doi:10.1111/j.1558-5646.2012.01809.x

# GENETIC CONSTRAINTS ON ADAPTATION TO A CHANGING ENVIRONMENT

**Luis-Miguel Chevin**[1,2]

[1]*Centre d'Ecologie Fonctionnelle et Evolutive, CNRS, Montpellier, France*

[2]*E-mail: chevin.lm@gmail.com*

Received February 7, 2012
Accepted September 10, 2012

**Genetic correlations between traits can constrain responses to natural selection. To what extent such correlations limit adaptation depends on patterns of directional selection. I derive the expected rate of adaptation (or evolvability) under randomly changing selection gradients. When directional selection gradients have an arbitrary covariance matrix, the average rate of adaptation depends on genetic correlations between traits, contrary to the isotropic case investigated in previous studies. Adaptation may be faster on average with more genetic correlation between traits, if these traits are selected to change jointly more often than the average pair of traits. However, natural selection maximizes the long-term fitness of a population, not necessarily its rate of adaptation. I therefore derive the average lag load caused by deviations of the mean phenotype from an optimum, under several forms of environmental changes typically experienced by natural populations, both stochastic and deterministic. Simple formulas are produced for how the G matrix affects long-term fitness in these contexts, and I discuss how their parameters can be estimated empirically.**

**KEY WORDS:** **Directional selection, environmental stochasticity, evolutionary demography, evolutionary quantitative genetics, fluctuating selection, punctuated equilibria, rate of adaptation.**

To establish a firm link with biological reality, evolutionary theory should be able to accurately predict phenotypic change as observed in wild populations (Endler 1986; Kinnison and Hendry 2001), notably for quantitative traits with polygenic inheritance such as body size, and other morphological, physiological, or phenological traits (Kruuk et al. 2008). This need has become more urgent with the realization that anthropogenic environmental alterations (notably global climate change) not only impose strong selective pressures on natural populations, but also affect their demography and extinction risk, and that the influence of the environment on population growth is largely mediated by the phenotype (Pelletier et al. 2007; Ozgul et al. 2009; Ozgul et al. 2010; Pelletier et al. 2011).

Evolutionary predictions for individual traits may fail when they are genetically correlated with other traits under conflicting selective pressures. Following the formalization of this idea (Lande 1979) and the development of associated estimation procedures (Lande and Arnold 1983; Arnold and Wade 1984), numerous studies have investigated the additive genetic covariance structure of quantitative traits, and confronted it with patterns of natural selection (reviewed in Kingsolver and Diamond 2011). A recent meta-analysis of reported estimates of phenotypic selection (directional gradients and differentials) showed that phenotypic correlations frequently limit directional selection on body size, but not on other traits (Kingsolver and Diamond 2011). Assuming genetic and phenotypic correlations are similar, this result suggests weak genetic constraints on responses to selection by the average trait. However, all traits are not equivalent to the organism, and the efficiency of adaptive evolution should instead be measured by how much the mean fitness increases in response to natural selection in any generation. Such an analysis was recently performed assuming constant directional selection (Agrawal and Stinchcombe 2009), but it is likely that the direction of strength of selection is temporally variable in wild populations.

Recent attempts to quantify temporal variation in directional selection (Siepielski et al. 2009, 2011) were complicated by the


© 2012 The Author.


EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

large error variance of each estimate (Morrissey and Hadfield 2012). However, several facts argue in favor of temporal variation in phenotypic selection. First for most organisms, many aspects of their environment, either abiotic (temperature, precipitation) or biotic (density of competitors, predators, parasites. . .) vary between generations. Second, population sizes often fluctuate substantially, even for organisms with vital rates inconsistent with deterministic cycling or chaos (Lande et al. 2003, pp. 3–6). Such demographic fluctuations are likely caused by random changes in the environment (environmental stochasticity), which should also cause fluctuating selection (Lande 2007; Engen et al. 2011). And third, meta-analyses of rates of phenotypic evolution combining data from different time scales (field studies, fossil record, and comparative analyses) show that rapid phenotypic change may occur over a few tens of generations, but rarely accumulate over longer time (Estes and Arnold 2007; Uyeda et al. 2011). Such rapid but noncumulative evolution indicates that mean phenotypes generally fluctuate within bounds, with rare events of irreversible evolution outside these bounds provoking long-term divergence (Uyeda et al. 2011). This pattern is consistent with a model where an optimum phenotype fluctuates moderately within an adaptive zone, with key innovations or strong environmental changes occasionally causing strong displacements toward new adaptive zones (Simpson 1944).

Evolutionary theory has recently investigated how the genetic covariances of traits affect the average increase in fitness in response to variable selection (Hansen and Houle 2008; Kirkpatrick 2009). However, the authors assumed selection with constant strength and uniform phenotypic orientation, implying all traits are equivalent for adaptation (isotropy). Although these analyses provided useful insights, they have limited meaning when directional selection varies nonuniformly across traits. Moreover, these studies made no explicit link between responses to selection and the long-term fitness of a population. Here our aim will be to identify general multivariate constraints on adaptation, to understand how the additive genetic (co)variances affect the rate of adaptation and long-term fitness under various types of environmental change. We will see that the **G** matrix favored by evolution in the long run need not be the one that maximizes the average response to directional selection.

## Constraints on Adaptation and Long-Term Fitness

Let us first illustrate the general argument with an arbitrary fitness function, and in terms of parameters amenable to direct empirical estimation.

### ASSUMPTIONS ON THE INHERITANCE OF TRAITS

We focus throughout on $n$ quantitative traits with polygenic inheritance. The phenotype of each individual is the sum of a genetic (heritable) component, and a normally distributed residual component caused by microenvironmental variation and developmental noise. The latter has constant mean 0, and same covariance matrix **E** for all genotypes, consistent with standard quantitative genetics (Falconer and Mackay 1996; Lynch and Walsh 1998). Phenotypic variation is determined many polymorphic loci, such that the additive genetic (or breeding) values have multivariate Gaussian distribution. Hence the phenotype distribution is also normal, with total phenotypic covariance matrix $\mathbf{P} = \mathbf{G} + \mathbf{E}$. We will further assume for simplicity that neither **G** nor **E** change in time. Neglecting changes in additive genetic (co)variances should be a good approximation when many loci contribute vanishingly small effects to the traits (infinitesimal model, Fisher 1918; Falconer and Mackay 1996), or when abundant mutations with continuous effects occur at loosely linked loci (Lande 1976a, 1980a). Genetic variances then rapidly reach an equilibrium set by the strength of stabilizing selection (Turelli 1988 provides a thorough discussion of the limitations and benefits of the assumption of constant genetic covariances). We will address changes in the **G** matrix in section Discussion.

### RATE OF ADAPTATION AND EVOLVABILITY

Denote as $\mathbf{z}$ the vector containing the phenotypic values of $n$ quantitative traits, with mean $\bar{\mathbf{z}}$ and additive genetic covariance matrix **G** in a population. With discrete nonoverlapping generations, $W(\mathbf{z})$ is the absolute multiplicative fitness of phenotype $\mathbf{z}$, and the mean fitness in the population is $\overline{W}$. Frequency-independent selection is assumed, such that mean fitness always increases in response to selection within a generation (Wright 1969; Crow and Kimura 1970), although it may decrease across generations if a change in the environment alters the fitness function (see below).

One way to quantify how much the responses of multiple traits to selection affect overall adaptation is by translating changes in mean additive genetic effects (or breeding values) of all traits into changes in mean fitness, under *current* conditions of selection. According to Frank and Slatkin (1992) and references therein, this is what Fisher (1930) originally meant by "rate of adaptation" in his fundamental theorem: a change in fitness measured in a common frame of reference. With unbiased environmental effects on the phenotype, the within-generation change in mean breeding values equals the between-generation change in mean phenotype $\Delta\bar{\mathbf{z}}$. To second order in $\Delta\bar{\mathbf{z}}$, the within-generation change in logarithmic mean fitness in response to natural selection, or rate of adaptation, is

$$\Delta^* \ln \overline{W} = \boldsymbol{\beta}^T \Delta\bar{\mathbf{z}} + 1/2 \Delta\bar{\mathbf{z}}^T \mathbf{H} \Delta\bar{\mathbf{z}} \qquad (1a)$$

(Lande 1979, eq. 8a), where superscripted "$T$" denotes matrix/vector transposition, and the star denotes change within a generation. The selection gradient $\boldsymbol{\beta} = \nabla \ln \overline{W}$ is the vector of

partial derivatives of $\ln \overline{W}$ with respect to the mean of each trait. It quantifies directional selection operating directly on each trait, correcting for phenotypic correlations with other traits. Geometrically, it points toward the direction of steepest slope in the adaptive landscape representing log mean fitness against the mean of each trait (Lande 1979). The $(i,j)$th element of the Hessian matrix $\mathbf{H}$ contains the second derivative of $\ln \overline{W}$ relative to the means of traits $i$ and $j$, $\partial^2 \ln \overline{W}/(\partial \overline{z_i} \partial \overline{z_j})$. It quantifies quadratic selection (stabilizing, disruptive, or correlative), and measures the curvature of the log mean fitness landscape close to the current mean phenotype.

With a multivariate Gaussian phenotype distribution, the genetic response to directional selection by the mean phenotype is $\Delta \bar{\mathbf{z}} = \mathbf{G}\boldsymbol{\beta}$, accounting for both direct responses and indirect responses caused by genetic correlations (Lande 1979). The within-generation change in the genetic covariance matrix in response to natural selection is $\Delta^* \mathbf{G} = \mathbf{GHG}$ (Lande and Arnold 1983). Inserting into equation (1a), the rate of adaptation becomes

$$\Delta^* \ln \overline{W} = \boldsymbol{\beta}^T \left(\mathbf{G} + \tfrac{1}{2} \Delta^* \mathbf{G}\right) \boldsymbol{\beta}. \tag{1b}$$

Under weak stabilizing selection ($tr(\mathbf{HG}) << n$), $\mathbf{G}$ changes little within a generation, leading to

$$\Delta^* \ln \overline{W} \approx \boldsymbol{\beta}^T \mathbf{G} \boldsymbol{\beta}. \tag{1c}$$

Equation (1a–c) quantify genetic constraints on adaptation directly through their effects on the rate of fitness increase within a generation. Approximation (1c) is a popular measure of "evolvability," described notably by Hansen and Houle (2008), and represents a projection of the vector of phenotypic change along the direction of the current selection gradient. But whenever stabilizing selection is not negligible, it should be more accurate to use (1b). All basic parameters of equation (1a,b) can be estimated using multivariate quadratic regression (Lande and Arnold 1983).

In a changing environment, another interesting quantity is the total change in log mean fitness over a generation, $\Delta \ln \overline{W} = \ln \overline{W}'(\bar{\mathbf{z}}') - \ln \overline{W}(\bar{\mathbf{z}})$ (primes denote values in the next generation). To first order in $\Delta \bar{\mathbf{z}} = \mathbf{G}\boldsymbol{\beta}$ around $\bar{\mathbf{z}}$, this change is

$$\Delta \ln \overline{W} \approx [\ln \overline{W}'(\bar{\mathbf{z}}) - \ln \overline{W}(\bar{\mathbf{z}})] + \boldsymbol{\beta}'^T \mathbf{G} \boldsymbol{\beta}. \tag{1d}$$

The term in brackets is the contribution from changes in the mean fitness function *for a given mean phenotype* $\bar{\mathbf{z}}$, owing to a change in the environment or in higher moments of the phenotype distribution. Fisher (1930, pp. 41–42) described this term as "environmental deterioration" of fitness, which he saw as the major cause of natural selection (see also Frank and Slatkin 1992). The second term is our main focus here, as it measures how much change in $\ln \overline{W}$ is caused by the responses of multiple traits to selection over one generation. Note that in (1d), $\boldsymbol{\beta}'$ is the gradient of $\ln \overline{W}'$ evaluated at $\bar{\mathbf{z}}$ instead of $\bar{\mathbf{z}}'$, which is a valid approximation under weak quadratic selection (small $tr(\mathbf{GH})$).

