kalisz 1990; MacColl 2011). In some instances, selection varies in a largely predictable way, for instance because of seasonal environmental






EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

change (Kingsolver 1995; Benkman and Miller 1996); but it also often appears random (Fisher and Ford 1947; Cain et al. 1990; Grant and Grant 2002; reviewed in Bell 2010). In the latter case, the strength of selection can be modeled as a stochastic process, and detected as such.

For quantitative traits with a (nearly) continuous distribution, directional phenotypic selection in a given generation is commonly measured using the directional (or linear) selection gradient $\beta$, the regression slope of relative fitness (individual over mean fitness) on trait value (Lande and Arnold 1983). An attractive feature of $\beta$ is that for a normally distributed trait, it is directly related to the change in the mean phenotype per generation through the equation $\Delta\bar{z} = G\beta$, with $G$ being the trait's additive genetic variance (or covariance matrix, for multiple traits, with $\beta$ a vector) (Lande 1976, 1979). The selection gradient $\beta$ is also the local slope of the surface relating population mean fitness to the mean phenotype, which extends Wright's fitness landscape (Wright 1937) to (normally distributed) quantitative traits (Lande 1976). When natural selection changes in time, it thus seems natural to use the temporal distribution of directional selection gradients to measure its fluctuations (Siepielski et al. 2009; Siepielski et al. 2011; Morrissey and Hadfield 2012). A benefit of this approach is that, under some assumptions about the genetic architecture of the trait, it can be used to predict temporal patterns of phenotypic change without making any assumption about the shape of the fitness surface. However, very different processes may lead to similar temporal distributions of $\beta$, making it difficult to infer the mechanism that is actually driving evolution, as we elaborate below.

Stationary random variation in selection is not only characterized by its variance, but also by its temporal autocorrelation, which determines its predictability over different time scales. Previous theory has made it clear that the autocorrelation of selection strongly affects whether (and how much) genetic responses to selection optimize long-term fitness and population growth in a fluctuating environment (Charlesworth 1993; Lande and Shannon 1996; Bürger and Gimelfarb 2002; Chevin 2013). A genetic response to selection in one generation is more likely to be beneficial in the next if the direction of selection does not change. For a given variance of selection gradients, genetic responses to selection are thus expected to improve long-term fitness when selection gradients are positively autocorrelated across generations (Chevin 2013). When they are negatively autocorrelated, on the other hand, genetic responses are detrimental on average. This means that arguments about improving or maintaining the "adaptive potential" of populations to favor their persistence in a changing environment depend on information about the predictability of selection on traits that strongly affect fitness; high "adaptive potential" could actually be detrimental to a population exposed to largely unpredictable environmental fluctuations.

We here investigate the temporal variance and autocorrelation of selection gradients for a population subject to a fitness function that has a maximum for an optimum value of a trait, under various patterns of movement of this optimum due to a changing environment. Several arguments, from basic functional considerations (Wright 1935; van Asch et al. 2007) to standard properties of dynamical systems (Otto and Day 2007), and analyses of empirical rates of evolution across different timescales (Estes and Arnold 2007; Uyeda et al. 2011), support the idea that many traits are under stabilizing selection for an optimum phenotype (see Charlesworth et al. 1982 for a critical review of such arguments). Direct measurements of selection also commonly find fitness functions with an optimum phenotype (e.g., Schluter and Nychka 1994; Benkman and Miller 1996; Gimenez et al. 2006; Garant et al. 2007; Martin and Wainwright 2013). Although meta-analyses did not find a prevalent role of negative over positive quadratic gradients in the literature (Kingsolver et al. 2001; Kingsolver and Diamond 2011), positive curvature is also detected when the mean phenotype is far from an optimum, and does not necessarily imply disruptive selection (Schluter 1988). Furthermore, fitness functions with a moving optimum have a long history in evolutionary theory, where they are the most common way to model adaptation to a changing environment, in both the population genetic (Fisher 1930; Orr 1998; Kopp and Hermisson 2007; Gordo and Campos 2012) and quantitative genetic traditions (Lynch and Lande 1993; Bürger and Lynch 1995; Gomulkiewicz and Holt 1995; Gomulkiewicz and Houle 2009; Jones et al. 2012; Zhang 2012; Chevin 2013), so it is natural to ask what the distribution of selection gradients would be in this context.

Under stabilizing selection, the temporal variation in selection gradients need not directly reflect the temporal variation of the optimum. The strength and direction of natural selection also depends on the current phenotype distribution, which is the result of all past genetic responses to selection and random genetic drift. This is very clear in the case of a single abrupt shift in the optimum (e.g., Lande 1976; Gomulkiewicz and Holt 1995): Although the optimum does not move in generations following the shift, the magnitude of the selection gradient nevertheless diminishes as the mean phenotype approaches the optimum. More generally, evolution of the mean phenotype needs to be included to interpret temporal variation of selection gradients in terms of temporal changes in the optimum phenotype—the latter being more useful for predicting the long-term evolution and persistence of a population (Lynch and Lande 1993; Bürger and Lynch 1995; Gomulkiewicz and Holt 1995; Lande and Shannon 1996; Gomulkiewicz and Houle 2009; Chevin 2013). By performing such an analysis in the case of randomly changing environments, we show that although some patterns are specific and distinguishable, very different processes may lead to similar temporal distributions of selection gradients, such that further information

would be needed to uncover the underlying cause of measured fluctuations in selection.

# Methods

### MODEL AND ANALYTICAL DERIVATIONS:

We focus on a quantitative trait $z = g + e$, where $g$ is the additive genetic value (for simplicity, we ignore dominance and epistasis effects), and $e$ is a normally distributed residual component of variation, with mean 0, variance $V_e$, and no covariance with $g$ (Falconer and MacKay 1996). We assume that substantial polymorphism across numerous loci causes the additive genetic values to be normally distributed, such that $z$ also has a normal distribution, with mean $\overline{z} = \overline{g}$ and phenotypic variance $P = G + V_e$ ($G$ is the additive genetic variance). Generations are discrete and nonoverlapping, and the fitness of an individual with phenotype $z$ is proportional to $W(z) = \exp\left(-(z-\theta)^2/(2\omega^2)\right)$, where $\theta$ is the optimum phenotype and $\omega$ is the width of the fitness function. We assume that the optimum phenotype may change with the environment, but not the width of the fitness function, in line with most theory of adaptation to a moving optimum (Lynch and Lande 1993; Bürger and Lynch 1995; Gomulkiewicz and Holt 1995; Hansen 1997; Kopp and Hermisson 2007; Jones et al. 2012; Chevin 2013); see Revell (2007) for a study of fluctuating peak width. The height of the fitness peak could also change with the environment but this would not affect phenotypic selection, so we standardize this height to 1 without loss of generality. We investigate several situations that can cause random changes in phenotypic selection: random genetic drift with a constant optimum; a stationary fluctuating optimum without or with autocorrelation (white noise and autoregressive process, respectively); and a nonstationary fluctuating optimum (random walk).

The response to selection in any given generation is $\Delta\overline{z} = G\beta$, where $\beta = \partial \ln \overline{W}/\partial\overline{z} = -S(\overline{z} - \theta)$ is the directional selection gradient, and $S = 1/(\omega^2 + P)$ is the strength of stabilizing selection (Lande 1976). Our analytical results rest on the assumption that the genetic variance $G$ is constant, which is a good approximation under substantial mutation with Gaussian stabilizing selection, even in a changing environment (Bürger 2000, and our simulation results below). All the scenarios we investigate lead to a stationary distribution of selection gradients $\beta$, and we derive the stochastic variance and autocorrelation function (ACF) of these gradients.

### POPULATION-BASED SIMULATIONS

To check the validity of some approximations used to derive the analytical results (notably the "weak selection" assumption required for the continuous-time approximation below), we ran population-based numerical simulations. Such simulations are rapid, and allow running several replicates for each parameter set. In these simulations, we used the recursion equation for evolution of a normally distributed quantitative trait under natural selection and random genetic drift, $\Delta\overline{z} = G\beta + \sqrt{G/N_e}\zeta$, where $\zeta$ is a standard normal random deviate and $N_e$ the effective population size (Lande 1976). We iterated this recursion over 5000 generations, under the required pattern of change in the optimum, and assuming a constant genetic variance $G$. In each generation, we recorded the selection gradient $\beta = -S(\overline{z} - \theta)$. We then computed the variance and ACF of these gradients over the last 4000 generations. For the ACF, we computed the correlation of $\beta$ in a given generation with $\beta$ in 100 future generations, over all generations between 1000 and 4900.

