SUPPLEMENTARY ANALYSIS

Variance of the effective optimum

By definition, the variance in the per-generation effective optimum is

$$\mathrm{Var}[\theta_{tot}] = Var[\sum_{k=1}^{n} e_k\theta_k] = \sigma_\theta^2 \sum_{k=1}^{n} e_k^2 + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} cov\left[e_i\theta_i, e_j\theta_j\right], \quad \text{(A1)}$$

where σ_θ^2 is the stationary variance of the per-episode optimum. Rewriting (A1) using the properties of the autocorrelation of the environment (eq. 11) leads to

$$\mathrm{Var}[\theta_{tot}] = \sigma_\theta^2(\sum_{k=1}^{n} e_k^2 + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} e_i e_j \rho^{(j-i)\alpha T/n}), \quad \text{(A2)}$$

with $e_i = \omega_{tot}^2/\omega_i^2$. The first term in the right-hand side of equation (A2) corresponds to the direct contribution of the variance of the environment. The second term on the right-hand side results from the autocorrelation of the environment between selection episodes within a generation, generating additional variation to the effective optimum. With the further assumption that the strength of selection is the same at all episodes of selection, $\omega_i^2 = \omega^2$, Equation A2 further simplifies in

$$\mathrm{Var}[\theta_{tot}] = \sigma_\theta^2\left(\frac{1}{n} + \frac{2}{n}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\rho_\theta^{(j-i)\alpha T/n}\right) = \sigma_\theta^2\frac{n - n\rho^{\,2\alpha T/n} + 2\rho^{\,\alpha T/n}(\rho^{\,\alpha T} - 1)}{n^2(\rho^{\,\alpha T/n} - 1)^2} \quad \text{(A3)}$$

The variance of the effective optimum θ_{tot} is always lower than the variance of the environment (set to 1 in figure A1), and decreases with increasing number of selection episodes (fig A1-A). However, $\mathrm{Var}[\theta_{tot}]$ does not decrease without bounds, instead reaching a limit value for a large number of episodes of selection, $\lim_{n\to\infty}(\mathrm{Var}[\theta_{tot}]) = 2\frac{\psi - \ln\psi - 1}{(\ln\psi)^2}$, which only depend on the autocorrelation of the environment over the time window for selection $\psi = \rho^{\,\alpha T}$. In addition, for a given number of selection episodes, the variance of the effective optimum increases with the autocorrelation of the environment (fig. A1-B).

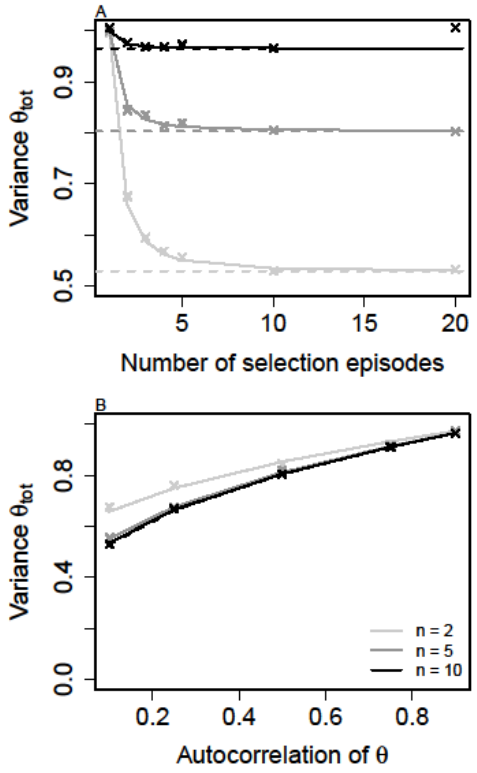

Figure A1: Variance of the effective optimum as a function of the number of selection episodes (A) and as a function of the autocorrelation of the environment (B). Crosses represent the results from simulations where the strength of selection is drawn from a probability distribution as described in fig. 2 caption. Full lines represent the expectation from eq. A3 and dashed lines represent the limit for a large number of selection episodes. Panel A: from light grey to dark $\rho_{\theta,T} = 0.2$, 0.5 and 0.9 respectively. Panel B: the gray scale represents different number of selection episode per generation (see within the panel). For both panels, $\bar{\omega}_{tot}^2 = 25$, $\sigma_\theta^2 = 1$, $\alpha = 1$.