## EXPECTED RATE OF ADAPTATION WITH RANDOM SELECTION GRADIENTS

To investigate how the additive genetic covariance $\mathbf{G}$ matrix influences the average rate of adaptation in a variable environment, let us consider that the directional selection gradient $\boldsymbol{\beta}$ is drawn in each generation from an arbitrary multivariate distribution, with mean $\bar{\boldsymbol{\beta}}$ and covariance matrix $\mathbf{B}$. The mean selection gradient $\bar{\boldsymbol{\beta}}$ gives the directional trend in selection, whereas $\mathbf{B}$ characterizes the fluctuations of directional selection across generations. The diagonal elements of $\mathbf{B}$ give the variances in strengths of directional selection for each trait. The off-diagonal elements of $\mathbf{B}$ give the covariances of directional selection between pairs of traits, indicating whether they tend to be selected in the same direction, and with similar strength, in any generation.

In such a fluctuating environment, the rate of adaptation in equation (1c) is a quadratic form in random vectors, whose properties are well characterized (Mathai and Provost 1992). In particular, assuming a constant $\mathbf{G}$ matrix, the expected rate of adaptation is

$$\mathrm{E}(\Delta^* \ln \overline{W}) = \bar{\boldsymbol{\beta}}^T \mathbf{G} \bar{\boldsymbol{\beta}} + \mathrm{tr}(\mathbf{BG}), \tag{2a}$$

where tr() denotes the trace of a matrix, the sum of its diagonal elements. Equation (2a) generalizes previous results by Kirkpatrick (2009) and Hansen and Houle (2008), who both considered selection gradients with mean 0, uniform directions in the phenotypic space, and constant norm. The first term in equation (2a) is caused by the trend in directional selection, and is similar to the effect of a constant selection gradient $\bar{\boldsymbol{\beta}}$. The second term in (2a) is the expected change in fitness in response to fluctuations in directional selection. Because $\mathbf{B}$ and $\mathbf{G}$ are positive (semi-)definite (by definition of a covariance matrix), tr($\mathbf{BG}$) is always positive. In other words, more adaptive evolution takes place in populations with, than without, fluctuations in the direction of selection.

Interpretations of equation (2a) in terms of variances and covariances of individual traits, or of directions in the phenotypic space (using eigenvectors), are provided in Appendix S1, illustrating the link to previous results derived under more restrictive conditions by Kirkpatrick (2009) and Hansen and Houle (2008). In particular, term-by-term expansion of matrix product $\mathbf{BG}$ shows that genetic correlations can increase the expected rate of adaptation, if they occur between traits that tend to be selected to change jointly stronger (and more often) than the average pair of traits. The eigenvalues of $\mathbf{BG}$ can be used to define an effective number of traits (effective complexity) similar to that proposed by Kirkpatrick (2009) for the $\mathbf{G}$ matrix, but that also takes into account patterns of directional selection. Finally, note that equation

(2a) also holds if **B** is positive semidefinite, such that only a subset of genetically variable traits (or linear combinations thereof) are under fluctuating selection.

Furthermore, if the distribution of $\boldsymbol{\beta}$ is multivariate normal (as arises naturally in a model of moving optimum described below), the variance in rates of adaptation is

$$\mathrm{Var}(\Delta^* \ln \overline{W}) = 4\bar{\boldsymbol{\beta}}^T \mathbf{G}\mathbf{B}\mathbf{G}\bar{\boldsymbol{\beta}} + 2\mathrm{tr}((\mathbf{B}\mathbf{G})^2), \qquad (2b)$$

which quantifies how much stochasticity in the evolutionary process is caused by fluctuations in directional selection. With a constant covariance matrix **G**, the time until the realized mean of $\Delta^* \ln \overline{W}$ reaches a given confidence interval around its expected value in equation (2a) is proportional to $\mathrm{Var}(\Delta^* \ln \overline{W})$. Note that the trend in directional selection (quantified by $\bar{\boldsymbol{\beta}}$) affects the *variance* in rates of adaptation only in the presence of fluctuations (non-null **B**, first term in eq. 2b), while it affects the *expected* rate of adaptation even without fluctuations (first term in eq. 2a).

To account more explicitly for environmentally induced changes in the fitness function, equation (1d) can be used to compute the contribution of phenotypic evolution to the overall change in log mean fitness over a generation. If environmental fluctuations cause a stationary distribution of selection gradients, then $\mathrm{E}(\boldsymbol{\beta}'^T \mathbf{G}\boldsymbol{\beta}) = \bar{\boldsymbol{\beta}}^T \mathbf{G}\bar{\boldsymbol{\beta}} + \mathrm{tr}(\mathbf{B}_g\mathbf{G})$ with $\mathbf{B}_g = \mathrm{E}((\boldsymbol{\beta}' - \bar{\boldsymbol{\beta}})(\boldsymbol{\beta} - \bar{\boldsymbol{\beta}})^T)$, the cross covariance matrix of selection gradients over one generation. In particular, when selection gradients have the same autocorrelation $\rho_\beta$ for all traits over a generation, $\mathrm{tr}(\mathbf{B}_g\mathbf{G}) = \rho_\beta \mathrm{tr}(\mathbf{B}\mathbf{G})$, showing that the contribution of selection responses to the overall change in mean fitness over a generation is larger when directional selection is highly autocorrelated across generations.

## RATE OF ADAPTATION VERSUS LONG-TERM FITNESS

Equation (2a) only deals with the *change* in mean fitness in response to natural selection (within or between generations), not with mean fitness per se. Yet, what is maximized by (frequency-independent) natural selection is mean fitness, and not necessarily its rate of change (Wright 1969; Crow and Kimura 1970; Lande 1976b). In a fluctuating environment, evolution is expected to maximize the long-term growth rate $r_e = E(\ln \overline{W})$, even in populations with density-dependent growth, as long as density-dependent regulation affects all phenotypes or genotypes similarly, such that selection is density independent (Lande 2007).

The reason why rates of adaptation are not sufficient to predict long-term fitness is made clear in equation (1d), where the first term captures the "environmental deterioration" (sensu Fisher 1930) caused by changes in the fitness function across generations. This term is independent of responses to selection, and needs to be accounted for to predict how the **G** matrix affects adaptation in the long run. In the next section, I analyze a popular model of moving optimum phenotype (reviewed in Arnold et al. 2001), where all changes in the fitness function result from movements of a fitness peak with constant shape, which naturally causes both environmental deterioration of fitness and directional selection. I use this model to find how the **G** matrix affects long-term fitness, under representative types of environmental change.

## MODELS OF MOVING OPTIMUM

Models where an optimum phenotype fluctuates randomly in time have been analyzed for single characters (Slatkin and Lande 1976; Bull 1987; Charlesworth 1993a,b; Hansen and Martins 1996; Lande and Shannon 1996), or multiple uncorrelated characters (Zhang 2012), as have been multivariate models with a deterministic trend in the optimum (Hellmann and Pineda-Krch 2007; Gomulkiewicz and Houle 2009; Duputié et al. 2012). I here generalize these results by allowing multiple correlated traits to be under randomly fluctuating selection, combined with a deterministic trend in directional selection. The aim is to find how the genetic covariance structure constrains adaptation, by affecting the expected deviation of the mean phenotype from the optimum in the long run.

## MODEL SPECIFICATION

### *Mean fitness and loads*

Consider optimizing selection, such that the fitness of individuals with phenotype $\mathbf{z}$ is a multivariate Gaussian function of their deviation from an optimum phenotype $\boldsymbol{\theta}$, $W(\mathbf{z}) = W_{\max} \exp\{-1/2(\mathbf{z} - \boldsymbol{\theta})^T \Omega^{-1}(\mathbf{z} - \boldsymbol{\theta})\}$ (Lande 1979). Diagonal terms in the positive definite matrix $\Omega$ give the width of the fitness function for each individual trait, which is inversely proportional to the strength of stabilizing selection on this trait. Off-diagonal terms in $\Omega$ give the width of the fitness function for pairs of traits; corresponding terms in $\Omega^{-1}$ quantify correlative selection between pairs of traits. The mean fitness in a given generation is obtained by integrating $W$ over the phenotype distribution, yielding $\overline{W} = W_{\max}\sqrt{\det(\mathbf{S}\Omega)} \exp\{-1/2(\bar{\mathbf{z}} - \boldsymbol{\theta})^T \mathbf{S}(\bar{\mathbf{z}} - \boldsymbol{\theta})\}$ (Gomulkiewicz and Houle 2009), where det() denotes the determinant of a matrix, and $\mathbf{S} = (\Omega + \mathbf{P})^{-1}$ quantifies stabilizing and correlative selection on the mean phenotype. The logarithmic mean fitness is thus

$$\ln \overline{W} = r_{\max} - L_G - L_\theta, \qquad (3a)$$

which is similar to earlier formulations (Pease et al. 1989; Lynch and Lande 1993; Kirkpatrick and Barton 1997; Gomulkiewicz and Houle 2009; Polechova et al. 2009). The term $r_{\max} = \ln(W_{\max})$ is the intrinsic rate of increase of a population where all individuals have the optimum phenotype. The two subtracted terms are loads that are always positive, and therefore always decrease fitness. The (co)variance load (or standing load) $L_G = -1/2 \ln(\det(\mathbf{S}\Omega))$ (or

$L_G \approx 1/2tr(\Omega^{-1}\mathbf{P})$ under weak stabilizing selection, such that $tr(\mathbf{\Omega}^{-1}\mathbf{P}) << tr(\mathbf{I})$) quantifies the decrease in $\ln \overline{W}$ caused by deviations of individual phenotypes from the mean phenotype. It is constant with respect to deviations of the mean phenotype from the optimum, and therefore occurs even when the mean phenotype is at the optimum. The last term is the lag load (Maynard-Smith 1976),

$$L_\theta = \tfrac{1}{2}(\overline{\mathbf{z}} - \boldsymbol{\theta})^T \mathbf{S}(\overline{\mathbf{z}} - \boldsymbol{\theta}), \tag{3b}$$

which increases with maladaptation of the mean phenotype (deviations from the optimum). This term is what we are most interested in, because it is affected by responses to natural selection.

### *Response to selection*

From the definition of the selection gradient (below eq. 1a), the response to selection in any generation is

$$\Delta\overline{\mathbf{z}} = -\mathbf{GS}(\overline{\mathbf{z}} - \boldsymbol{\theta}), \tag{4a}$$

where the minus sign indicates that evolution always acts to reduce the mean deviation from the optimum in the ongoing generation (Lande 1979; Gomulkiewicz and Houle 2009). The cumulative evolutionary change over multiple generations can be computed assuming constancy of the $\mathbf{G}$ matrix and of the shape and orientation of the fitness peak (fluctuations of peak shape were investigated by Revell 2007). Changing the variable from $\overline{\mathbf{z}}$ to $\mathbf{x} = (\overline{\mathbf{z}} - \boldsymbol{\theta})$, and approximating the evolutionary dynamics in continuous time (which is valid under weak directional selection), equation (4a) becomes a multivariate differential equation,

$$d\mathbf{x}/dt + \mathbf{GSx} = -d\boldsymbol{\theta}/dt, \tag{4b}$$

which can be solved using standard techniques (Higham 2008; Boyce and DiPrima 2009). Changing back to $\overline{\mathbf{z}}$ yields as an asymptotic solution (for large $t$)

$$\overline{\mathbf{z}}_t = \mathbf{GS}\int_0^\infty \exp(-\mathbf{GS}\tau)\boldsymbol{\theta}_{t-\tau}d\tau, \tag{4c}$$

where the matrix exponential can be found by diagonalizing the matrix, taking the exponential of its nonzero eigenvalues, and changing back to the original space (Higham 2008). Equation (4c) is the multivariate analog to equation (7) in Lande and Shannon (1996) (and a similar one in discrete time by Charlesworth 1993a). It shows that in the long run, the mean phenotype integrates the responses to all past environments, weighting each optimum by how far in the past it was encountered. This cumulative process results in a buffering of environmental fluctuations, and in effect "filters" the environmental "signal." For instance in Figure 2, the mean of each trait tracks the major movements of its corresponding optimum, but more rapid fluctuations are buffered (compare same colors in Fig. 2C and D).