### INDIVIDUAL-BASED SIMULATIONS

Despite their efficiency, our population-based simulations still rest on the somewhat strong assumptions of normality of phenotype and breeding value distributions, constant genetic variance $G$, and constant effective population size $N_e$. These assumptions are necessarily violated in reality, although they may be good approximations in some cases. To investigate the robustness of predictions to these assumptions, we also ran extensive individual-based simulations with explicit loci. The simulation program we used is described in detail in Haller and Hendry (2014), and has basic assumptions similar to the mathematical model above, notably discrete nonoverlapping generations in a single panmictic population subject to stabilizing selection. However, the simulations also allow the population size to change because of variation in mean fitness across generations, with a ceiling type of density regulation (e.g., Lande 1993) at carrying capacity $K = 1000$.

Here, we modified this framework to fit the requirements of the present model. Specifically, we modeled a trait determined by 32 freely recombining additive loci with a continuum of alleles, with mutation rate $\mu = 10^{-5}$, $10^{-4}$, or $10^{-3}$, and standard deviation of mutational phenotypic effects $a = 2$ per locus (the environmental variance was set to $V_e = 1$). The probability that individuals survive to reproduce was determined by a Gaussian fitness function as described above, with width $\omega = 1$, 2, 4, or 8 (larger values cause weaker stabilizing selection). The pattern of fluctuating selection was modified by changing the variance ($\sigma^2_\theta = 0.05$, 0.1, 0.2, or 0.5) and the autocorrelation time ($T = 1$, 4, 15, 50, or 200) of the optimum (the autocorrelation of two optima separated by lag $\tau$ being $\rho(\theta, \tau) = \exp(-|\tau|/T)$, see below). Simulations were run for 15,000 generations (30,000 generations for $T = 200$). For each generation (past a burn-in time of 10,000 generations), directional selection gradients were computed by linear regression of relative fitness on the phenotype (Lande and Arnold 1983; Lande 1993), over the whole population. The temporal variance and autocorrelation of these selection gradients were then computed. To compare these values to the analytical predictions, we computed the mean genetic variance across generations

**Table 1. Summary of the temporal distribution of selection gradients under different patterns of environmental change in the optimum phenotype.**

| Pattern of change in optimum | No drift | | Drift | |
|---|---|---|---|---|
| | Variance V(β) | ACF ρ(β,τ) | Variance V(β) | ACF ρ(β,τ) |
| Constant | *NA* | *NA* | $\frac{S}{(2-SG)N_e}$ | $(1-SG)^{\tau}$ |
| Random walk | $\frac{SB^2}{(2-SG)G}$ | $(1-SG)^{\tau}$ | $S\frac{N_eB^2+G}{(2-SG)N_eG}$ | $(1-SG)^{\tau}$ |
| White noise | $\frac{2S^2\sigma_\theta^2}{2-SG}$ | $-\frac{SG}{2}(1-SG)^{\tau-1}$ | $S\frac{1+2\alpha}{(2-SG)N_e}$ | $\frac{1-\alpha SG(1-SG)^{-1}}{1+2\alpha}(1-SG)^{\tau}$ |
| Autoregressive | $\frac{S^2\sigma_\theta^2}{1+\kappa}$ | $\frac{e^{-\tau/T}-\kappa e^{-SG\tau}}{1-\kappa}$ | $S\frac{2\alpha+\kappa+1}{2N_e(1+\kappa)}$ | $\frac{2\alpha\frac{e^{-\tau/T}-\kappa e^{-SG\tau}}{1-\kappa}+(1+\kappa)e^{-SG\tau}}{2\alpha+\kappa+1}$ |

**ACF, autocorrelation function; $N_e$, effective population size; $G$, additive genetic variance; $S = 1/(\omega^2 + P)$, strength of stabilizing selection, with ω the width of the fitness function and $P$ the phenotypic variance; $\sigma_\theta^2$, variance of fluctuations in the optimum; $T$, characteristic autocorrelation time of the optimum, such that the autocorrelation over one generation is $\rho = e^{-(1/T)}$; $\alpha = N_eS\sigma_\theta^2$, compound parameter comparing the relative importance of fluctuating selection and drift as stochastic factors; $\kappa = SGT$, compound parameter characterizing responses to selection under an autoregressive optimum. Formulas for the autoregressive optimum with drift were obtained by using the continuous-time approximation for the effect of genetic drift, $V(\beta) = S/(2N_e)$ and $\rho(\beta, \tau) = e^{-SG\tau}$.**

$E(G)$ (where $E()$ denotes the expectation of a stochastic process) for each parameter set, using $E(S) = (\omega^2 + E(G) + V_e)^{-1}$ for the strength of stabilizing selection. The effective population size $N_e$ was computed as the harmonic mean of population sizes across generations, following standard population genetic theory (e.g., Hartl and Clark 2007, p. 122). We simulated only two replicates per parameter set, but averaged measurements over many generations, such as to get accurate stochastic expectations owing to the ergodicity property of our stationary systems.

# Results

Our main results are the analytical solutions for the variance and ACF of selection gradients under various regimes of environmental change in the optimum phenotype; the corresponding formulas are summarized in Table 1. The robustness of approximations underlying these derivations is then checked by population- and individual-based simulations. For brevity, we only apply these simulations to one scenario of environmental change, the autoregressive optimum, because this regime encompasses a wide variety of situations that approach other scenarios as limits, as detailed below. Throughout the results, $V(x)$ denotes the variance of a stochastic variable $x$, whereas $\rho(x, \tau)$ denotes its ACF $E(x_t, x_{t+\tau})/V(x)$.

## DRIFT IN A FINITE POPULATION

In finite populations, random genetic drift causes changes in the mean phenotype, resulting in directional selection to bring it back toward the optimum. This produces changes in β, even when the optimum does not move. Defining $x = \overline{z} - \theta$, and denoting values in the next generation with primes, we have (from Lande 1976) $x' = (1 - SG)x + \sqrt{G/N_e}\zeta$, where ζ is a standard normal random deviate. The first term is the deterministic force of selection bringing the mean phenotype back toward the optimum, whereas the second term is the effect of random genetic drift. From this, the recursion for the variance of $x$ is $V(x') = (1 - SG)^2V(x) + G/N_e$. Solving for the equilibrium and replacing $\beta = -Sx$, the stationary variance of selection gradients caused by drift around a constant optimum is

$$V(\beta|\text{drift}) = \frac{S}{(2 - SG)N_e}, \tag{1a}$$

It is easily shown that the autocorrelation of selection gradients τ generations apart is

$$\rho(\beta, \tau|\text{drift}) = (1 - SG)^{\tau}. \tag{1b}$$

Under weak stabilizing selection ($SG << 1$), equations (1a–b) converge to $V(\beta|\text{drift}) = S/(2N_e)$ and $\rho(\beta, \tau|\text{drift}) = e^{-SG\tau}$, which can also be obtained as solutions of the continuous-time Orstein–Uhlenbeck process for $x$ (as described in Lande 1976).

Even with a constant optimum, random genetic drift in a finite population thus causes temporally variable directional selection. The magnitude of these apparent fluctuations is larger under small effective population size, as expected, but also increases with the strength of stabilizing selection $S$. This may seem paradoxical, as a large $S$ prevents phenotypes from deviating strongly from the optimum. However, a given deviation from the optimum results in stronger directional selection with a large $S$, and this effect dominates. With a constant optimum, all autocorrelation in β is generated by evolutionary inertia determined by responses to selection, which have a time scale of $1/(SG)$: Selection gradients measured over a given time interval have stronger positive autocorrelation when stabilizing selection is weak, or genetic variance is low (see Hansen 1997 for the effect of the same process on the

covariance between mean phenotypes across species). For a given $SG$, autocorrelation decreases exponentially with the interval $\tau$ between the generations in which the selection gradients are measured.

If environmental change causes the optimum phenotype to fluctuate in time following a stationary Gaussian process, then the mean phenotype and the selection gradient also follow a Gaussian process (e.g., Bürger and Lynch 1995). Hence, by standard properties of Gaussian processes (Feller 1968; Karlin and Taylor 1981), the stochastic variance of $\beta$ is the sum of the variances caused by drift and by the fluctuating optimum, and similarly for the autocovariance, which can be used to derive the autocorrelation. Below, we therefore derive the variance and autocorrelation caused by fluctuating selection in a changing environment using an assumption of infinite population size for simplicity, and then combine those results with equation (1a) to include the effect of drift. The final equations including the effect of drift appear in the right column of Table 1.