Variance of the selection gradient

The variance of β_{tot} is proportional to the variance of the distance of the mean phenotype to the optimum

$$\mathrm{Var}[\beta_{tot}] = (S_{tot})^2 \mathrm{Var}[(\bar{z} - \theta_{tot})] = (S_{tot})^2 \mathrm{E}[(\bar{z} - \theta_{tot})^2]. \quad \text{(A4)}$$

By rearranging equation 10 in the main text (Charlesworth 1993), the mean phenotypic value at generation g is

$$\bar{z}_g = (1 - S_{tot}G_0)^g \bar{z}_0 + S_{tot}G_0 \sum_{i=1}^{g} (1 - S_{tot}G_0)^{i-1} \theta_{tot,g-i}. \quad \text{(A5)}$$

After a sufficiently large number of generations, the first term in equation (A5), corresponding to the initial conditions, can be neglected. We thus obtain

$$\mathrm{E}[(\bar{z} - \theta_{tot})^2] = \mathrm{E}\left[\left(S_{tot}G_0 \sum_{i=1}^{g} (1 - S_{tot}G_0)^{i-1} \theta_{tot,g-i} - \theta_{tot,g}\right)^2\right], \quad \text{(A6)}$$

By developing terms, using the autocorrelation of the effective optimum (eq. 12) and assuming that $g \to \infty$, we obtain equation (16) in the main text (see also Charlesworth 1993; eq. 19a)

$$\mathrm{Var}[\beta_{tot}] = 2\mathrm{Var}[\theta_{tot}] S_{tot}^2 \frac{1-\rho\left(1+S_{tot}G_0(\varphi-1)\right)}{(2-S_{tot}G_0)(1-\rho(1-S_{tot}G_0))}.$$

Autocorrelation of the selection gradient

The autocovariance of the selection gradient over τ generations is

$$\mathrm{Cov}\left[\beta_{tot,g}, \beta_{tot,g-\tau}\right] = S_{tot}^2\left(\mathrm{E}\left[\theta_{tot,g}\theta_{tot,g-\tau}\right] + \mathrm{E}\left[\bar{z}_g\bar{z}_{g-\tau}\right] - \mathrm{E}\left[\bar{z}_{g-\tau}\theta_{tot,g}\right] - \mathrm{E}\left[\theta_{tot,g-\tau}\bar{z}_g\right]\right). \quad \text{(A7)}$$

Continuous time approach

Following Lande and Shannon (1996), the dynamics can be approximated in continuous time. The continuous time approximation eases the analysis. After long enough time so that the we can neglect the initial conditions, the mean phenotype at time *t* is

$$\bar{z}_t = S_{tot}G_0 \int_0^{\infty} e^{-S_{tot}G_0 x} \theta_{tot,t-x} dx. \quad \text{(A8)}$$

Equation (A8) is the continuous time equivalent to equation (A6) once the effect of initial conditions has been erased. From equations (A7) and (A8), the method to derive the

autocorrelation of the selection gradient is described in Chevin and Haller (2014, appendix for the autoregressive optimum). This approach cannot a priori be used in the context of the present model because the inflation factor is 1 for $\tau = 0$ and is not defined for $0 < \tau < 1$. However, we found that the naïve approach to use equation (12) in the main text directly in the derivations from Chevin and Haller (2014) matches closely the exact discrete time approach that we develop in the next section. As we could not demonstrate this observation formally, we chose not to present this result in our main text. We nevertheless provide this result here as a heuristic:

$$\rho_{\beta tot,\tau} = \varphi \frac{\rho^{\tau T} - S_{tot} G_0 \frac{T_\theta}{T} e^{-S_{tot} G_0 \tau}}{1 - S_{tot} G_0 \frac{T_\theta}{T}}, \quad \text{(A9)}$$

where $\frac{T_\theta}{T}$ measures the characteristic timescale of the autocorrelation of the environment in units of generation time.

Discrete time approach

In contrast with previous studies (Lande and Shannon 1996, Chevin and Haller 2014), we calculate the autocorrelation of the selection gradient in discrete time. Similarly to the calculation for the variance of the selection gradient, we use equation (A5) assuming that many generations occurred so that the first term can be neglected. We also use equation (12) in the main text for the autocorrelation of the effective optimum $\rho_{\theta tot,\tau} = \rho^{\tau T} \varphi$, reminding that $\varphi = 1$ for $\tau = 0$. The general procedure follows Chevin and Haller (2014, appendix for the autoregressive optimum) by calculating each expectation from equation (A7). Even though we kept notations concise, each sum in the expectations below can be resolved exactly. The first expectation is straightforward

$$\mathrm{E}\left[\theta_{tot,g} \theta_{tot,g-\tau}\right] = \varphi \mathrm{Var}[\theta_{tot}] \rho^{\tau T}. \quad \text{(A10)}$$