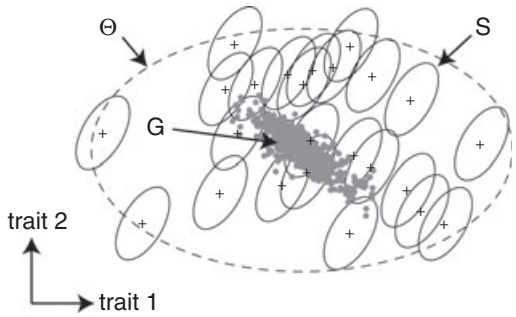


**Figure 1. Schematic illustration of a fluctuating optimum phenotype. The matrices involved in the evolution of multiple traits in a fluctuating environment are represented for the particular case of two traits under selection. The cloud of gray dots shows the distribution of breeding values (inherited component of the phenotype) in the population; its shape and orientation are governed by the G matrix. The continuous ellipses show isofitness curves, determined by the S matrix, assuming Gaussian optimizing selection. Lastly, fluctuations in the optimum phenotype (crosses) follow a pattern controlled by the covariance matrix Θ (dashed line).**

Equations (3) and (4) can be combined to find the expected lag load, and hence the long-term growth rate of a population, under various types of environmental change, which I describe below.

### *Pattern of environmental change*

Current global change combines a long-term deterministic trend with shorter term, largely random, fluctuations (Trenberth and Josey 2007), a pattern that is probably typical. We therefore consider that environmentally induced changes in the optimum phenotype can be partitioned in two terms, $\boldsymbol{\theta}_t = \mathbf{d}_t + \mathbf{s}_t$. The deterministic component $\mathbf{d}_t$ represents a trend such as global warming, rising $CO_2$, etc... The stochastic (random) component $\mathbf{s}_t$ describes fluctuations in the optimum, and is taken to be a stationary Gaussian process with mean $\mathbf{0}$ (without loss of generality) and covariance matrix $\mathbf{\Theta}$ among traits. Figure 1 illustrates graphically the meaning of $\mathbf{\Theta}$, in relation to the two other matrices of interest, $\mathbf{G}$ and $\mathbf{S}$. Because the deterministic and stochastic components are independent, their effects on the lag load are additive (Lynch and Lande 1993), so the latter can be partitioned as $L_\theta = L_d + L_s$, with $L_d$ the deterministic component and $L_s$the stochastic component.

## LAG LOAD CAUSED BY STOCHASTIC FLUCTUATIONS

### *Uncorrelated fluctuations*

Let us first assume that environmental change causes the optimum phenotype to undergo multivariate white noise movement, with mean 0 and covariance matrix $\mathbf{\Theta}$ (Box et al. 2008, p. 553), such that the position of the optimum is independent from one generation to the next. Equation (4b) then is a stochastic differential equation, whose solution is a multivariate Ornstein–Uhlenbeck

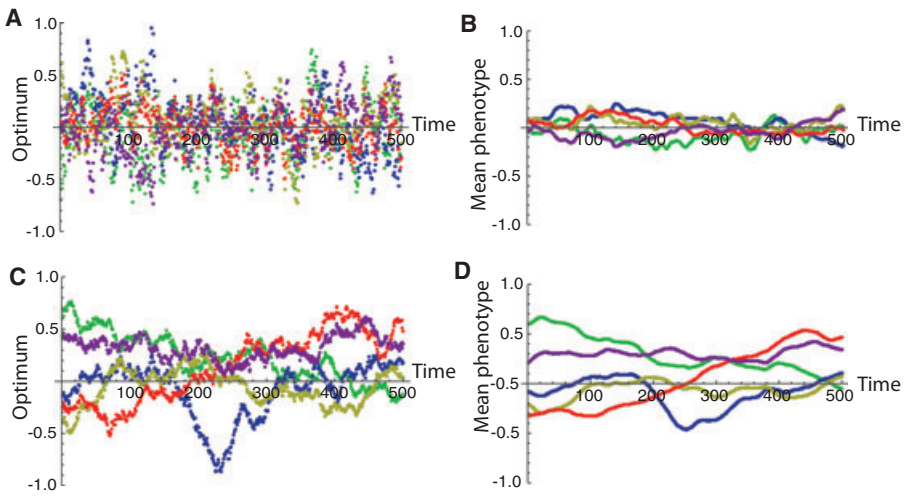


**Figure 2.** **Response to autocorrelated fluctuating selection. The optimum phenotypes (A, C) and mean phenotypes (B, D) for five traits (colors) are represented under weak ($T_\rho = 10$, A, B) or strong ($T_\rho = 200$, C, D) environmental autocorrelation. Note that the mean of each trait is more able to track its optimum under autocorrelated fluctuations. Simulations were run with matrices G, S, and Θ drawn randomly as detailed in the main text, with the conditions $tr(\mathbf{GS}) = 0.1$, and $tr(\mathbf{S\Theta}) = 0.4$.**

process (Karlin and Taylor 1981). Both the optimum and the mean phenotype tend to stationary multivariate normal distributions. The stochastic expectation of the lag load contributing to the long-term population growth rate is

$$E_w(L_s) = \tfrac{1}{2} tr\left(\mathbf{S\Theta}\left(\mathbf{I} + \tfrac{1}{2}\mathbf{GS}\right)\right), \tag{5a}$$

where the subscript "$w$" is for "white noise" (details of the derivation are given in Appendix S2). Equation (5a) is the multivariate analog to previous expressions for a single trait (Lynch and Lande 1993; Lande and Shannon 1996). Under strong stabilizing selection (large $tr(\mathbf{GS})$), (5a) becomes

$$E_{w,\text{strong}}(L_s) \approx \tfrac{1}{4} tr(\mathbf{GS}^2\mathbf{\Theta}), \tag{5b}$$

where use has been made of the fact that $tr(\mathbf{AB}) = tr(\mathbf{BA})$ for symmetric matrices (cyclic property of traces).

Equation (5a,b) show that with uncorrelated fluctuations in the optimum, the lag load increases with increasing $tr(\mathbf{GS})$, that is, when trait combinations under stronger stabilizing selection have more genetic variance. In this context, genetic constraints on the rate of adaptation are thus *favored* in the long run. The reason is that in a completely unpredictable environment, responding to selection in any given generation usually implies being farther from the optimum in the next (Charlesworth 1993b; Lande and Shannon 1996). The best option is then for the mean phenotype to remain constant, and equal to the average optimum. It is noteworthy that in such unpredictable environments, the expected rate of adaptation is actually higher in more maladapted populations (Supporting Information). Responses to selection increase the lag load all the more as fluctuations in the optimum are large along phenotypic directions where deviations from the optimum strongly reduce fitness (large $tr(\mathbf{S\Theta})$, see Fig. 1), causing strong potential for directional selection.

### *Autocorrelated fluctuations*

Let us now turn to autocorrelated fluctuations. Specifically, assume that the optimum is again a stationary Gaussian process with covariance matrix $\mathbf{\Theta}$, but autocorrelated in time. Autocorrelation is assumed to be positive and decrease geometrically (or exponentially) with the number of generations, which is a form of autoregressive process (Feller 1968; Box et al. 2008). Analogous to the univariate version in Lande and Shannon (1996), we define the autocovariance between two multivariate optima separated by $\tau$ generations as

$$E\left\{\boldsymbol{\theta}_t \boldsymbol{\theta}_{t-\tau}^T\right\} = \mathbf{\Theta} \exp(-|\tau| / T_\rho). \tag{6a}$$

The parameter $T_\rho > 0$ is a characteristic time for autocorrelation; for instance autocorrelation is half its maximum for two optima separated by $\tau = T_\rho \ln 2$ generations.

Figure 2 illustrates how changing $T_\rho$ affects patterns of fluctuations in the optimum. With $T_\rho$ on the scale of the generation time, fluctuations are almost unpredictable, resembling white noise (Fig. 2A). In contrast for large $T_\rho >> 1$ generations, the optimum changes in the same direction over many generations, such that the overall magnitude of fluctuations (quantified by $\mathbf{\Theta}$) is only observed on the long run (Fig. 2C). In the limit of very large $T_\rho$, changes in the optimum become similar to Brownian motion, with infinitely small displacement per generation. Hence, $T_\rho$ tunes the pattern of environmental fluctuations and their predictability, and is also directly connected to the speed of changes in the optimum. Over time intervals much shorter than $T_\rho$, the expected

speed of the optimum is (Appendix S2)

$$v_\theta = \sqrt{2tr(\boldsymbol{\Theta})/T_\rho}, \tag{6b}$$

so all things being equal, the optimum moves slower when the critical time $T_\rho$ increases (more autocorrelation).

The autocorrelation time influences the response of the mean phenotype to natural selection. Only when $T_\rho$ is sufficiently large can responses to selection accumulate over several generations, such that the mean phenotype of each trait is able to track large movements of its optimum (Fig. 1D). To what extent the mean of each trait is able to track its particular optimum also depends on the genetic correlations with other traits, and on how fast the optimum changes for these traits.

With autocorrelated fluctuations in the optimum, the expected lag load is

$$E_\rho\{L_s\} = \tfrac{1}{2}tr(\mathbf{S\Theta}(\mathbf{I} + T_\rho\mathbf{GS})^{-1}), \tag{7a}$$

where the subscript ρ is for autocorrelated optimum (Appendix S2). For a given **G** matrix, the load again increases with stronger fluctuating selection (large $tr(\mathbf{S\Theta})$). However, in contrast to white noise, with autocorrelated fluctuations in the optimum the lag load *decreases* with the overall potential to respond to selection, as indicated by the fact that the *inverse* of the matrix containing **GS** appears in (7a). This decrease in load occurs because autocorrelation increases the predictability of fluctuations, implying that the response to selection in one generation is likely to cause higher fitness in the next, as highlighted previously in a univariate context (Charlesworth 1993a,b; Lande and Shannon 1996). Equation (7a) can be generalized to the case where the optimum phenotype has different autocorrelation times for different traits, such that the predictability of directional selection varies among traits. In that case, the scalar $T_\rho$ in equation (7a) should simply be replaced by $\mathbf{T_\rho}$, the diagonal matrix harboring the autocorrelation times for the optima of each trait (Fig. S1, and eq. B7 in Supporting Information). Multivariate responses to selection more strongly reduce the lag load in environments that are more autocorrelated (because **GS** is multiplied by$T_\rho$). For large $T_\rho$ (strong autocorrelation), the lag load becomes

$$E_{\rho,T_\rho\to\infty}\{L_s\} = \frac{tr(\mathbf{G}^{-1}\boldsymbol{\Theta})}{2T_\rho}. \tag{7b}$$

In this regime, where the optimum moves very slowly (from eq. 6b), the lag load is not at all affected by the shape of the fitness landscape (continuous ellipses in Fig. 1), but instead only depends on the pattern of fluctuations in the optimum (dashed ellipse in Fig. 1). If the genetic covariance matrix is collinear with the matrix of variation in the optimum phenotype (shared eigenvectors), the lag load is minimum, for a given total genetic variance $tr(\mathbf{G})$, when the eigenvalues of **G** are proportional to those of **Θ**, such that **G** and **Θ** have the same correlation structure (dashed ellipse proportional to gray cloud in Fig. 1; see also Appendix S1).