## RANDOM WALK OF THE OPTIMUM

A model of random change in the optimum that shares similarities with random genetic drift is one in which the optimum follows a random walk (the discrete-time analog of Brownian motion), such that $\theta' = \theta + B\zeta$, where $\zeta$ is a standard normal random increment (Hansen 1997). The recursion for the selection gradient then is $\beta' = S(\theta' - \bar{z}') = (1 - SG)\beta + SB\zeta$, which leads to the stationary variance and ACF

$$V(\beta|RW) = \frac{SB^2}{(2 - SG)G} \tag{2a}$$

$$\rho(\beta, \tau|RW) = (1 - SG)^{\tau}, \tag{2b}$$

where the "*RW*" stands for "random walk." Under weak stabilizing selection, these can again be approximated by the solutions of a continuous-time OU process, $V_{\beta,RW} = SB^2/(2G)$ and $\rho_{\beta,RW}(\tau) = e^{-SG\tau}$. When the optimum phenotype undergoes a random walk, the variance of selection gradients thus increases with the squared magnitude $B^2$ of increments of the optimum, but decreases with the genetic variance of the trait (as long as $SG$ is small). Interestingly, the ACF of selection gradients in this case is the same as under genetic drift with a constant optimum.

## WHITE NOISE IN THE OPTIMUM

A random walk is a nonstationary stochastic process, wherein the squared distance of the optimum phenotype from its initial value is expected to increase with time. This does not accurately describe scenarios in which an optimum fluctuates within constant bounds. Stationary fluctuating selection in a randomly changing environment can be modeled by assuming for instance that the optimum phenotype has constant mean (set to 0 without loss of generality) and variance $\sigma_\theta^2$, and no correlation across generations (Lynch and Lande 1993; Bürger and Lynch 1995), which corresponds to a white noise process in continuous time (Karlin and Taylor 1981). Under this assumption, we show in the Appendix that

$$V(\beta|WN) = \frac{2S^2\sigma_\theta^2}{2 - SG} \tag{3a}$$

$$\rho(\beta, \tau > 0|WN) = -\frac{SG}{2}(1 - SG)^{\tau-1}, \tag{3b}$$

where the "*WN*" stands for "white noise." Equation (3a) shows that the variance of selection gradients increases with the stochastic variance $\sigma_\theta^2$ of the optimum phenotype and the strength $S$ of stabilizing selection, but that it *also* increases with the amount of response to selection, as measured by $SG$ for a given deviation from the optimum. This is because in an unpredictable environment, responding to selection in one generation generally means being farther from the optimum in the next, resulting in stronger directional selection, as highlighted in previous work about the lag load (Charlesworth 1993; Lynch and Lande 1993; Lande and Shannon 1996; Chevin 2013).

Strikingly, the autocorrelation of $\beta$ in equation (3b) is negative for all nonzero intervals, contrary to the drift and random walk cases. This is a consequence of unpredictable fluctuations in the optimum: Whenever the population responds strongly to selection in a given generation, the mean phenotype remains displaced from the average optimum for some time, and selection gradients in subsequent generations thus tend to be in the opposite direction. The strength of this autocorrelation decreases in a geometric series, but with a maximum absolute value (for $\tau = 1$) of $SG/2$ rather than 1. This is because autocorrelation of $\beta$ is here generated by persistent deviations from the average optimum caused by responses to selection, which scale with the evolutionary potential $SG$.

Summing equations (1a) and (3a) yields the variance of selection gradients under white noise in a finite population subject to drift, as given in Table 1. It depends on a compound parameter $\alpha = N_e S\sigma_\theta^2$, which determines the relative importance of fluctuations in the optimum versus random genetic drift. In populations small enough that $\alpha << 1$, most of the variation in selection gradients is caused by drift in the mean phenotype, whereas in the opposite case ($\alpha >> 1$), most of this variation is due to the movement of the phenotypic optimum. The ACF is similarly obtained as a weighted average of the ACFs caused by the two processes, leading to the formula in Table 1, which under weak stabilizing selection ($SG << 1$) can be approximated as

$$\rho(\beta, \tau > 0|WN + \text{drift}) \approx \frac{1 - \alpha SG}{1 + 2\alpha} e^{-SG\tau}. \tag{4}$$

This shows that even under completely uncorrelated fluctuations in the optimum, random genetic drift is sufficiently strong to

make the autocorrelation of β positive for all $\tau > 0$ whenever $\alpha < (SG)^{-1}$.

## AUTOREGRESSIVE OPTIMUM

Although white noise is a good approximation for rapid fluctuations, any changing environment will be autocorrelated over relevant time scales, leading to some degree of predictability. One of the simplest ways to include autocorrelation in a stationary process is through a first-order autoregressive process (AR(1), Box et al. 2008). We thus assume that $\theta' = \rho\theta + \sqrt{1-\rho^2}\sigma_\theta\zeta$, where $\rho = \rho(\theta, 1)$ is the autocorrelation of the optimum over one generation (and we assume $0 < \rho < 1$), $\sigma_\theta^2$ its stochastic variance, and ζ is a standard normal random deviate. It is convenient to parameterize the autocorrelation as $\rho = \exp(-1/T)$, such that the ACF between two optima separated by τ generations is $\rho(\theta, \tau) = \exp(-|\tau|/T)$. The parameter $T$ is the characteristic timescale for the autocorrelation of the optimum: Optima are similar over periods of time on the order of (or shorter than) $T$ generations. Dynamics with $T \leq 1$ resemble white noise, whereas with very large $T$ the dynamics are similar to (bounded) Brownian motion. Simple expressions can be derived for the stochastic variance and ACF of selection gradients in this context by using a continuous-time approximation that assumes weak stabilizing selection ($SG << 1$, Appendix), leading to

$$V(\beta|AR_1) = \frac{S^2\sigma_\theta^2}{1+\kappa}, \tag{5a}$$

$$\rho(\beta, \tau|AR_1) = \frac{e^{-\tau/T} - \kappa e^{-SG\tau}}{1-\kappa}, \tag{5b}$$

where $\kappa = SGT$ and $AR_1$ stands for first-order autoregressive process. Note that the expression for the variance is similar to that for the expected lag load in the same context, but replacing $S/2$ by $S^2$, as expected from the definition of variance in selection gradients (e.g., Lande and Shannon 1996, replacing $S$ by $\gamma$, or Appendix 1 in Chevin 2013).

Equation (5b) shows that autocorrelation of selection gradients is produced by two processes: (1) autocorrelation in the optimum phenotype (first term in numerator) and (2) persistent deviations of the mean phenotype from the average optimum caused by responses to selection, as in the case of white noise (second term in numerator). The latter term can even cause the autocorrelation to become negative for $\tau > T\ln(\kappa)/(\kappa-1)$, although these negative values approach 0 for large $T$. Under strong autocorrelation of the optimum ($T \to \infty$), the ACF of β tends to $e^{-SG\tau}$ as under Brownian motion of the optimum (see eq. (2b)). This sets an upper limit to the ACF: Even in highly autocorrelated environments, the autocorrelation of selection gradients cannot exceed that caused by the inertia of evolutionary responses to selection. For time intervals much shorter than $T$ and $(SG)^{-1}$, the ACF can be approximated as $e^{-(SG+1/T)\tau}$. This indicates that under weak stabilizing selection and when the optimum is autocorrelated over more than a few generations, the ACF of selection gradients over a short timescale (i.e., a few generations) is a simple exponential decay with rate $SG + 1/T$. This confirms that fluctuations in the optimum decrease the autocorrelation in selection gradients below the baseline of $e^{-SG\tau}$, particularly when fluctuations in the optimum are not strongly autocorrelated (small $T$). Combining these equations with the equations for drift in a finite population yields the formulas given in Table 1. Autocorrelation of β is mostly determined by drift when $\alpha << 1$ and $\alpha << \kappa$, in which case $\rho(\beta, \tau|AR_1 + \text{drift}) \to e^{-SG\tau}$.

## SIMULATIONS

Our analytical results rest on a number of assumptions (normality of phenotype and breeding value distributions, constant additive genetic variance $G$, and effective population size $N_e$), as well as approximations that produce more easily interpretable formulas (notably the continuous-time approximation). We investigated the robustness of predictions to these simplifications using two types of simulations: population-based simulations with constant $G$ and $N_e$, and individual-based simulations with explicit loci. The first type allows us to check the validity of our derivations for evolution in a stochastic environment, conditional on a constant genetic variance. The second type introduces further realism by allowing for changes in $G$, but is also hugely more computationally intensive, which necessarily limits the number of replicates that can be run. We focused on the case of an autoregressive optimum because it can approximate white noise (for $T \leq 1$) and random walk (for $T \to \infty$), as well as more intermediate scenarios.

The variance of selection gradients under genetic drift and a fluctuating optimum is plotted against the autocorrelation time of the optimum in Figure 1. The prediction for an autoregressive optimum with drift (summing eqs. (5a) and (1a), formula in Table 1) matches population-based numerical simulations better for $T > 2$, as expected because the continuous-time approximation is not valid for very rapid changes. For $T << 1$, simulation results are accurately described by the prediction under white noise with drift (summing eqs. (3a) and (1a), see Table 1). For a given $N_e$, the variance is largest under essentially nonautocorrelated fluctuations ($T << 1$), because the mean phenotype is then unable to track the optimum, and when it does it results in larger deviations from the optimum (and stronger directional selection) in the future, on average. In contrast, with a very slow and predictable optimum ($T \to \infty$), the variance of β converges to its minimum value, the constant-optimum case in which all deviations from the optimum are caused by genetic drift (eq. (1a)).