The expectation for the first cross product is

836 $\mathrm{E}\left[\bar{z}_{g-\tau}\theta_{tot,g}\right] = S_{tot}G_0 \sum_{i=1}^{g-\tau}(1 - S_{tot}G_0)^{i-1}E\left[\theta_{tot,g}\theta_{tot,g-i-\tau}\right]$, taking into account the

837 discontinuity of the inflation factor I_f in $i = g - \tau$, it results in

838 $E\left[\bar{z}_{g-\tau}\theta_{tot,g}\right] = \varphi\mathrm{Var}[\theta_{tot}]S_{tot}G_0 \sum_{i=1}^{g-\tau-1}(1 - S_{tot}G_0)^{i-1}\rho^{(\tau+i)T} + S_{tot}G_0(1 -$

839 $S_{tot}G_0)^{g-\tau-1}\mathrm{Var}[\theta_{tot}]$. (A11)

840 The expectation for second cross product is

841 $E\left[\theta_{tot,g-\tau}\bar{z}_g\right] = S_{tot}G_0 \sum_{i=1}^{g}(1 - S_{tot}G_0)^{i-1}E(\theta_{tot,g-i}\theta_{tot,g-\tau})$.

842 We need to distinguish 3 cases, $i < \tau$, $i > \tau$ and $i = \tau$:

843 $E\left[\theta_{tot,g-\tau}\bar{z}_g\right] = \varphi\mathrm{Var}[\theta_{tot}]S_{tot}G_0 \left(\sum_{i=1}^{\tau-1}(1 - S_{tot}G_0)^{i-1}\rho^{(\tau-i)T} + \sum_{i=\tau+1}^{g}(1 -$

844 $S_{tot}G_0)^{i-1}\rho^{(i-\tau)T} + \varphi^{-1}(1 - S_{tot}G_0)^{\tau-1}\right)$. (A12)

845 Lastly,

846 $E\left[\bar{z}_g\bar{z}_{g-\tau}\right] = (S_{tot}G_0)^2 \sum_{i=1}^{g}\sum_{k=1}^{g-\tau}(1 - S_{tot}G_0)^{i+k-2}E\left[\theta_{tot,g-i}\theta_{tot,g-\tau-k}\right]$.

847 To calculate the double sum, it is convenient to perform a change of variables

848 $\sum_{i=1}^{g}\sum_{k=1}^{g-\tau}(1 - S_{tot}G_0)^{i+k-2}E\left[\theta_{tot,g-i}\theta_{tot,g-\tau-k}\right] = \sum_{i=1}^{g}\sum_{j=\tau+1}^{g}(1 -$

849 $S_{tot}G_0)^{i+j-\tau-2}E\left[\theta_{tot,g-i}\theta_{tot,g-j}\right]$.

850 We then need to distinguish 4 cases,

851 for $i < \tau+1$, $A = \varphi\mathrm{Var}[\theta_{tot}] \sum_{i=1}^{\tau}\sum_{j=\tau+1}^{g}(1 - S_{tot}G_0)^{i+j-\tau-2}\rho^{(j-i)T}$.

852 for $i > \tau$ and $i = j$, $B = \mathrm{Var}[\theta_{tot}] \sum_{i=\tau+1}^{g}(1 - S_{tot}G_0)^{2i-\tau-2}$.

for $i > \tau$ and $j > i$, $C = \varphi \mathrm{Var}[\theta_{tot}] \sum_{j=\tau+2}^{g} \sum_{i=\tau+1}^{j-1} (1 - S_{tot} G_0)^{i+j-\tau-2} \rho^{(j-i)T}$.

for $i > \tau$ and $i > j$, $D = \varphi \mathrm{Var}[\theta_{tot}] \sum_{i=\tau+2}^{g} \sum_{j=\tau+1}^{i-1} (1 - S_{tot} G_0)^{i+j-\tau-2} \rho^{(i-j)T}$,

with C = D, and resulting in

$$E\left[\bar{z}_g \bar{z}_{g-\tau}\right] = (S_{tot} G_0)^2 (A + B + 2C). \quad \text{(A13)}$$