## LAG LOAD CAUSED BY A DETERMINISTIC CHANGE

### *Sustained trend of change*

We will now investigate deterministic environmental change. The situation where a sustained trend causes the optimum to move at constant speed, with phenotypic direction and intensity given by the vector **k**, was investigated by Gomulkiewicz and Houle (2009), extending univariate results by Lynch and Lande (1993). They showed that the mean phenotype asymptotically moves at the same speed as the optimum, and in the same phenotypic direction, but at a constant distance behind it. The asymptotic phenotypic lag vector is $(\bar{\mathbf{z}}_t - \mathbf{k}t)_{t\to\infty} = (\mathbf{GS})^{-1}\mathbf{k}$. Inserted into eq. (3b), this gives the asymptotic deterministic lag load (Gomulkiewicz and Houle 2009),

$$L_{d,k} = \tfrac{1}{2}\mathbf{k}^T\mathbf{G}^{-1}\mathbf{S}^{-1}\mathbf{G}^{-1}\mathbf{k}. \tag{8}$$

Note that because **k** is a vector, $L_k$ is a quadratic form, and hence a scalar. Averaged over all possible trends of directional change **k**, this deterministic load is proportional to $tr((\mathbf{GS})^{-1}\mathbf{G}^{-1})$, so long-term fitness is higher with more potential to respond to selection (larger $tr(\mathbf{GS})$). However, for a given occurrence of deterministic trend (constant **k**), the lag load crucially depends on how the **G** matrix interacts with the direction of change (captured by **k**); an insightful geometric interpretation of this scenario is given by Duputié et al. (2012).

### *Abrupt shift in the optimum*

Another situation of interest is that where the optimum phenotype suddenly shifts to a new value, and then stays constant for a long time, as would occur following an abrupt change in abiotic conditions, or massive introduction of a competitor. This scenario, first investigated by Lande (1976b), and described as the "displaced optimum model" by Estes and Arnold (2007), has received considerable theoretical and empirical attention in the context of evolutionary rescue from potential extinctions caused by environmental stress (Gomulkiewicz and Holt 1995; Bell and Gonzalez 2009; Chevin and Lande 2010; Gomulkiewicz et al. 2010). Here although the fitness function is constant, the selection gradient does change in time as the population evolves toward the optimum. Hence, the displaced optimum model can be considered as a null scenario of cumulative evolution, for empirical studies where directional selection is found in one generation and the environment is thought to remain relatively constant (while a constant selection gradient is often assumed in this context).

The evolutionary dynamics can be analyzed easily in this scenario by setting the right member of equation (4b) to 0

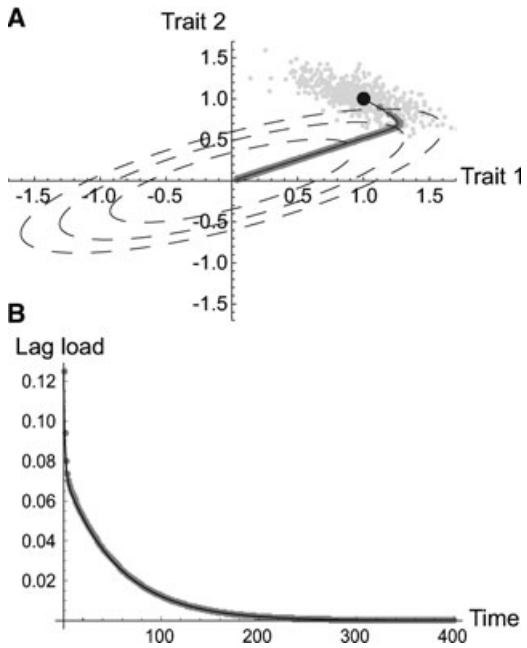


**Figure 3.** **Response to an abrupt environmental shift. The evolution of two traits in response to a sudden displacement in their optimum phenotype (A) is represented together with the dynamics of the corresponding lag load caused by deviations from the optimum (B). The optimum is displaced at generation 0 to a new value used as the origin for the graph in A. The mean phenotype is initially at a distance of 1 unit on each trait (large black dot in A). Because of the genetic constraints set by the interaction of the genetic covariance G matrix (materialized by the light gray dots) and the selection S matrix (isofitness represented by the dashed lines), the mean phenotype initially moves away from the optimum, but this still decreases the lag load. The continuous line in A is equation (9) with $\bar{\mathbf{z}}_0 - \theta = (1, 1)$, $\mathbf{G} = \begin{pmatrix} 0.6 & -0.3 \\ -0.3 & 0.3 \end{pmatrix}$ and $\mathbf{S} = \begin{pmatrix} 0.15 & -0.2 \\ -0.2 & 0.5 \end{pmatrix}$. For the continuous line in B, equation (9) was inserted into (3b) with the same parameters. Numerical recursions of (3b) and (4a) are also shown as dark gray dots.**

(no change in the optimum). This gives, for the vector of deviations of the mean phenotype from the optimum,

$$(\bar{\mathbf{z}}_t - \theta) = \exp(-\mathbf{GS}t)(\bar{\mathbf{z}}_0 - \boldsymbol{\theta}). \quad (9)$$

Note that the exponential in equation (9) is a function of a matrix, so it is itself a matrix that changes through time (Higham 2008). This becomes clearer when noting that $\exp(-\mathbf{GS}t) = \sum_i \exp(-\lambda_i t)\mathbf{U}_i$, where $\lambda_i$ is the $i$st eigenvalue of $\mathbf{GS}$, and $\mathbf{U}_i$ is the matrix obtained as the self outer product of the associated eigenvector. Hence, $\exp(-\mathbf{GS}t)$ is a linear combination of matrices, with coefficients that change in time at different exponential rates, given by the eigenvalues of $\mathbf{GS}$. The direction of phenotypic evolution thus also changes through time. Figure 3A illustrates this process with two traits. Because of the constraints set by the $\mathbf{G}$ matrix, the population initially seems to move away from the optimum, but eventually changes direction and moves toward it; this complex trajectory is fully captured by equation (9).

Inserting equation (9) into equation (3b) gives the dynamics of the lag load, an example of which is illustrated in Figure 3B. In this example, the lag load first decreases rapidly for about five generations, but then in a second phase decreases more slowly. That the dynamics may include several phases is a consequence of anisotropic selection on multiple traits (some combinations of traits are more strongly selected than others), combined with genetic correlations between traits; each phase in the dynamics of the lag load corresponds to a direction of phenotypic evolution in Figure 2A. Integrating this load over time produces the overall reduction in fitness incurred by maladaptation of the mean phenotype, along the complete trajectory toward a new optimum. This quantity is equivalent to the "cost of natural selection" introduced by Haldane (1957) for the substitution of a beneficial mutation (also termed substitution load). For multivariate quantitative traits, it has a remarkably simple expression (derived in Appendix S2),

$$L_{d,sh} = \tfrac{1}{4}\mathbf{x}_0^T \mathbf{G}^{-1}\mathbf{x}_0, \quad (10)$$

where $\mathbf{x}_0$ is the vector of initial displacement from the optimum (the subscript "sh" is for "shift"). This quantity offers a metric for how much the genetic constraints set by the $\mathbf{G}$ matrix affect the long-term fitness of a population, following an abrupt environmental shift. Under density-independent population growth, the ratio of the sizes of two populations having adapted to the same abrupt environmental shift, but with different $\mathbf{G}$ matrices, is just the exponential of the difference in their $L_{d,sh}$. Strikingly, this total load only depends on the inverse of the $\mathbf{G}$ matrix, and on the initial phenotypic direction (and intensity) of the shift in the optimum, not on the shape of the fitness landscape. Indeed, widening the fitness function reduces the initial cost of the environmental shift, but also slows down adaptation, and both effects exactly cancel out over the dynamics. Interestingly, this independence to the strength of selection parallels Haldane (1957)'s finding that the total number of selective deaths required for the substitution of a beneficial allele (cost of selection, or substitution load) does not depend on the selection coefficient, but only on the initial frequency of this allele. The accuracy of equation (10) was checked with numerical recursions similar to Figure 3 (not shown).

### *Punctuated equilibria*

A final scenario of interest is that where abrupt environmental shifts, as described above, occur episodically in the evolutionary history of a species. If the associated displacements of the optimum phenotype are well beyond those caused by the usual stochastic fluctuations in the optimum, such catastrophic events can cause major phenotypic changes, with the mean phenotype

moving to a new "adaptive zone." This process has been termed "quantum evolution" (Simpson 1944), or "punctuated equilibria" (Eldredge and Gould 1972). Despite some debate on its underlying mechanism and on the role of the environment in this process (Eldredge and Gould 1972; Charlesworth et al. 1982; Kirkpatrick 1982), recent analysis of rates of phenotypic evolution across multiple time scales and measurement methodologies indicate the robustness of this pattern, for body size traits at least (Uyeda et al. 2011).

The influence of the **G** matrix on adaptation under punctuated equilibria can be found by calculating the cumulated cost of natural selection. Consider abrupt environmental changes that are sufficiently rare that the population has reached evolutionary equilibrium before a new major change in the optimum occurs. If the distribution of adaptive zones in the long run is multivariate Gaussian, with mean 0 and covariance matrix **X**, then this is also the distribution of initial displacements of the optimum in equation (10). From the properties of quadratic forms, the expected net cost of natural selection with a given matrix **G**, after ξ changes in the optimum, is

$$L_{d,punct} = \frac{\xi}{4} tr(\mathbf{G}^{-1}\mathbf{X}). \quad (11)$$

This gives the overall reduction in fitness incurred by a species throughout its evolutionary history. Species with larger $L_{d,punct}$ are more likely, on average, to be exposed to extinction risk from abrupt environmental changes (Gomulkiewicz and Holt 1995).

## SIMULATIONS

The accuracy of the stochastic predictions, and the domains of validity of the approximations, were tested against quantitative genetic simulations (on Mathematica), where five genetically correlated traits adapt to a fluctuating optimum. For each set of parameters, the matrices of genetic covariance **G**, of stabilizing selection strengths **S**, and of environmental fluctuations **Θ**, were drawn at the beginning of each replicate from Wishart distributions, a family of random covariance matrices. These matrices were then scaled so they produced specified values for the potential for adaptation (or twice the standing load) $tr(\mathbf{GS})$, and for the expected intensity of fluctuating selection $tr(\mathbf{S\Theta})$. White noise fluctuations were simulated by drawing the optimum vector from a multivariate normal distribution with mean 0 and covariance matrix **Θ** at the beginning of each generation. Autocorrelated fluctuations were modeled as a first-order autoregressive process, where the optimum vector in a given generation depends on that in the previous generation. More precisely, $\boldsymbol{\theta}_t = \rho\boldsymbol{\theta}_{t-1} + \sqrt{1-\rho^2}\boldsymbol{\chi}$, where **χ** was randomly drawn from a multivariate normal distribution with mean 0 and covariance matrix **Θ**, and $\rho = \exp(-1/T_\rho)$ (from 6a). This sampling method produces a stationary distributed optimum, with covariance matrix **Θ** and autocorrelation ρ (Box et al. 2008). The response to selection in each generation was then computed using equation (3b), and the lag load using equation (4a), with a constant **G** matrix. The optimum and mean phenotype were both initiated at the average optimum (the null vector). The simulation then ran for 5000 generations, and the mean lag load was computed over the last 4000 generations, such that the influence of initial conditions was removed by a burn-in period of 1000 generations (chosen by visual inspection of the dynamics for the largest $T_\rho$, where the stationary phase is longest to reach).