Figure 2 shows how the ACF of selection gradients changes with the characteristic timescale of autocorrelation in the

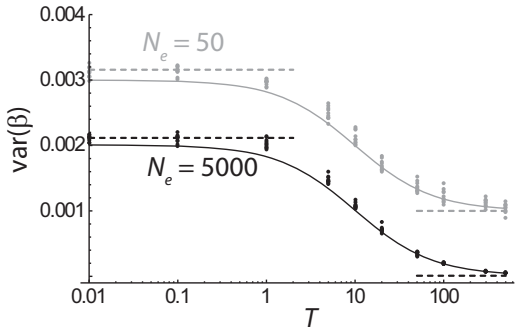


**Figure 1. Stationary variance of selection gradients. The variance of selection gradients under an autoregressive optimum is plotted against the autocorrelation time *T* of the optimum, for an effective population size of $N_e = 50$ (gray) or $N_e = 5000$ (black). The continuous lines show the analytical predictions for an autoregressive optimum with drift (Table 1). The dashed lines on the right represent the limit for drift and a constant optimum (eq. 1a), whereas those on the left are the results assuming white noise optimum with drift (Table 1). The dots show results of population-based simulations (10 replicates) of evolution of a quantitative trait with constant genetic variance in response to an autoregressive optimum. Parameters are $G = 1$, $S = 0.1$, and $\sigma_\theta^2 = 0.2$.**

optimum. The results of population-based simulations with constant variance and weak stabilizing selection are captured well by the analytical prediction for an autoregressive process when $T \geq 1$, and by the prediction for a white noise process when $T < 1$. (Note that the effects of genetic drift can be ignored here because $\alpha = 100 >> 1$.) For $T < 1$ (top left panel), the ACF is similar to that under white noise, with weak negative autocorrelation (dashed line). When $T$ is on the order of the generation time, the ACF shows a rapid exponential-like decrease (with rate $SG + 1/T$) for small $\tau$, followed by a dip into negative values, until a minimum is reached and autocorrelation increases back toward zero (top right panel). As $T$ increases, the initial decay in autocorrelation becomes less pronounced, and negative values are smaller in magnitude (bottom left panel in Fig. 2). Finally, for very large $T$, the ACF tends toward a simple exponential decay with rate $SG$ determined only by responses to selection (dotted line in Fig. 2).

In individual-based simulations with explicit loci, the genetic variance may change substantially across generations (e.g., Bürger and Lande 1994). We thus used the long-term mean genetic variance from the simulations $E(G)$ in the analytical formulas for comparison. The variance of selection gradients in individual-based simulations is accurately predicted by the analytical formula for an autoregressive optimum with drift (all points close to the $x = y$ line in Fig. 3). All else being equal, $V(\beta)$ increases with the strength of stabilizing selection (from blue to red) and the variance in the optimum (size of dots within each color), as predicted by the analytical theory. The prediction in Table 1 performs less well under weak stabilizing selection and small variance in the optimum (small blue dots to the left), that is, under conditions leading to the smallest predicted $V(\beta)$. Indeed in simulations (and in actual finite populations), the randomness of birth and death

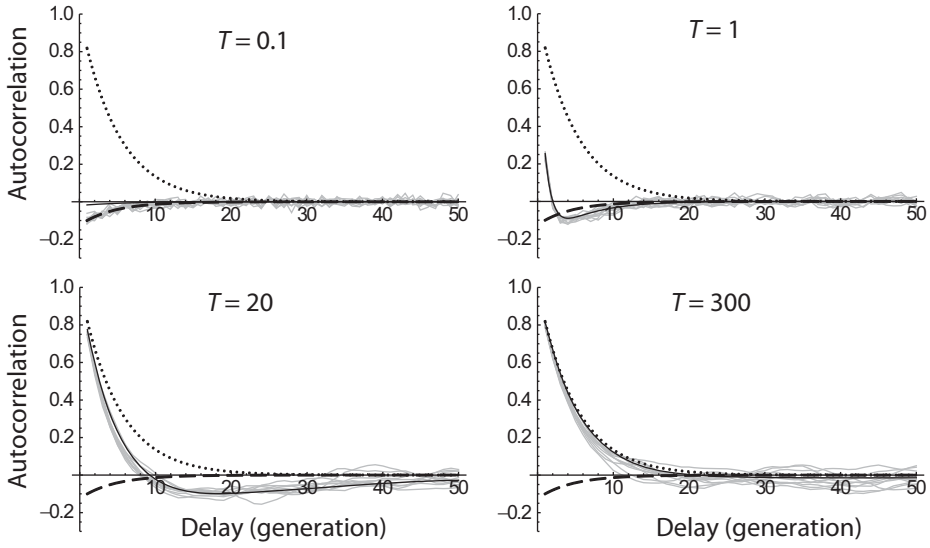


**Figure 2. Autocorrelation function of selection gradients. The ACF of selection gradients under an autoregressive optimum is illustrated for several values of the autocorrelation time *T* of the optimum. The black line shows the prediction for an autoregressive optimum neglecting drift (eq. 5b), whereas the gray lines show results of population-based simulations of evolution of a quantitative trait with constant genetic variance in responses to an autoregressive optimum (10 replicates). Also shown in each graph are the limits for high autocorrelation of the optimum, which equals that for drift with a constant optimum (eq. 1b, dotted), and the limit for a white noise optimum neglecting drift (eq. 3b, dashed). Parameters are $G = 1$, $S = 0.2$, $\sigma_\theta^2 = 0.2$, and $N_e = 5000$.**

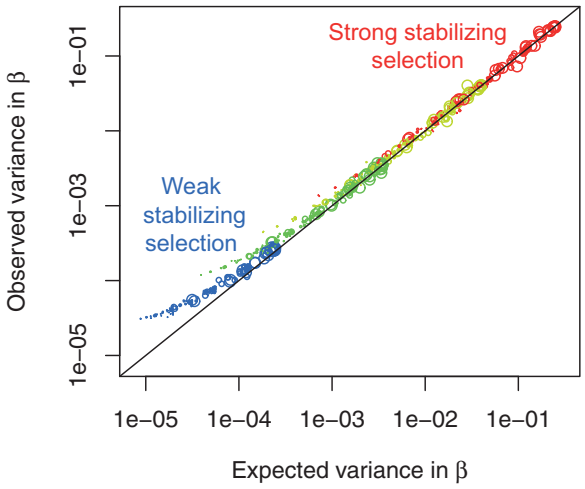


**Figure 3.** Variance of β in individual-based simulations. The observed variance in β in individual-based simulations with explicit loci is plotted against its expectation under autoregressive optimum with drift (Table 1). Colors indicate different strengths of stabilizing selection (red: $\omega = 1$, yellow: $\omega = 2$, green: $\omega = 4$, blue: $\omega = 8$), and the size of the dots is proportional to the variance of optimum phenotypes ($\sigma_\theta^2 = 0.05$, 0.1, 0.2, or 0.5). Other variable parameters are the autocorrelation time of the optimum ($T = 1$, 4, 15, 50, or 200), and the mutation rate per locus ($\mu = 10^{-5}$, $10^{-4}$, or $10^{-3}$). The black line shows $y = x$.

events (demographic stochasticity) causes variance in selection gradient for a given phenotype distribution and fitness surface (e.g., Engen and Saether 2014), and this variance acts as noise for the detection of $V(\beta)$: theoretical predictions do not perform well below this precision threshold.

The ACF in individual-based simulations is also predicted well by the theory, under parameters allowing maintenance of substantial genetic variance (Fig. 4). However, this ACF may be noisy in individual realizations, despite being computed over several thousand generations. This noise is reduced under higher mutational variance (with $\mu = 10^{-3}$ and $a = 10$, Fig. S1), probably because this limits the stochastic changes in the genetic variance. Our prediction for the ACF becomes less accurate under smaller fluctuations ($\sigma_\theta^2 = 0.05$ instead of 1, Fig. S2), because genetic drift then contributes more to the autocorrelation (Table 1, autoregressive with drift), and our approximations neglect its effects on the genetic variance. Overall, the temporal distribution of selection gradients is thus reasonably well predicted by a model that neglects temporal changes in genetic variance, as long as the trait maintains substantial variance in every generation, and fluctuations in the optimum are sufficiently large that drift is not a strong driver of phenotypic change.