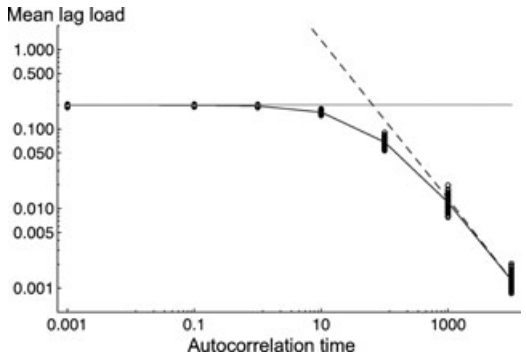


**Figure 4. Asymptotic lag load under fluctuating selection. The average lag load is shown as a function of the autocorrelation time $T_\rho$. Dots show the average over 4000 generations of simulations for 100 replicates. The black line is the prediction from equation (7a) averaged over replicates, whereas the dashed line is the approximation for high autocorrelation (eq. 7b). The gray line shows the approximation assuming no autocorrelation (white noise, eq. 5A). Parameters were drawn as in Figure 2.**

Figure 4 shows how the average lag load changes with autocorrelation in the optimum phenotype. The lag load decreases with increasing autocorrelation, as expected. Simulations agree well with the most general prediction from equation (7a). The expected lag load assuming no autocorrelation (white noise, eq. 5a) is a good approximation as long as the autocorrelation time is of the order of a generation, or less. In this range of autocorrelation times, and in the absence of any deterministic trend, long-term fitness is higher with more genetic constraints the response to selection. The approximation assuming strong autocorrelation (eq. 7b) is accurate for very large $T_\rho$, of the order of 1000. Recall that in that regime, only the pattern of fluctuations in the optimum (contained in **Θ**) affects long-term fitness for a given **G** matrix, while the orientation and width of the fitness peak (summarized in **S**) have no effect. Simulations with variable autocorrelation among traits, such that the predictability of the optimum phenotype changes from one trait to another, are presented in Figure S1; they conform well to the analytical prediction (B7) in the online Appendix S2.

Figure 5 shows the average lag load when combining environmental fluctuations to a deterministic trend of sustained change.

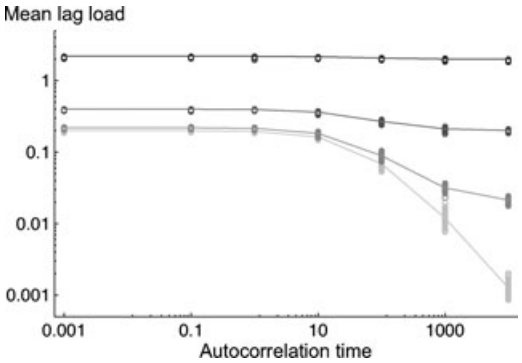


**Figure 5. Deterministic versus stochastic lag load. The expected lag load is shown as a function of the autocorrelation time $T_\rho$ for various rates of sustained deterministic change in the optimum. The vector of deterministic changes was drawn randomly at the beginning of each replicate, and scaled such that the expected asymptotic lag load is $L_{d,k} = 0$, $L_{d,k} = 0.02$, $L_{d,k} = 0.2$, or $L_{d,k} = 2$ (from light gray to black). Dots show for each of 100 replicates the average over 4000 generations of simulations. The lines show the prediction obtained by summing equations (7a) and (8), averaged over replicates. Other parameters were drawn as in Figure 4.**

At the beginning of each simulation, the vector **k** of directional change in the optimum phenotype was drawn randomly from a multivariate normal distribution with the identity as covariance matrix (no correlation), thus favoring no particular direction on average, and divided by a scalar so as to produce a given asymptotic deterministic lag load (from eq. 8). This deterministic trend of environmental change was then combined to autocorrelated fluctuations as described above. Also shown as continuous lines in Figure 5 are the predictions obtained by summing equations (7a) and (8) for each parameter set. Theoretical predictions fit the simulations well. Increasing the deterministic load while keeping the pattern of stochastic fluctuations unchanged reduces the relative impact of autocorrelation time on the expected lag load (darker curves are flatter in Fig. 5). Whether the deterministic trend contributes more or less than the stochastic fluctuations to the overall lag load does not directly depend on the relative speeds of deterministic versus stochastic changes in the optimum, $\sqrt{\mathbf{k}^T\mathbf{k}}/v_\theta$, but instead on the more complex comparison of equations (7a) and (8). Hence, when multiple correlated traits respond to selection in a changing optimum, the overall speed of changes in the optimum is less informative about long-term fitness than the directionality of these changes relative to the shape of the adaptive landscape and to the axes of major variation in the **G** matrix.

# *Discussion*

### THEORY OF GENETIC CONSTRAINTS ON ADAPTATION

Current predictions about how genetic correlations between traits constrain responses to natural selection are somewhat limited. With constant directional selection, it is best for a population to have all genetic variance along the direction of the selection gradient. However, this may hamper future responses to selection along other phenotypic directions. This consideration has motivated research aiming at understanding how the **G** matrix affects the rate of adaptation, taking into account the inherent variability of natural selection. Up to now, theory has focused on idealized distributions of selection gradients, where all traits—or combinations thereof—are equally likely to be under directional selection (isotropy) with a constant strength (Hansen and Houle 2008; Kirkpatrick 2009). Under these assumptions, genetic correlations have no effect, on average, on the expected rate of adaptation. But this expectation does not hold when selection gradients have arbitrary orientation and intensity, as shown here. Instead, the genetic correlation structure that maximizes the expected responses to selection depends on temporal correlations of directional selection between pairs of traits. However, higher rate of adaptation does not mean higher fitness, and whether a **G** matrix that facilitates responses to selection yields higher fitness in the long run depends on the pattern and predictability of environmental change, as illustrated here with models of a moving optimum phenotype. Constraints on adaptation should therefore be analyzed in reference to patterns of environmental changes.

The results above are for an effectively infinite population without random genetic drift, but their extension to finite population size is straightforward, if genetic drift does not change the **G** matrix much. Random genetic drift causes additional deviations of the mean phenotype from the optimum. Under drift and Gaussian optimizing selection, the mean phenotype in a population of effective size $N_e$ reaches a stationary Gaussian distribution, with mean $\bar{\boldsymbol{\theta}}$ (the mean optimum) and covariance matrix $1/(2N_e)\mathbf{S}^{-1}$ (i.e., with same shape as the fitness landscape, Lande 1976b, 1980b). Integrating mean fitness over this stationary distribution produces the expected drift load, which expressed in log fitness as in equation (3a) is

$$L_{drift} = \frac{m}{2}\ln\left(1+\frac{1}{2N_e}\right) \approx \frac{m}{4N_e}$$

(Lande 1980b), where $m$ is the number of independent linear combinations of traits under selection, and the approximation is for $N_e >> 1$. That the drift load increases with the number of selected traits is a consequence of the higher difficulty in "finding" the optimum in a phenotypic space with higher dimensionality (see also Tenaillon et al. 2007, Gros and Tenaillon 2009). This drift load, which does not depend on the **G** matrix, simply adds up to the other loads defined in equations (3–11). Note that $m$ is actually the minimum of the ranks (number of non-null eigenvalues) of the **G** and **S** matrices, which can be much lower than the apparent number of traits under selection $n$ (see Chevin et al. 2010b for consequences of mutation effects with $m << n$).

Several lines of empirical evidence suggest that the effective dimensionality of selection, mutation, and genetic variation, may often be quite low, typically below 10 (Martin and Lenormand 2006; Kirkpatrick 2009; but see Tenaillon et al. 2007). Therefore, in most populations of sizes above a few tens to hundreds of individuals (for which it makes sense to study adaptation), drift load of quantitative traits should not have a major effect on long-term fitness, except in essentially constant environments, and the fate of populations will therefore largely depend on their **G** matrix.

## TESTING THE PREDICTIONS

The parameters in most formulas provided here are amenable to empirical estimation. Multivariate quadratic regression can be used to estimate both the vector of linear selection gradients $\boldsymbol{\beta}$, and the matrix of quadratic selection gradients $\boldsymbol{\gamma}$, measuring how fitness changes with squares (or cross products) of deviations from the mean phenotype (Lande and Arnold 1983). Those can then be combined to compute the Hessian matrix defined in equation (1a), $\mathbf{H} = \boldsymbol{\gamma} - \boldsymbol{\beta}\boldsymbol{\beta}^T$ (Lande and Arnold 1983). However, the covariance matrix **B** of directional selection gradients cannot be obtained from each individual $\boldsymbol{\beta}$ estimated separately, because this approach would not account for the error variances in the estimates for each trait, and the error covariances of estimates among traits (Morrissey and Hadfield 2012 discussed the one-trait case). Instead, the covariance matrix of $\boldsymbol{\beta}$ should be estimated from the whole dataset (phenotypes and fitnesses of all individual for all time steps) using multivariate random regression, a class of mixed linear models widely used in animal breeding (Schaeffer 2004), and more recently in evolutionary quantitative genetics (e.g., Nussey et al. 2007; Brommer et al. 2012). Random regression would treat the selection gradients directly as multivariate random variables, and would estimate both the parameters of their distribution and a residual covariance matrix due to measurement error, without explicitly estimating each individual $\boldsymbol{\beta}$. (Note that a spatial covariance matrix **B** of selection gradients was recently used to analyze phenotypic divergence under sexual selection in the fruit fly [Chenoweth et al. 2010], testing predictions from a model by Zeng (1988), but without estimating **B** by random regression.) Given an estimated **B**, equation (2a) can be solved with chosen conditions on the **G** matrix (e.g., constant total variance $tr(\mathbf{G})$, or constant genetic variances but varying genetic correlations, as in Agrawal and Stinchcombe 2009) to find which genetic covariance matrix maximizes the expected rate of adaptation. Comparing this rate for the "best" **G** matrix to that for the measured **G** matrix gives a measurement of genetic constraints on the rate of adaptation under fluctuating selection. A confidence interval for rates of adaptation would be provided by the trace of the product of **G** with the error covariance matrix of **B**.

In most cases the position and movements of an optimum phenotype are not directly observed. However, assuming a Gaussian fitness function, it can be shown that $\mathbf{S} = -\mathbf{H}$. The selection gradient is thus $\boldsymbol{\beta} = -\mathbf{S}(\bar{\mathbf{z}} - \boldsymbol{\theta}) = \mathbf{H}(\bar{\mathbf{z}} - \boldsymbol{\theta})$, and the multivariate deviation of the mean phenotype from the optimum can be estimated as $(\bar{\mathbf{z}} - \boldsymbol{\theta}) = \mathbf{H}^{-1}\boldsymbol{\beta}$ (e.g., Estes and Arnold 2007). Assuming a constant shape of the fitness function (unchanged **H**), the temporal distribution of $\boldsymbol{\beta}$ can thus be used to estimate the covariance matrix of optima $\boldsymbol{\Theta}$, and predict the influence of the **G** matrix on the expected lag load (eqs. 5–7). In principle, the autocorrelation time of the optimum also can be estimated using procedures of time-series analysis such as autoregressive moving average (ARMA) fitting (Box et al. 2008). However, the data available to date make it difficult to reliably detect temporal variation in directional gradient even for a single trait, owing to the large error in the estimator of each gradient (Kingsolver and Diamond 2011; Morrissey and Hadfield 2012). This situation can only be improved by increasing sample size, and designing data collection or experiments specifically to optimize the required measurements.

## ON THE EVOLUTION OF G MATRICES

For simplicity I considered constant **G** matrices, which is a reasonable approximation at least in some situations (see section Model specification), but future work should investigate the evolution of the **G** matrix in similar contexts. With a stable **G** matrix evolving much slower than the mean, the lag load is likely to reach its asymptotic value conditional on **G**, as provided here. The **G** matrix may then evolve to reduce the total load, by second-order selection on epistatic modifiers of the covariance structure. Kawecki (2000) showed that the genetic variance of a single trait can evolve by indirect selection on canalization modifiers, and that lower genetic variance evolves in environments that fluctuate with low predictability. It was also shown theoretically that modifier genes can cause the evolution of genetic correlations and modularity among multiple traits (Wagner and Altenberg 1996; Wagner et al. 2005; Jones et al. 2007; Pavlicev et al. 2011). Alternatively, **G** may be unstable and dynamically evolve on the same timescale as the mean phenotype, without requiring additional modifier genes, as investigated theoretically in a variety of cases (Jones et al. 2003, 2004; Guillaume and Whitlock 2007). Such rapid changes in the **G** matrix will affect how the mean phenotype responds to selection in a changing environment, and invalidate some of the results presented here. An extreme example is when a gene of major effect transiently changes the **G** matrix and modifies the rate and direction of adaptation (Agrawal et al. 2001).

Which of these models (modifiers of a stable **G** matrix, vs flexible **G** matrix) is more accurate for **G** matrix evolution in a changing environment will largely depend on the genetic architecture of traits (linkage, pleiotropy, and epistasis of mutations), which is set by the molecular and physiological mechanisms underpinning the **G** matrix. Even under abundant mutation with

continuously distributed additive effects—a situation that should lead to a relatively stable covariance structure (Lande 1976a, 1980)—the stability of the **G** matrix should depend on whether (and to what extent) the pleiotropic effects of mutations vary from one locus to another. This question can be investigated using a recently developed model for the heterogeneity of mutation phenotypic effects across loci (Chevin et al. 2010b).

## Conclusion

I investigated how genetic correlations affect the rate of adaptation and long-term fitness in a changing environment. Other factors may also constrain responses to natural selection, such as lack of genetic variance (Kellermann et al. 2009), indirect genetic effects (Wolf et al. 1998), or sexually antagonistic trait expression/selection (Foerster et al. 2007). More fundamentally, standard predictions about evolutionary responses of quantitative traits to a changing environment may fail if the **G** matrix itself changes in direct response to the environment (i.e., without genetic change), as is increasingly documented (Wilson et al. 2006; Tonsor and Scheiner 2007; Husby et al. 2010; Husby et al. 2011). When such phenotypically plastic changes in the **G** matrix occur, accurate predictions can only be reached by explicitly modeling the genetic (co)variances of reaction norm parameters for multiple traits. Evolutionary dynamic predictions for the evolution of correlated plastic traits are all the more needed as both theoretical models (Chevin and Lande 2010; Chevin et al. 2010a; Reed et al. 2010) and data from natural populations (Ozgul et al. 2009; Ozgul et al. 2010; Pelletier et al. 2011) have recently highlighted the importance of phenotypic plasticity in the interaction between evolution and demography in changing environments.

### ACKNOWLEDGMENTS

I thank C. Teplitsky for discussions that inspired this project, J. D. Hadfield and M. B Morrissey for help with statistical issues, and S. J. Arnold, A. F. Agrawal, M. W. Blows, F. Blanquart, J. D. Hadfield, F. Massol, C. Teplitsky, and two anonymous reviewers for comments and suggestions that improved this manuscript. This work was supported by the grant "ContempEvol" from the "retour post-doctorants" program of the French Agence Nationale de la Recherche.

### LITERATURE CITED

Agrawal, A. F., and J. R. Stinchcombe. 2009. How much do genetic covariances alter the rate of adaptation? Proc. Biol. Sci. 276:1183–1191.

Agrawal, A. F., E. D. Brodie, and L. H. Rieseberg. 2001. Possible consequences of genes of major effect: transient changes in the G-matrix. Genetica 112:33–43.

Arnold, S. J., and M. J. Wade. 1984. On the measurement of natural and sexual selection—theory. Evolution 38:709–719.

Arnold, S. J., M. E. Pfrender, and A. G. Jones. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. Genetica 112:9–32.

Bell, G., and A. Gonzalez. 2009. Evolutionary rescue can prevent extinction following environmental change. Ecol. Lett. 12:942–948.

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

Boyce, W. E., and R. C. DiPrima. 2009. Elementary differential equations and boundary value problems. Wiley, New York.

Brommer, J. E., P. Kontiainen, and H. Pietiäinen. 2012. Selection on plasticity of seasonal life-history traits using random regression mixed model analysis. Ecol. Evol. 2:695–704.

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

Charlesworth, B. 1993a. Directional selection and the evolution of sex and recombination. Genet. Res. 61:205–224.

———. 1993b. 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.

Chenoweth, S. F., H. D. Rundle, and M. W. Blows. 2010. The contribution of selection and genetic constraints to phenotypic divergence. Am. Nat. 175:186–196.

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

Chevin, L. M., R. Lande, and G. M. Mace. 2010a. Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. PLoS Biol. 8:e1000357.

Chevin, L. M., G. Martin, and T. Lenormand. 2010b. Fisher's model and the genomics of adaptation: restricted pleiotropy, heterogeneous mutation and parallel evolution. Evolution 64:3213–3231.

Crow, J. F., and M. Kimura. 1970. An introduction to population genetics theory. Harper & Row, New York.

Duputié, A., F. Massol, I. Chuine, M. Kirkpatrick, and O. Ronce. 2012. How do genetic correlations affect species range shifts in a changing environment? Ecol. Lett. 15:251–259.

Eldredge, N., and S. J. Gould. 1972. Punctuated equilibria: an alternative to phyletic gradualism. Models Paleobiology 82:115.

Endler, J. 1986. Natural selection in the wild. Princeton Univ. Press, Princeton, NJ.

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.

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.

Fisher, R. A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. Trans. R. Soc. Edin. 52:35–42.

———. 1930. The genetical theory of natural selection. Oxford Univ. Press, Oxford, UK.

Foerster, K., T. Coulson, B. C. Sheldon, J. M. Pemberton, T. H. Clutton-Brock, and L. E. Kruuk. 2007. Sexually antagonistic genetic variation for fitness in red deer. Nature 447:1107–1110.

Frank, S. A., and M. Slatkin. 1992. Fisher's fundamental theorem of natural selection. Trends Ecol. Evol. 7:92–95.

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.

Gomulkiewicz, R., R. D. Holt, M. Barfield, and S. L. Nuismer. 2010. Genetics, adaptation, and invasion in harsh environments. Evol. Appl. 3:97–108.

Gros, P. A., H. Le Nagard, and O. Tenaillon. 2009. The evolution of epistasis and its links with genetic robustness, complexity and drift in a phenotypic model of adaptation. Genetics 182:277–293.

Guillaume, F., and M. C. Whitlock. 2007. Effects of migration on the genetic covariance matrix. Evolution 61:2398–2409.

Haldane, J. B. S. 1957. The cost of natural selection. J. Genet. 55:511–524.

Hansen, T. F., and D. Houle. 2008. Measuring and comparing evolvability and constraint in multivariate characters. J. Evol. Biol. 21:1201–1219.

Hansen, T. F., and E. P. Martins. 1996. Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. Evolution 50:1404–1417.

Hellmann, J. J., and M. Pineda-Krch. 2007. Constraints and reinforcement on adaptation under climate change: selection of genetically correlated traits. Biol. Conserv. 137:599–609.

Higham, N. J. 2008. Functions of matrices: theory and computation. Society for Industrial Mathematics, Philadelphia, PA.

Husby, A., D. H. Nussey, M. E. Visser, A. J. Wilson, B. C. Sheldon, and L. E. Kruuk. 2010. Contrasting patterns of phenotypic plasticity in reproductive traits in two great tit (Parus major) populations. Evolution 64:2221–2237.

Husby, A., M. E. Visser, and L. E. Kruuk. 2011. Speeding up microevolution: the effects of increasing temperature on selection and genetic variance in a wild bird population. PLoS Biol. 9:e1000585.

Jones, A. G., S. J. Arnold, and R. Burger. 2003. Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. Evolution 57:1747–1760.

———. 2004. Evolution and stability of the G-matrix on a landscape with a moving optimum. Evolution 58:1639–1654.

———. 2007. The mutation matrix and the evolution of evolvability. Evolution 61:727–745.

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

Kawecki, T. J. 2000. The evolution of canalization under fluctuating selection. Evolution 54:1–12.

Kellermann, V., B. van Heerwaarden, C. M. Sgro, and A. A. Hoffmann. 2009. Fundamental evolutionary limits in ecological traits drive *Drosophila* species distributions. Science 325:1244–1246.

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

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.

Kirkpatrick, M. 1982. Quantum evolution and punctuated equilibria in continuous genetic characters. Am. Nat. 119:833–848.

———. 2009. Patterns of quantitative genetic variation in multiple dimensions. Genetica 136:271–284.

Kirkpatrick, M., and N. H. Barton. 1997. Evolution of a species' range. Am. Nat. 150:1–23.

Kruuk, L. E. B., J. Slate, and A. J. Wilson. 2008. New answers for old questions: the evolutionary quantitative genetics of wild animal populations. Annu. Rev. Ecol. Evol. Syst. 39:525.

Lande, R. 1976a. The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genet. Res. 26:221–235.

———. 1976b. 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.

———. 1980a. The genetic covariance between characters maintained by pleiotropic mutations. Genetics 94:203–215.

———. 1980b. Genetic variation and phenotypic evolution during allopatric speciation. Am. Nat. 116:463–479.

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

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

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

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

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.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits, Sinauer Associates. Sunderland, MA, USA.

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.

Mathai, A. M., and S. B. Provost. 1992. Quadratic forms in random variables: theory and applications. Marcel Dekker, New York.

Maynard-Smith, J. 1976. What determines the rate of evolution? Am. Nat. 110:331–338.

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

Nussey, D. H., A. J. Wilson, and J. E. Brommer. 2007. The evolutionary ecology of individual phenotypic plasticity in wild populations. J. Evol. Biol. 20:831–844.

Ozgul, A., S. Tuljapurkar, T. G. Benton, J. M. Pemberton, T. H. Clutton-Brock, and T. Coulson. 2009. The dynamics of phenotypic change and the shrinking sheep of St. Kilda. Science 325:464–467.

Ozgul, A., D. Z. Childs, M. K. Oli, K. B. Armitage, D. T. Blumstein, L. E. Olson, S. Tuljapurkar, and T. Coulson. 2010. Coupled dynamics of body mass and population growth in response to environmental change. Nature 466:482-U485.

Pavlicev, M., J. M. Cheverud, and G. P. Wagner. 2011. Evolution of adaptive phenotypic variation patterns by direct selection for evolvability. Proc. R. Soc. Lond. B 278:1903–1912.

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

Pelletier, F., T. Clutton-Brock, J. Pemberton, S. Tuljapurkar, and T. Coulson. 2007. The evolutionary demography of ecological change: linking trait variation and population growth. Science 315:1571–1574.

Pelletier, F., K. Moyes, T. H. Clutton-Brock, and T. Coulson. 2012. Decomposing variation in population growth into contributions from environment and phenotypes in an age-structured population. Proc. R. Soc. Lond. B 279:394–401.

Polechova, J., N. Barton, and G. Marion. 2009. Species' range: adaptation in space and time. Am. Nat. 174:E186–E204.