Although our analytical formulas make no assumptions about relationships between variables, the graphs in Figures 1 and 2 were obtained by keeping the same genetic variance, while varying the pattern of environmental fluctuations. This amounts to assuming that the long-term mean genetic variance $E(G)$ and strength of stabilizing selection $E(S)$ do not depend on patterns of fluctuations in the optimum, which is not true in general (Bürger and Gimelfarb 2002). The validity of this assumption in individual-based simulations is assessed in the left panel of Figure 5, in which we used a single value for the mutation rate, such that all heterogeneity of genetic variance is due to the parameter values used for the strength of stabilizing selection and the pattern of environmental change. More genetic variance is maintained under weaker stabilizing selection (lighter gray), as expected. But for a given strength of stabilizing selection, the mean genetic variance across generations also changes with $S\sigma_\theta^2$, the effective magnitude of fluctuations. Genetic variance is larger under stronger fluctuations, and the slope of the relationship between $E(G)$ and $E(S)\sigma_\theta^2$ depends little on the width $\omega$ of the fitness function: The two darker set of dots in Figure 5 (left panel) have similar slopes over the same range of $E(S)\sigma_\theta^2$. For large $E(S)\sigma_\theta^2$, there is also a tendency for $E(G)$ to increase with the autocorrelation time $T$ (size of dots). This is consistent with the fact that strongly autocorrelated random change in the optimum is similar to directional change, and the latter is known to cause increased genetic variance (relative to a constant optimum) in individual-based simulations (Fig. 6 in Bürger and Lynch 1995; Fig. 1 in Bürger 1999).

In all scenarios but the random walk, the genetic variance enters the analytical formulas only through the product $SG$. Because $S$ is a decreasing function of the phenotypic (and genetic) variance, changes in $G$ are not sufficient to understand how the accuracy of analytical predictions depends on patterns of change in the optimum. We report changes of $E(SG)$, the long-term mean of $SG$, with the intensity of fluctuations of the optimum, in the right panel of Figure 5. Weaker stabilizing selection (in lighter gray) results in smaller $E(SG)$, which varies little with the magnitude of fluctuations in the optimum. But under stronger stabilizing selection, $E(SG)$ increases with the effective intensity of fluctuations $S\sigma_\theta^2$, and to some extent with the autocorrelation time $T$.

## Discussion and Conclusions

Measuring the temporal variance of β, the directional selection gradient, is perhaps the most straightforward way to quantify temporal variation of phenotypic selection (as done by Siepielski et al. 2009 on individual estimates of β, and corrected for sampling error by Morrissey and Hadfield 2012). However, for any temporal measurement, variation is not only characterized by its variance, but also by its autocorrelation, which determines its predictability. This is particularly important for measurements of natural selection, for which the predictability of fluctuations is at least as important as their variance for predicting how responses to

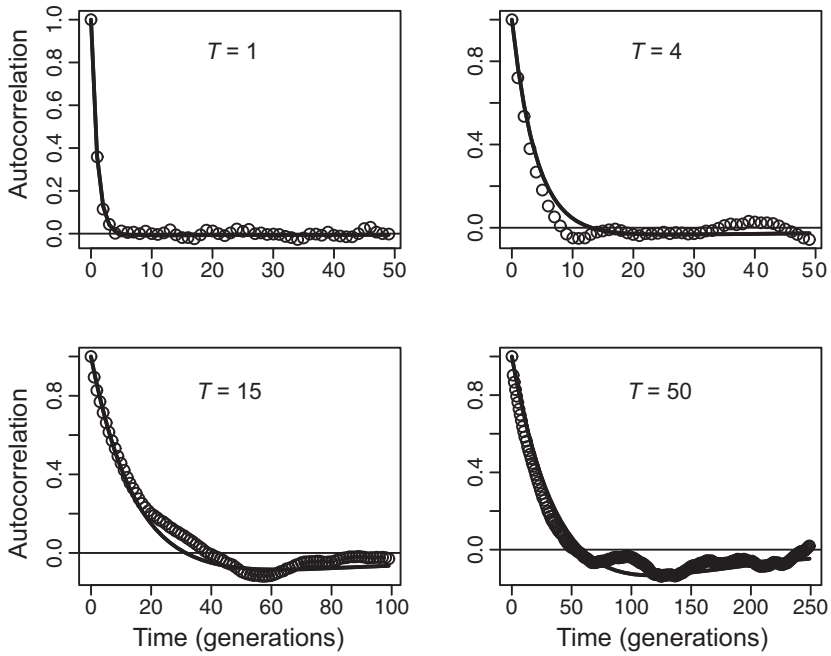


**Figure 4.** Autocorrelation of β in individual-based simulations. The observed ACF of β in single realizations of individual-based simulations is shown as dots, together with its expectation under autoregressive optimum with drift (Table 1), as continuous line. Parameters are as indicated in the text, with $\omega = 4$, $\mu = 10^{-4}$, and $\sigma_\theta^2 = 1$.

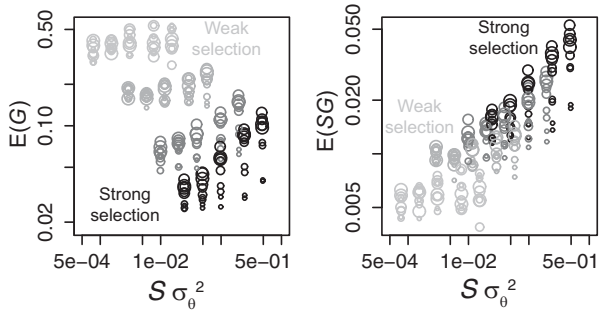


**Figure 5.** Effect of fluctuations on the evolutionary potential. The long-term mean for the additive genetic variance *G* and the product *SG* that determines the potential to respond to selection are plotted against the effective intensity of fluctuations in the optimum, $S\sigma_\theta^2$. Each gray level corresponds to a strength of stabilizing selection, with $\omega = 1$, 2, 4, and 8 from black to light gray. The dot sizes increase with autocorrelation time *T*. Only mutation rate $\mu = 10^{-4}$ is shown for clarity.

selection affect population long-term fitness (Charlesworth 1993; Lande and Shannon 1996; Chevin 2013). Ideally, phenotypic selection should thus be treated explicitly as a time series to estimate not only the variance, but also the ACF, of selection gradients. Here, we computed β over the entire population for simplicity, but in principle time series of selection gradients can also be analyzed from population subsamples, by using state-space models such as dynamic linear regression (Petris et al. 2009; Shumway and Stoffer 2010). However, in practice, estimating the autocorrelation of selection gradients may require a large sample size and many time points. In any case, a proper use of such methods would warrant a careful investigation of the properties of the underlying statistical approach, which is beyond the scope of the present article, and will be addressed elsewhere.

Even assuming that the variance and ACF of selection gradients can be estimated accurately, it is worth asking what can be inferred from them about the mechanisms that drive apparent fluctuating selection. It is well known from quantitative genetic theory that changes in directional selection gradients do not necessarily imply changes in the phenotype-fitness map. For instance, β decreases deterministically when the mean trait in a population approaches a constant optimum phenotype (Lande 1976; Gomulkiewicz and Holt 1995). And random genetic drift can also cause temporally variable phenotypic selection, by moving the mean phenotype away from a constant optimum. Here, we derived the temporal distribution of directional selection gradients β under several scenarios of adaptation to a fixed or moving optimum. All situations we investigated produced both variance and autocorrelation of β. Even a nonautocorrelated optimum produces weak negative autocorrelation of β, because genetic responses to occasional large environmental changes cause persistent deviations from the average optimum, resulting in an average selection gradient in the opposite direction. An autoregressive optimum may in some cases by inferred through a characteristic pattern, where the autocorrelation of β changes from positive to negative with increasing time lag (Fig. 2). In contrast, a similar exponential

decay of positive autocorrelation across time is produced under drift with a constant optimum, random walk of the optimum, or a highly autocorrelated autoregressive optimum, the first of which is causally very different from the other two.

Our results thus suggest that measurements of the temporal variance and autocorrelation of selection gradients, although informative about the overall amount of directional selection taking place in a population, are not always sufficient to infer the underlying processes driving variation in that selection. A more powerful approach for mechanistic inference would be to jointly analyze phenotypic selection and changes in the phenotype distribution, to directly estimate temporal changes in the fitness surface. For instance, Johnson et al. (2014) recently analyzed temporally changing phenotypic selection by assuming a fitness surface with a constant optimum, but this is likely to be an oversimplification in many cases. A Gaussian fitness function with a moving optimum, as modeled here, can be estimated in a generalized linear-mixed model that includes a quadratic term in the relationship between log fitness and phenotype, and a random effect with the relevant temporal structure, such as $AR_1$ (L.–M. Chevin and J. Tufto, unpubl. ms.). Such an approach would afford a better understanding and prediction of the evolution and demography of populations, by relating more explicitly to theoretical predictions (Lynch and Lande 1993; Bürger and Lynch 1995; Gomulkiewicz and Holt 1995; Lande and Shannon 1996; Gomulkiewicz and Houle 2009; Chevin 2013).

Our model assumes that the environment only causes changes in the position of the optimum phenotype, but in reality the environment might also affect the width of the fitness peak (as modeled by Revell 2007). If changes in peak width are moderate and uncorrelated to those in the optimum, then using the mean peak width over time in the formulas of Table 1 should still provide good approximations for the variance and autocorrelation of selection gradients, although more time points would then be required for accurate estimation. Note that in all our results, the peak width only acts through the compound parameter $S = 1/(\omega^2 + G + V_e)$, which is allowed to vary in individual-based simulations (through changes in $G$), so we addressed changes in the strength of stabilizing selection to some extent.

Our analytical predictions are for a quantitative trait with substantial standing genetic variance, and in simulations we also mostly investigated cases in which substantial genetic variance can be maintained—although some results also hold under somewhat strong stabilizing selection and weak mutation (Fig. 3). In the opposite limit of very little polymorphism ("strong selection weak mutation," Gillespie 1991), adaptation mostly relies on new mutations, and has different properties, even in environments that change deterministically and at a constant rate (genetically limited adaptation regime, Kopp and Hermisson 2007, 2009). In a random environment, whether the population will respond to each chance displacement of the optimum in that regime depends on the availability of mutations in the required direction. This adds the randomness of mutation events as another layer of evolutionary stochasticity, beyond those caused by genetic drift and fluctuating selection (Lenormand et al. 2009), which should substantially complicate the analysis. In this low polymorphism regime, it should be more fruitful to focus on the temporal distribution of selection coefficients of mutations (Gillespie 1991), rather than on phenotypic selection gradients as studied here. However, these genotypic selection coefficients may still ultimately arise from stabilizing selection on quantitative traits (Martin and Lenormand 2006; Chevin and Hospital 2008), so it might be possible to use the temporal changes in the distribution of fitness effects of mutations to infer fluctuations in the phenotypic fitness surface.

## ACKNOWLEDGMENTS

We thank J. Tufto and R. Bürger for suggesting more accurate derivations for some of our results, and J. Ashander, R. Bürger, A. Charmantier, E. Kisdi, M. B. Morrissey, R. G. Shaw, and J. Tufto for comments on earlier versions of this manuscript. This project was funded by the ContempEvol grant (ANR 11 PDOC 005 01) from the Agence Nationale de la Recherche.

## LITERATURE CITED

Bell, G. 2010. Fluctuating selection: the perpetual renewal of adaptation in variable environments. Philos. Trans. R Soc. B 365:87–97.

Benkman, C. W., and R. E. Miller. 1996. Morphological evolution in response to fluctuating selection. Evolution 50:2499–2504.

Box, G. E. P., G. M. Jenkins, and G. C. Reinsel. 2008. Time series analysis. Wiley, Hoboken, NJ.

Bull, J. J. 1987. Evolution of phenotypic variance. Evolution 41:303–315.

Bürger, R. 1999. Evolution of genetic variability and the advantage of sex and recombination in changing environments. Genetics 153:1055–1069.

———. 2000. The mathematical theory of selection, recombination, and mutation. Wiley, Chichester, U.K.

Bürger, R., and A. Gimelfarb. 2002. Fluctuating environments and the role of mutation in maintaining quantitative genetic variation. Genet. Res. 80:31–46.

Bürger, R., and R. Lande. 1994. On the distribution of the mean and variance of a quantitative trait under mutation-selection-drift balance. Genetics 138:901–912.

Bürger, R., and M. Lynch. 1995. Evolution and extinction in a changing environment—a quantitative-genetic analysis. Evolution 49:151–163.

Cain, A. J., L. M. Cook, and J. D. Currey. 1990. Population-size and morph frequency in a long-term study of *Cepaea nemoralis*. Proc. R. Soc. B Biol. Sci. 240:231–250.

Charlesworth, B. 1993. The evolution of sex and recombination in a varying environment. J. Hered. 84:345–350.

Charlesworth, B., R. Lande, and M. Slatkin. 1982. A neo-Darwinian commentary on macroevolution. Evolution 36:474–498.

Chesson, P. 2000. Mechanisms of maintenance of species diversity. Annu. Rev. Ecol. Syst. 31:343–366.

Chevin, L. M. 2013. Genetic constraints on adaptation to a changing environment. Evolution 67:708–721.

Chevin, L. M., and F. Hospital. 2008. Selective sweep at a quantitative trait locus in the presence of background genetic variation. Genetics 180:1645–1660.

Dempster, E. R. 1955. Maintenance of genetic heterogeneity. Cold Spring Harb. Symp. Quant. Biol. 20:25–32.

Eldredge, N., and S. J. Gould. 1972. Punctuated equilibria: an alternative to phyletic gradualism. Pp. 82–115 *in* T. J. M. Schopf, ed. Models in paleobiology. Freeman, Cooper and Company, San Francisco, CA.

Ellner, S., and N. G. Hairston. 1994. Role of overlapping generations in maintaining genetic-variation in a fluctuating environment. Am. Nat. 143:403–417.

Engen, S., R. Lande, and B. E. Saether. 2011. Evolution of a plastic quantitative trait in an age-structured population in a fluctuating environment. Evolution 65:2893–2906.

Engen, S., and B. E. Saether. 2014. Evolution in fluctuating environments: decomposing selection into additive components of the Robertson-Price equation. Evolution 68:854–865.

Estes, S., and S. J. Arnold. 2007. Resolving the paradox of stasis: models with stabilizing selection explain evolutionary divergence on all timescales. Am. Nat. 169:227–244.

Falconer, D. S., and T. F. MacKay. 1996. Introduction to quantitative genetics. Longman Group, Harlow, U.K.

Feller, W. 1968. An introduction to probability theory and its applications. Wiley, New York.

Felsenstein, J. 1974. The evolutionary advantage of recombination. Genetics 78:737–756.

Fisher, R. A. 1930. The genetical theory of natural selection. Oxford Univ. Press, Oxford, U.K.

Fisher, R. A., and E. B. Ford. 1947. The spread of a gene in natural conditions in a colony of the moth *Panaxia dominula*. Heredity 1:143–174.

Garant, D., L. E. Kruuk, R. H. McCleery, and B. C. Sheldon. 2007. The effects of environmental heterogeneity on multivariate selection on reproductive traits in female great tits. Evolution 61:1546–1559.

Gavrilets, S. and S. M. Scheiner. 1993. The genetics of phenotypic plasticity. 5. Evolution of reaction norm shape. J. Evol. Biol. 6:31–48.

Gillespie, J. H. 1974. Nautural selection for within-generation variance in offspring number. Genetics 76:601–606.

———. 1991. The causes of molecular evolution. Oxford Univ. Press, Oxford, U.K.

Gimenez, O., R. Covas, C. R. Brown, M. D. Anderson, M. B. Brown, and T. Lenormand. 2006. Nonparametric estimation of natural selection on a quantitative trait using mark-recapture data. Evolution 60:460–466.

Gingerich, P. D. 2001. Rates of evolution on the time scale of the evolutionary process. Genetica 112:127–144.

Gomulkiewicz, R., and R. D. Holt. 1995. When does evolution by natural selection prevent extinction. Evolution 49:201–207.

Gomulkiewicz, R., and D. Houle. 2009. Demographic and genetic constraints on evolution. Am. Nat. 174:E218–E229.

Gordo, I., and P. R. Campos. 2012. Evolution of clonal populations approaching a fitness peak. Biol. Lett. 9:20120239.

Grant, P. R., and B. R. Grant. 2002. Unpredictable evolution in a 30-year study of Darwin's finches. Science 296:707–711.

Haldane, J. B. S., and S. D. Jayakar. 1963. Polymorphism due to selection of varying direction. J. Genet. 58:237–242.

Haller, B. C., and A. P. Hendry. 2014. Solving the paradox of stasis: squashed stabilizing selection and the limits of detection. Evolution 68:483–500.

Hansen, T. F. 1997. Stabilizing selection and the comparative analysis of adaptation. Evolution 51:1341–1351.

Hartl, D. L., and A. G. Clark. 2007. Principles of population genetics. Sinauer Associates, Sunderland, MA.

Hedrick, P. W. 2006. Genetic polymorphism in heterogeneous environments: the age of genomics. Annu. Rev. Ecol. Evol. Syst. 37:67–93.

Hedrick, P. W., M. E. Ginevan, and E. P. Ewing. 1976. Genetic-polymorphism in heterogeneous environments. Annu. Rev. Ecol. Syst. 7:1–32.