Reed, T. E., R. S. Waples, D. E. Schindler, J. J. Hard, and M. T. Kinnison. 2010. Phenotypic plasticity and population viability: the importance of environmental predictability. Proc. R. Soc. Lond. B 277:3391–3400.

Revell, L. J. 2007. The G matrix under fluctuating correlational mutation and selection. Evolution 61:1857–1872.

Schaeffer, L. 2004. Application of random regression models in animal breeding. Livest. Prod. Sci. 86:35–45.

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., and R. Lande. 1976. Niche width in a fluctuating environment-density independent model. Am. Nat. 110:31–55.

Tenaillon, O., O. K. Silander J. P. Uzan, and L. Chao. 2007. Quantifying organismal complexity using a population genetic approach. PLoS ONE 2:e217.

Tonsor, S. J., and S. M. Scheiner. 2007. Plastic trait integration across a CO2 gradient in *Arabidopsis thaliana*. Am. Nat. 169:E119–E140.

Trenberth, K. E., and S. Josey. 2007. Observations: surface and atmospheric climate change. Pp. 747–845 *in* S. Solomon, ed. Climate change 2007: the physical science basis. Cambridge Univ. Press, Cambridge, U.K.

Turelli, M. 1988. Phenotypic evolution, constant covariances, and the maintenance of additive variance. Evolution 42:1342–1347.

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.

Wagner, G. P., and L. Altenberg. 1996. Perspective: complex adaptations and the evolution of evolvability. Evolution 50:967–976.

Wagner, G. P., J. Mezey, and R. Calabretta. 2005. Natural selection and the origin of modules. *In* W. Callebaut, and D. Rasskin-Gutman, eds. Modularity. Understanding the development and evolution of complex natural systems. The MIT Press, Cambridge, MA.

Wilson, A. J., J. M. Pemberton, J. G. Pilkington, D. W. Coltman, D. V. Mifsud, T. H. Clutton-Brock, and L. E. B. Kruuk. 2006. Environmental coupling of selection and heritability limits evolution. PLoS Biol. 4: 1270–1275.

Wolf, J. B., E. D. Brodie, J. M. Cheverud, A. J. Moore, and M. J. Wade. 1998. Evolutionary consequences of indirect genetic effects. Trends Ecol. Evol. 13:64–69.

Wright, S. 1969. Evolution and the genetics of populations. Volume 2: The theory of gene frequencies. The University of Chicago Press, Chicago.

Zeng, Z. B. 1988. Long-term correlated response, interpopulation covariation, and interspecific allometry. Evolution 42:363–374.

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: M. Blows

## *Supporting Information*

The following supporting information is available for this article:

**Appendix S1.** Individual traits and geometrical interpretations.

**Appendix S2.** Derivation of the lag load.

**Figure S1.** Asymptotic lag load with variable autocorrelation.

Supporting Information may be found in the online version of this article.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

# *Appendix 1: Individual traits and geometrical interpretations*

In this appendix, I provide elements of interpretation for the trace of a matrix product. These are applied here to the expected evolvability, but are valid for the trace of the product of any positive definite symmetric matrices (such as $\boldsymbol{\Theta}$, $\mathbf{S}$, or $\mathbf{G}$).

Let us denote E($e$) (for evolvability *sensu* Hansen and Houle 2008) the expected rate of adaptation in response to fluctuations in the selection gradient only, with no directional trend, that is (from eq. 2a), $\mathrm{E}(e) = tr(\mathbf{BG})$. This can be written in terms of variances and covariances of individual traits, by noting that the trace of the products of two matrices $\mathbf{A}$ and $\mathbf{B}$ is the sum of the element-wise products of $\mathbf{A}$ and $\mathbf{B}^T$. Hence $\mathrm{tr}(\mathbf{BG}) = \sum_{i,j} B_{ij} G_{ji}$ which, recalling that $\mathbf{B}$ and $\mathbf{G}$ are symmetric, leads to:

$$\mathrm{E}(e) = n\overline{B_{ii}G_{ii}} + 2n(n-1)\overline{B_{ij}G_{ij}} \qquad \text{(A1)}$$

where overbar denote average across traits. This can be further simplified by noting that $n\overline{B_{ii}} = \mathrm{tr}[\mathbf{B}]$, which is also the expectation of the squared norm of the selection gradient, $\|\boldsymbol{\beta}\|^2 = \boldsymbol{\beta}^T\boldsymbol{\beta}$ (by standard properties of quadratic forms). We thus have

$$\mathrm{E}(e) = \mathrm{E}(\|\boldsymbol{\beta}\|^2)\overline{G_{ii}} + n\mathrm{Cov}(B_{ii}G_{ii}) + 2n(n-1)\left[\overline{B_{ij}}\,\overline{G_{ij}} + \mathrm{Cov}(B_{ij}G_{ij})\right], \qquad \text{(A2)}$$

where E() denotes expectation over the distribution of gradients, while Cov() denotes covariance across traits. Formula (A2) relates the average rate of adaptive evolution to directly measurable aspects of genetic variation and natural selection. The first term in (A2) was previously found by Hansen and Houle (2008), and occurs even under isotropic (spherical) variation in the gradient, *i.e.* when directional selection occurs equivalently along all directions of the phenotypic space, on average. In that case, the mean selection response is proportional to (i) the average genetic variance of traits and (ii) the expected squared norm of the selection gradient (squared overall strength of selection). As might be expected, genetic covariances among traits are not involved here, because their effects cancel out on average under isotropic variation in $\boldsymbol{\beta}$.

The other terms in (A2) are effects of anisotropy, which were not included in Hansen and Houle (2008). The second term shows that the average rate of adaptation $\mathrm{E}(e)$ is higher when the genetic variance of traits covaries with variation in their selection gradient. This

occurs when traits for which directional selection fluctuates more across generations also have more genetic variance. The third term in (A2) gathers the effects of genetic correlations among traits, which often are the prime interest in empirical studies of **G** matrices. The first term in the square brackets shows that $\mathrm{E}(e)$ is higher when the average genetic correlation of traits has the same sign as the average correlation of directional selection among traits. The second term shows that $\mathrm{E}(e)$ is higher when pairs of traits with larger than average genetic covariances are more often selected jointly and in the same direction (and reciprocally).

While eq. (A2) offers an interpretation of eq. (2a) in terms of the directly measurable genetic variances and covariances of traits, it is sometimes more useful to project multivariate phenotypic evolution along independent linear combination of traits. This can be reached by diagonalizing the **G** matrix, which amounts to rotating the axes of the phenotypic space, allowing a geometrical interpretation of multivariate evolution (reviewed in Walsh and Blows 2009). Because the choice of traits to measure is always arbitrary to some extent, it can even be argued that the traits obtained this way are biologically more meaningful (Martin and Lenormand 2006). Variation in the selection gradient also can be decomposed into independent axes of variation.

Diagonalizing the **G** matrix results in $\mathbf{G} = \sum \lambda_i \mathbf{p}_i \mathbf{p}_i^T$, where the $\lambda_i$ are the (unranked) eigenvalues of **G**, and $\mathbf{p}_i$ is the eigenvector associated to the $i^{\text{th}}$ eigenvalue (note that the outer product $\mathbf{p}_i \mathbf{p}_i^T$ yields a matrix). Geometrically, this means that in the basis whose axes are the eigenvectors of **G**, the variances are given by the $\lambda_i$, and there are no covariances. Similarly, $\mathbf{B} = \sum \kappa_i \mathbf{q}_i \mathbf{q}_i^T$, with $\kappa_i$ the $i^{\text{th}}$ eigenvalues of **B**, and $\mathbf{q}_i$ the corresponding eigenvector. We can then write

$$\mathrm{E}(e) = tr(\mathbf{BG}) = \sum_{i,j} \kappa_i \lambda_j tr(\mathbf{q}_i \mathbf{q}_i^T \mathbf{p}_j \mathbf{p}_j^T). \qquad \text{(A3)}$$

This can be further simplified by noting that $\mathbf{q}_i^T \mathbf{p}_j = \cos(\alpha_{ij})$, with $\alpha_{ij}$ the angle between $\mathbf{q}_i$ and $\mathbf{p}_j$, such that $tr(\mathbf{q}_i \mathbf{q}_i^T \mathbf{p}_j \mathbf{p}_j^T) = \cos(\alpha_{ij}) tr(\mathbf{q}_i \mathbf{p}_j^T)$. The trace of the outer product of two vectors equals their scalar product, so $\mathrm{tr}(\mathbf{q}_i \mathbf{p}_j^T) = \mathbf{q}_i^T \mathbf{p}_j = \cos(\alpha_{ij})$. Defining the matrix **C** whose $(i,j)^{\text{th}}$ element is $\cos^2(\alpha_{ij})$, the average rate of adaptation can then be written as a bilinear form,

$$\mathrm{E}(e) = \boldsymbol{\kappa} \mathbf{C} \boldsymbol{\lambda}^T = \sum_{i,j} \kappa_i \lambda_j \cos^2(\alpha_{ij}), \qquad \text{(A4)}$$

where $\boldsymbol{\lambda}$ and $\boldsymbol{\kappa}$ are vectors that contain the eigenvalues of **G** and **B** respectively, in arbitrary order.

Equation (A4) has a simple interpretation. Let us first focus on the case where the inherent axes of genetic variation are the same as those of variation of the selection gradient. Since the eigenvalues are in arbitrary order, this collinearity condition is, without loss of generality, equivalent to $\cos^2(\alpha_{ii}) = 1$ for all *i*, and $\cos^2(\alpha_{ij}) = 0$ for all $i \neq j$, such that **C** is the identity matrix. We can then rewrite (A4) in a form similar to (A1), yielding

$$\begin{aligned} \mathrm{E}(e) &= \sum_{i,j} \kappa_i \lambda_j \\ &= n\,\bar{\kappa}\,\bar{\lambda} + n\mathrm{Cov}(\kappa_i, \lambda_i). \end{aligned} \tag{A5}$$

Recalling that the sum of eigenvalues equals the sum of diagonal terms of the original non-diagonalized matrix, $n\,\bar{\kappa}\,\bar{\lambda} = \mathrm{E}(\|\boldsymbol{\beta}\|^2)\overline{G_{ii}}$, so the first term in (A9) is the same as that in (A2). The second term is caused by anisotropy. It shows that with collinear axes of variation for genetics and selection, the mean rate of adaptation is higher if the directions with larger than average variation in strength of selection (contributing more than average to the selection gradient) also have larger than average genetic variance. This generalizes earlier arguments by Schluter (1996) about evolution along lines of least resistance in response to a constant selection gradients. Without collinearity between **G** and **B**, the expected rate of adaptation is necessarily lower, since $0 \leq \cos^2(\alpha) \leq 1$ for any $\alpha$.