Johnson, D. W., K. Grorud-Colvert, S. Sponaugle, and B. X. Semmens. 2014. Phenotypic variation and selective mortality as major drivers of recruitment variability in fishes. Ecol. Lett. 17:743–755.

Jones, A. G., R. Bürger, S. J. Arnold, P. A. Hohenlohe, and J. C. Uyeda. 2012. The effects of stochastic and episodic movement of the optimum on the evolution of the G-matrix and the response of the trait mean to selection. J. Evol. Biol. 25:2210–2231.

Karlin, S., and H. M. Taylor. 1981. A second course in stochastic processes. Academic Press, San Diego, CA.

Kingsolver, J. G. 1995. Viability selection on seasonally polyphenic traits—wing melanin pattern in western white butterflies. Evolution 49:932–941.

Kingsolver, J. G., and S. E. Diamond. 2011. Phenotypic selection in natural populations: what limits directional selection? Am. Nat. 177:346–357.

Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. E. Hill, A. Hoang, P. Gibert, and P. Beerli. 2001. The strength of phenotypic selection in natural populations. Am. Nat. 157:245–261.

Kinnison, M. T., and A. P. Hendry. 2001. The pace of modern life II: from rates of contemporary microevolution to pattern and process. Genetica 112:145–164.

Kopp, M., and J. Hermisson. 2007. Adaptation of a quantitative trait to a moving optimum. Genetics 176:715–719.

———. 2009. The genetic basis of phenotypic adaptation II: the distribution of adaptive substitutions in the moving optimum model. Genetics 183:1453–1476.

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

———. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. Evolution 33:402–416.

———. 1993. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. Am. Nat. 142:911–927.

———. 2007. Expected relative fitness and the adaptive topography of fluctuating selection. Evolution 61:1835–1846.

———. 2009. Adaptation to an extraordinary environment by evolution of phenotypic plasticity and genetic assimilation. J. Evol. Biol. 22:1435–1446.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.

Lande, R., S. Engen, and B. Saether. 2003. Stochastic population dynamics in ecology and conservation: an introduction. Oxford Univ. Press, Oxford, U.K.

Lande, R., and S. Shannon. 1996. The role of genetic variation in adaptation and population persistence in a changing environment. Evolution 50:434–437.

Lenormand, T., D. Roze, and F. Rousset. 2009. Stochasticity in evolution. Trends Ecol. Evol. 24:157–165.

Lynch, M., and R. Lande. 1993. Evolution and extinction in response to environmental change. Pp. 234–250 *in* P. Kareiva, J. Kingsolver, and R. Huey, eds. Biotic interactions and global change. Sinauer, Sunderland, MA.

MacColl, A. D. 2011. The ecological causes of evolution. Trends Ecol. Evol. 26:514–522.

Martin, C. H., and P. C. Wainwright. 2013. Multiple fitness peaks on the adaptive landscape drive adaptive radiation in the wild. Science 339:208–211.

Martin, G., and T. Lenormand. 2006. A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. Evolution 60:893–907.

Morrissey, M. B., and J. D. Hadfield. 2012. Directional selection in temporally replicated studies is remarkably consistent. Evolution 66:435–442.

Orr, H. A. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution 52:935–949.
Otto, S. P., and T. Day. 2007. A biologist's guide to mathematical modeling in ecology and evolution. Princeton Univ. Press, Princeton, NJ.
Petris, G., S. Petrone, and P. Campagnoli. 2009. Dynamic linear models with R. Springer, New York, NY.
Reimchen, T. E., and P. Nosil. 2002. Temporal variation in divergent selection on spine number in threespine stickleback. Evolution 56:2472–2483.
Revell, L. J. 2007. The G matrix under fluctuating correlational mutation and selection. Evolution 61:1857–1872.
Scheiner, S. M. 2014. The genetics of phenotypic plasticity. XIII. Interactions with developmental instability. Ecol. Evol. 4:1347–1360.
Schluter, D. 1988. Estimating the form of natural-selection on a quantitative trait. Evolution 42:849–861.
Schluter, D., and D. Nychka. 1994. Exploring fitness surfaces. Am. Nat. 143:597–616.
Shumway, R. H., and D. S. Stoffer. 2010. Time series analysis and its applications: with R examples. Springer, New York, NY.
Siepielski, A. M., J. D. DiBattista, and S. M. Carlson. 2009. It's about time: the temporal dynamics of phenotypic selection in the wild. Ecol. Lett. 12:1261–1276.
Siepielski, A. M., J. D. DiBattista, J. A. Evans, and S. M. Carlson. 2011. Differences in the temporal dynamics of phenotypic selection among fitness components in the wild. Proc. Biol. Sci. 278:1572–1580.
Simpson, G. G. 1944. Tempo and mode in evolution. Columbia Univ. Press, New York.
Slatkin, M. 1974. Hedging ones evolutionary bets. Nature 250:704–705.
Svardal, H., C. Rueffler, and J. Hermisson. 2011. Comparing environmental and genetic variance as adaptive response to fluctuating selection. Evolution 65:2492–2513.
Uyeda, J. C., T. F. Hansen, S. J. Arnold, and J. Pienaar. 2011. The million-year wait for macroevolutionary bursts. Proc. Natl Acad. Sci. USA 108:15908–15913.
van Asch, M., P. H. van Tienderen, L. J. M. Holleman, and M. E. Visser. 2007. Predicting adaptation of phenology in response to climate change, an insect herbivore example. Glob. Chang. Biol. 13:1596–1604.
Wade, M. J., and S. Kalisz. 1990. The causes of natural selection. Evolution 44:1947–1955.
Wright, S. 1935. The analysis of variance and the correlations between relatives with respect to deviations from an optimum. J. Genet. 30:243–256.
———. 1937. The distribution of gene frequencies in populations. Science 85:504.
Zhang, X. S. 2012. Fisher's geometrical model of fitness landscape and variance in fitness within a changing environment. Evolution 66:2350–2368.

Associate Editor: E. Kisdi

# Appendix

## Derivation of the Variance and ACF of Selection Gradients

### *WHITE NOISE OPTIMUM*

To derive the variance and autocorrelation of directional selection gradients under white noise in the optimum, we rely on the fact that optimum phenotypes are independent across generations in this case, and the mean phenotype in one generation is also independent from the optimum in future generations. From the definition of the selection gradient in generation $t$ as $\beta_t = -S(\overline{z}_t - \theta_t)$ in the main text, we have directly $V(\beta_t) = S^2[\sigma_\theta^2 + V(\overline{z}_t)]$, where $V(\overline{z}_t)$ is the variance of the stochastic distribution of mean phenotypes in generation $t$. In the next generation, the selection gradient is $\beta_{t+1} = -S(\overline{z}_{t+1} - \theta_{t+1})$, which using $\overline{z}_{t+1} = \overline{z}_t(1 - SG) + SG\theta_t$ can be written as $\beta_{t+1} = -S[(1 - SG)\overline{z}_t - (\theta_{t+1} - SG\theta_t)]$. Because the distribution of $\theta$ is assumed to be stationary, we have $V(\theta_{t+1}) = V(\theta_t) = \sigma_\theta^2$, leading to

$$V(\beta_{t+1}) = S^2\left[(1 - SG)^2 V(\overline{z}_t) + (1 + (SG)^2)\sigma_\theta^2\right],$$

or replacing with the expression for $V(\beta_t)$ above,

$$V(\beta_{t+1}) = (1 - SG)^2 V(\beta_t) + 2S^3 G\sigma_\theta^2.$$

This is a simple arithmetico-geometric sequence, which for $SG < 1$ (i.e., moderately strong stabilizing selection) converges to an equilibrium variance $V(\beta)$. We can thus replace $V(\beta_{t+1}) = V(\beta_t) = V(\beta)$ in the equation above, leading to $V(\beta)[1 - (1 - SG)^2] = 2S^3 G\sigma_\theta^2$. Solving for $V(\beta)$ yields equation (3a)) in the main text.

The covariance of selection gradients across generations can be derived in a similar way. First note that for all time lags $\tau$, the autocovariance is $\mathrm{cov}(\beta_t, \beta_{t+\tau}) = E(\beta_t\beta_{t+\tau})$ (where $E$ denotes expectation of the stochastic process), since $E(\beta_t) = 0$ for all $t$. For $\tau = 1$ (one generation of lag), the autocovariance is the expectation of

$$\beta_t\beta_{t+1} = S^2(\overline{z}_t - \theta_t)\left[(1 - SG)\overline{z}_t - (\theta_{t+1} - SG\theta_t)\right].$$

Taking the expectation accounting for all independent stochastic variables (optimum and mean phenotype), we have

$$\mathrm{cov}(\beta_t, \beta_{t+1}) = S^2[(1 - SG)V(\overline{z}_t) - SG\sigma_\theta^2],$$

or replacing with the expression for $V(\beta_t)$ above,

$$\mathrm{cov}(\beta_t, \beta_{t+1}) = (1 - SG)V(\beta_t) - S^2\sigma_\theta^2.$$

For all delays $\tau \geq 1$ we have in a similar way

$$\begin{aligned}\beta_t\beta_{t+\tau+1} &= S^2(\overline{z}_t - \theta_t)[(1 - SG)\overline{z}_{t+\tau} - (\theta_{t+\tau+1} - SG\theta_{t+\tau})] \\ &= S^2(\overline{z}_t - \theta_t)[(1 - SG)(\overline{z}_{t+\tau} - \theta_{t+\tau}) - \theta_{t+\tau+1} + \theta_{t+\tau}].\end{aligned}$$

Taking the expectations on both sides yields

$$\mathrm{cov}(\beta_t, \beta_{t+\tau+1}) = (1 - SG)\mathrm{cov}(\beta_t, \beta_{t+\tau}).$$

Induction from this recursion indicates that for all $\tau \geq 1$, we have

$$\mathrm{cov}(\beta_t, \beta_{t+\tau}) = (1 - SG)^{\tau-1}\mathrm{cov}(\beta_t, \beta_{t+1}).$$

The final step to find the stationary autocorrelation function is to replace $V(\beta_t)$ by its equilibrium value $V(\beta)$ in the expression for $\mathrm{cov}(\beta_t, \beta_{t+1})$. This yields

$$\mathrm{cov}(\beta_t, \beta_{t+1}) = \frac{2S^2\sigma_\theta^2(1 - SG)}{2 - SG} - S^2\sigma_\theta^2,$$

or grouping in a single fraction,

$$\mathrm{cov}(\beta_t, \beta_{t+1}) = \frac{2S^2\sigma_\theta^2}{2 - SG}(1 - SG - 1 + \frac{SG}{2}).$$

We then have for all $\tau \geq 1$

$$\mathrm{cov}(\beta_t, \beta_{t+\tau}) = -\frac{SG}{2}(1 - SG)^{\tau-1}V(\beta),$$

which produces equation (3b) in the main text for the autocorrelation function.

### AUTOREGRESSIVE OPTIMUM

Although the analysis of responses to selection with an autoregressive optimum also can be analyzed in discrete time (Charlesworth 1993), it is more convenient to approximate the dynamics in continuous time (Lande and Shannon 1996; Chevin 2013). After long enough time that the effect of initial conditions have vanished, assuming constant $G$ the mean phenotype in generation $t$ can be written $\overline{z}_t = SG\int_0^\infty e^{-SGx}\theta_{t-x}dx$ (Lande and Shannon 1996). The autocorrelation of selection gradients is derived by plugging this into the expression for the autocovariance,

$$\mathrm{cov}(\beta_t, \beta_{t-\tau}) = S^2\left[E(\theta_t\theta_{t-\tau}) + E(\overline{z}_t\overline{z}_{t-\tau}) - E(\overline{z}_{t-\tau}\theta_t) - E(\theta_{t-\tau}\overline{z}_t)\right].$$

Note that the last two terms in the equation above are not the same because $\tau$ introduces some asymmetry. Let us assume that $\tau > 0$ (but the results hold for $\tau < 0$). For the first expectation, we have by definition $E(\theta_t\theta_{t-\tau}) = \sigma_\theta^2\exp(-\tau/T)$. For the first cross product of $\overline{z}$ and $\theta$, we have

$$\begin{aligned} E(\overline{z}_{t-\tau}\theta_t) &= SG\int_0^\infty e^{-SGx}E(\theta_t\theta_{t-(\tau+x)})dx \\ &= SG\sigma_\theta^2\int_0^\infty \exp\left\{-\left(SGx + \frac{\tau + x}{T}\right)\right\}dx \\ &= \frac{SGT}{1 + SGT}\sigma_\theta^2 e^{-\tau/T}, \end{aligned}$$

where we have used the fact that $\tau + x > 0$ between the first and second line. For the second cross product, the expectation inside the integral is $E(\theta_{t-x}\theta_{t-\tau}) = \sigma_\theta^2\exp(-|\tau - x|/T)$. Because of the absolute value, the integral needs to be split in two at $x = \tau$,

$$\begin{aligned} \frac{E(\theta_{t-\tau}\overline{z}_t)}{SG\sigma_\theta^2} &= \int_0^\tau \exp\left\{-\left(SGx + \frac{\tau - x}{T}\right)\right\}dx \\ &\quad + \int_\tau^\infty \exp\left\{-\left(SGx + \frac{x - \tau}{T}\right)\right\}dx \\ &= \frac{T}{SGT - 1}\left(e^{-\tau/T} - e^{-SG\tau}\right) + \frac{T}{SGT + 1}e^{-SG\tau}. \end{aligned}$$

For the integral involving the mean phenotype in two generations, we have

$$\begin{aligned} &\frac{E(\overline{z}_{t-\tau}\overline{z}_t)}{(SG)^2\,\sigma_\theta^2} \\ &= \int_0^\infty\int_0^\infty \exp\left\{-\left(SG(x + y) + \frac{|x - (\tau + y)|}{T}\right)\right\}dxdy. \end{aligned}$$

Here again, the absolute value entails that the double integral needs to be split into three parts: one for $x > \tau + y$ (which we denote $A$), another for $\tau + y > x$ and $x < \tau$ (integral $B$), and the other for $\tau < x < \tau + y$ (integral $C$). Solving the integrals yields,

$$\begin{aligned} A &= \int_0^\infty\left(\int_{y+\tau}^\infty \exp\left\{-\left(SG(x + y) + \frac{x - (\tau + y)}{T}\right)\right\}dx\right)dy \\ &= \frac{Te^{-SG\tau}}{2SG(SGT + 1)} \end{aligned}$$

$$\begin{aligned} B &= \int_0^\tau\left(\int_0^\infty \exp\left\{-\left(SG(x + y) + \frac{\tau + y - x}{T}\right)\right\}dy\right)dx \\ &= \frac{T^2}{(SGT)^2 - 1}\left(e^{-\tau/T} - e^{-SG\tau}\right) \end{aligned}$$

$$\begin{aligned} C &= \int_\tau^\infty\left(\int_{x-\tau}^\infty \exp\left\{-\left(SG(x + y) + \frac{\tau + y - x}{T}\right)\right\}dy\right)dx \\ &= A. \end{aligned}$$

Summing $A$+$B$+$C$, we thus have

$$\frac{E(\overline{z}_{t-\tau}\overline{z}_t)}{(SG)^2\,\sigma_\theta^2} = \frac{1}{SGT + 1}\left[\frac{Te^{-SG\tau}}{SG} + \frac{T^2}{SGT - 1}\left(e^{-\tau/T} - e^{-SG\tau}\right)\right],$$

which yields for the autocovariance of selection gradients caused by an autoregressive optimum,

$$\begin{aligned} \frac{\mathrm{cov}(\beta_t, \beta_{t-\tau})}{S^2\sigma_\theta^2} &= e^{-\tau/T} - SG\left(\frac{T}{1 + SGT}e^{-\tau/T}\right. \\ &\left. + \frac{T}{SGT - 1}\left(e^{-\tau/T} - e^{-SG\tau}\right)\right. \end{aligned}$$

$$+\frac{T}{SGT+1}e^{-SG\tau}\Bigg)$$
$$+(SG)^2\left(\frac{1}{SGT+1}\left[\frac{Te^{-SG\tau}}{SG}+\frac{T^2}{SGT-1}\left(e^{-\tau/T}-e^{-SG\tau}\right)\right]\right)$$

or writing $\kappa = SGT$,

$$\frac{\kappa^2-1}{S^2\sigma_\theta^2}\mathrm{cov}(\beta_t,\beta_{t-\tau}) = (\kappa-1)(1+\kappa)e^{-\tau/T}$$
$$-\kappa(\kappa-1)e^{-\tau/T}-(1+\kappa)\kappa\left(e^{-\tau/T}-e^{-SG\tau}\right)-(\kappa-1)\kappa e^{-SG\tau}$$
$$+\kappa(\kappa-1)e^{-SG\tau}+\kappa^2\left(e^{-\tau/T}-e^{-SG\tau}\right).$$
$$=-(e^{-\tau/T}-\kappa e^{-SG\tau})$$

The stochastic variance is obtained by setting $\tau = 0$ above, and the autocorrelation function is found by dividing the covariance by the variance, yielding the expressions equations (5a–b) in the main text.

## *Supporting Information*

Additional Supporting Information may be found in the online version of this article at the publisher's website:

**Figure S1**. Autocorrelation of $\boldsymbol{\beta}$ under high mutational variance.
**Figure S2**. Autocorrelation of $\boldsymbol{\beta}$ under weak fluctuations.