# *Appendix 2: Derivation of the lag load*

### STOCHASTIC ENVIRONMENT

To find the expected lag load in a stochastic environment, it is useful to expand eq. (3b) into three quadratic forms,

$$L_\theta = \frac{1}{2}\boldsymbol{\theta}_t^T \mathbf{S}\boldsymbol{\theta}_t + \frac{1}{2}\bar{\mathbf{z}}_t^T \mathbf{S}\bar{\mathbf{z}}_t - \bar{\mathbf{z}}_t^T \mathbf{S}\boldsymbol{\theta}_t \,. \qquad \text{(B1),}$$

The first term is a quadratic form in Gaussian vectors, whose expectation is in all cases

$$\mathrm{E}\left(\frac{1}{2}\boldsymbol{\theta}_t^T \mathbf{S}\boldsymbol{\theta}_t\right) = \frac{1}{2} tr(\mathbf{S\Theta}) \,. \qquad \text{(B2)}$$

The other two terms involve $\bar{\mathbf{z}}$, which from eq. (4c) integrates values of $\boldsymbol{\theta}$ over multiple generations. Their expectations are

$$\mathrm{E}\left(\bar{\mathbf{z}}^T \mathbf{S}\boldsymbol{\theta}\right) = \mathbf{GS}\int \exp(-\mathbf{GS}\tau)^T \,\mathrm{E}(\boldsymbol{\theta}_{t-\tau}^T \mathbf{S}\boldsymbol{\theta}_t^T) d\tau \qquad \text{(B3)}$$

$$\mathrm{E}\left(\frac{1}{2}\bar{\mathbf{z}}^T \mathbf{S}\bar{\mathbf{z}}\right) = \frac{1}{2}\mathbf{GS}\int\int \exp(-\mathbf{GS}\tau_1)^T \,\mathrm{E}(\boldsymbol{\theta}_{t-\tau_1}^T \mathbf{S}\boldsymbol{\theta}_{t-\tau_2}^T) \exp(-\mathbf{GS}\tau_2) d\tau_1 d\tau_2 \,. \qquad \text{(B4)}$$

#### *Uncorrelated fluctuations*

When the optimum undergoes multivariate white noise, $\mathrm{E}(\boldsymbol{\theta}_{t-\tau}^T \mathbf{S}\boldsymbol{\theta}_t^T) = 0$ for all $\tau \neq 0$, so eq. (B3) cancels out, and only terms with $\tau_1 = \tau_2$ are kept in eq. (B4). Solving (B4) with these conditions gives

$$\mathrm{E}_w\left(\frac{1}{2}\bar{\mathbf{z}}^T \mathbf{S}\bar{\mathbf{z}}\right) = \frac{1}{2} tr[\mathbf{S\Theta}(\mathbf{GS})^2 (2\mathbf{GS})^{-1}] = \frac{1}{4} tr[\mathbf{GS}^2\mathbf{\Theta}] \,. \qquad \text{(B5)}$$

Combining (B5) with (B2), yields eq. (5a) in the main text.

Because the stationary distributions of $\bar{\mathbf{z}}$ and $\boldsymbol{\theta}$ are both multivariate normal, the selection gradient $\boldsymbol{\beta} = -\mathbf{S}(\bar{\mathbf{z}} - \boldsymbol{\theta})$ is also multivariate normally distributed, with mean $\mathbf{0}$ and covariance matrix $\mathbf{B} = \mathrm{E}[(\bar{\mathbf{z}} - \boldsymbol{\theta})^T \mathbf{S}^2 (\bar{\mathbf{z}} - \boldsymbol{\theta})]$. Note that $\mathbf{B}$ is similar to twice the expected lag load $L_\theta$, except that the selection matrix $\mathbf{S}$ is here squared. Hence combining eqs. (2a) and (5a) in the main text, the average rate of adaptation (or evolvability) is thus $\mathrm{E}_w\left\{\Delta \ln \overline{W}\right\} = tr[\mathbf{GS}^2\mathbf{\Theta}(\mathbf{I} + \mathbf{GS}/2)]$, which increases with $\mathrm{E}_w(L_s)$. This means that more adaptive evolution takes place in populations that are on average more maladapted, which is a generic property of adaptation to an unpredictable fluctuating optimum.

*Autocorrelated fluctuations*

With autocorrelated noise, both (B3) and (B4) include terms due to correlations between the optima in different generations. From the autocorrelation function in eq. (6a), $\mathrm{E}(\boldsymbol{\theta}_{t-\tau}^T \mathbf{S} \boldsymbol{\theta}_t^T) = tr(\mathbf{S}\boldsymbol{\Theta}) \exp(-|t|/T_\rho)$. Inserting into equations (B3) and (B4) and integrating gives

$$\mathrm{E}_\rho(\bar{\mathbf{z}}^T \mathbf{S} \boldsymbol{\theta}) = \mathrm{E}_\rho(\bar{\mathbf{z}}^T \mathbf{S} \bar{\mathbf{z}}) = tr[\mathbf{G}\mathbf{S}^2\boldsymbol{\Theta}(\mathbf{G}\mathbf{S} + \frac{1}{T_\rho}\mathbf{I})^{-1}] \qquad \text{(B6)}$$

Combining with equations (B1) and (B2), this simplifies after some algebra to eq. (7a) in the main text.

The expected speed of the optimum under autocorrelated noise can be found by studying the random variable $\boldsymbol{\Delta}_\tau \boldsymbol{\theta} = \boldsymbol{\theta}_{t+\tau} - \boldsymbol{\theta}_t$, the vector of changes in the multitrait optimum phenotype over time interval $\tau$. The squared Euclidian distance moved by the optimum over this time frame is simply the scalar product of $\boldsymbol{\Delta}_\tau \boldsymbol{\theta}$ with itself, whose expectation is $\mathrm{E}(\boldsymbol{\Delta}_\tau \boldsymbol{\theta} \boldsymbol{\Delta}_\tau \boldsymbol{\theta}^T) = \mathrm{E}(\boldsymbol{\theta}_t \boldsymbol{\theta}_t^T) + \mathrm{E}(\boldsymbol{\theta}_{t+\tau} \boldsymbol{\theta}_{t+\tau}^T) - 2\mathrm{E}(\boldsymbol{\theta}_t \boldsymbol{\theta}_{t+\tau}^T)$. From the autocorrelation function in eq. (6a), this is $2tr(\boldsymbol{\Theta})[1 - \exp(-\tau/T_\rho)]$. The expected squared speed of moves in the optimum is thus $\mathrm{E}(\boldsymbol{\Delta}_\tau \boldsymbol{\theta} \boldsymbol{\Delta}_\tau \boldsymbol{\theta}^T)/\tau \approx 2tr(\boldsymbol{\Theta})/T_\rho$, where the approximation is for time intervals much shorter than the critical autocorrelation time, $\tau << T_\rho$. Taking the square root of this expression gives the speed of the optimum in (6b).

The expected lag load can be generalized to the situation where autocorrelation in the optimum varies among traits, such that directional selection does not have the same predictability for different traits. Assume that autocorrelation in the optimum still decreases exponentially in time (autoregressive process), but at different rates for different traits, while the cross-correlation among traits is a weighted product of their autocorrelations. More precisely, consider that the autocovariance matrix between two multivariate optima separated by $\tau$ generations is $\mathrm{E}\{\boldsymbol{\theta}_t \boldsymbol{\theta}_{t-\tau}^T\} = \mathbf{R}_\tau^T \boldsymbol{\Theta} \mathbf{R}_\tau$, where $\mathbf{R}_\tau = \exp(-\frac{1}{2}|\tau|\mathbf{T}_\rho^{-1})$, and $\mathbf{T}_\rho$ is a diagonal matrix whose entries are the autocorrelation times of each trait, the inverse of rate of exponential decrease of autocorrelation. Over $\tau$ generations, the autocorrelation in the optimum for a given trait *i* then is $\exp(-|\tau|/T_{\rho,i})$, where $T_{\rho,i}$ is the $i^{\text{th}}$ diagonal element of the matrix $\mathbf{T}_\rho$. The temporal cross-correlation between the optima for two traits *i* and *j* is

$\exp[-\frac{1}{2}|\tau|(1/T_{\rho,i}+1/T_{\rho,j})]$, that is, their cross-correlation time is the harmonic mean of their autocorrelation times. Inserting into equations (B3-4), and integrating over the time lag $\tau$, yields the expected lag load, which after some rearrangement simplifies to

$$\mathrm{E}_{\rho}\{L_s\} = \frac{1}{2} tr\left(\mathbf{S\Theta(I+GST_{\rho})}^{-1}\right). \qquad \text{(B7)}$$

Equation (B7) was checked by simulation. The program was similar to that described in the main text, except for the autocorrelations in the optimum phenotype. At the beginning of each simulation, an autocorrelation time was drawn for each trait from an exponential distribution with mean $\overline{T_{\rho}}$. These trait-specific autocorrelation times were placed in a diagonal matrix $\mathbf{T}_{\rho}$, used to compute the matrix $\mathbf{R}_1 = \exp(-\frac{1}{2}\mathbf{T}_{\rho}^{-1})$ (taking the exponential of diagonal terms only, by definition of the matrix exponential). The optimum at generation *t* was then computed as $\boldsymbol{\theta}_t = \mathbf{A}\boldsymbol{\theta}_{t-1} + \mathbf{B}\boldsymbol{\chi}$, with $\boldsymbol{\chi}$ a random vector drawn from a multivariate normal distribution with mean 0 and covariance matrix $\boldsymbol{\Theta}$. The two matrices **A** and **B** were chosen so that the covariance matrix of $\boldsymbol{\theta}_t$, and its temporal cross covariance matrix, both conform to their definitions, that is, $\mathrm{E}\{\boldsymbol{\theta}_t\boldsymbol{\theta}_t^T\} = \boldsymbol{\Theta}$ and $\mathrm{E}\{\boldsymbol{\theta}_t\boldsymbol{\theta}_{t-\tau}^T\} = \mathbf{R}_{\tau}^T\boldsymbol{\Theta}\mathbf{R}_{\tau}$ as defined above. It can be shown with some algebra that this is verified if $\mathbf{A} = \mathbf{R}_1\boldsymbol{\Theta}\mathbf{R}_1\boldsymbol{\Theta}^{-1}$ and $\mathbf{B} = (\mathbf{I}-\mathbf{QQ}^T)^{1/2}$, where $\mathbf{Q} = \boldsymbol{\Theta}^{-1/2}\mathbf{R}_1\boldsymbol{\Theta}\mathbf{R}_1(\boldsymbol{\Theta}^{-1/2})^T$ (some superscripts 'T' were omitted, making use of the fact that $\mathbf{R}_{\tau}$ equals its transpose, since it is a diagonal matrix). The asymptotic lag load was then computed for each simulation as detailed in the main text, and compared to equation (B7), for different mean autocorrelation times $\overline{T_{\rho}}$. The results are shown in Figure S1. Equation (B7) (black line in Fig. S1) more accurately captures simulation results than its counterpart that assumes same autocorrelation for all traits (eq. 7a, gray line in fig. S1).

## SUDDEN ENVIRONMENTAL SHIFT

After a sudden environmental shift, combining equations (3b) and (9), the lag load is

$$L_{\theta} = \frac{1}{2}\mathbf{x}_0^T\left(\int \exp(-\mathbf{GS}t)^T\,\mathbf{S}\exp(-\mathbf{GS}t)dt\right)\mathbf{x}_0. \qquad \text{(B8)}$$

The integral in (B8) equals $\mathbf{S}(2\mathbf{GS})^{-1} = \frac{1}{2}\mathbf{G}^{-1}$ which, inserted into (B8), yields (10) in the main text.

**Supplementary Figure S1: Asymptotic lag load with variable autocorrelation.** The average lag load is shown for simulations where the optimum of each trait has its own autocorrelation time, such that the environmental predictability of selection varies among traits. The diagonal matrix of autocorrelation times $\mathbf{T}_\rho$ was created at the beginning of each simulation, with each element in the diagonal drawn from an exponential distribution with mean $\overline{T_\rho}$. This mean autocorrelation time $\overline{T_\rho}$ was taken in multiples of 10, from $10^{-3}$ to $10^4$, and 100 replicate simulations were carried out for each $\overline{T_\rho}$. Dots show the average over 4000 generations for each simulation, and their abscissa is $tr(\mathbf{T}_\rho)/n$, the realized mean autocorrelation time for each simulation. The black line is the prediction from eq. (B7) in Appendix 2 averaged over replicates, with $\overline{T_\rho}$ as abscissa. The gray line shows the prediction assuming same autocorrelation time $\overline{T_\rho}$ for all traits (eq. 7).