Adding equations (A10) to (A13), finding the limit of this sum for large *g*, dividing by the variance of the selection gradient (eq. 16) and rearranging leads to the autocorrelation of the selection gradient

$$\rho_{\beta tot,\tau} = \frac{\varphi \varepsilon \rho^{\tau T} + \gamma (1 - S_{tot} G_0)^{\tau}}{\vartheta}, \quad \text{(A14)}$$

where,

$$\varepsilon = (\rho^T - 1)^2 (2 - S_{tot} G_0),$$

$$\gamma = \frac{S_{tot} G_0}{(1 - S_{tot} G_0)} \left(\rho^T (1 - \varphi) (S_{tot} G_0)^2 - (1 - \rho^T) \left(1 + \rho^T (2\varphi - 1)\right) (1 - S_{tot} G_0) \right),$$

$$\vartheta = 2(\rho^T - 1 + S_{tot} G_0)(\rho^T - 1 + \rho^T (\varphi - 1) S_{tot} G_0).$$

Equation (A14) is written in a form similar to equation (A9), decomposing the effects of the autocorrelation of the environment and of the deviations of the mean phenotype to the optimum. An interpretation for equation (A14) is proposed in the main text.

In the special case where there is a unique episode of selection per generation, i.e. $\varphi = 1$, equation (A14) is

$$\rho_{\beta tot,\tau} = \frac{\rho_{\theta}^{\tau T} - \kappa (1 - S_{tot} G_0)^{\tau}}{1 - \kappa}, \quad \text{(A15)}$$

where $\kappa = \frac{(1+\rho^T)S_{tot}\,G_0}{(1-\rho^T)(2-S_{tot}\,G_0)}$.

Equation (A15) is very similar to equation 5b in Chevin and Haller (2014). Comparing equation (A14) and (A15) shows that the inflation factor, i.e. the autocorrelation of selection within generation, complicates the analysis by introducing new terms at the interface between the autocorrelation of the environment and the dynamics of selection within and across generations. A Taylor expansion performed on the reciprocals of equation (A15) and 5b in Chevin and Haller (2014) shows that both equations are identical to at least the second order. This result can be confirmed by representing both functions for different autocorrelation of the environment. Both functions match exactly on most of the autocorrelation range with minor deviation at very weak autocorrelation (not shown).

Changes in (co)variances under selection

Here we derive the changes in the variances and covariances of phenotypic components after each selection episode. Following classical quantitative genetics, we assume that the phenotype z of any individual prior to any selection in a generation is the sum of two normally distributed components: the breeding value x and environmental effect e. At the beginning of a generation, these phenotypic components have means $\bar{x}$ and $\bar{e} = 0$, variances G and E, and are uncorrelated, such that their covariance is $C_{x,e} = 0$ and the total phenotypic variance is $P = G + E$. Changes in the mean breeding value under selection are obtained from changes in mean phenotype using the regression of breeding values on phenotypes, which at the beginning of a generation has slope $\frac{\mathrm{Cov}(x,z)}{P} = \frac{G}{G+E} = h^2$, following the usual notation for heritability.

We consider episodes of selection caused by a Gaussian fitness peak, such that individuals with phenotype z have fitness $(z) \propto \exp(-\frac{(z-\theta)^2}{2\omega^2})$ (where we omit indices for selection episodes and generation for simplicity). It is informative to start by investigating the first selection episode in a generation, where x and e are independent and $\bar{e} = 0$, $\bar{z} = \bar{x}$. The fitness function on breeding values is obtained by integrating over the distribution of environmental effects, $\widetilde{W}(x) \propto \exp\left(-\frac{(x-\theta)^2}{2(\omega^2+E)}\right)$, and the fitness function on environmental effects is obtained similarly by integrating over breeding values, $W_e(e) \propto \exp(-\frac{(\bar{x}+e-\theta)^2}{2(\omega^2+G)})$. The fitness function on breeding values (respectively, on environmental effects) is thus also Gaussian, with squared width $\omega^2 + V_e$ (respectively $\omega^2 + G$) and optimum θ (respectively $\theta - \bar{x}$). With a Gaussian fitness peak, the distribution of phenotypes, breeding values and environmental effects after selection remain Gaussian. Denoting the per-episode strength of stabilizing selection as $S = \frac{1}{\omega^2+P}$, the means after selection are

$$\bar{z}^* = \theta + \frac{\omega^2}{\omega^2+P}(\bar{z} - \theta) = \theta + (1 - SP)(\bar{z} - \theta) = \bar{z} - SP(\bar{z} - \theta) \quad \text{(A16a)}$$

$$\bar{x}^* = \theta + \frac{\omega^2+V_e}{\omega^2+P}(\bar{x} - \theta) = \theta + (1 - SG)(\bar{x} - \theta) = \bar{z} - SG(\bar{z} - \theta) \quad \text{(A16b)}$$

$$\bar{e}^* = \theta - \bar{x} + \frac{\omega^2+G}{\omega^2+P}(\bar{x} - \theta) = \theta - \bar{x} + (1 - SE)(\bar{x} - \theta) = -SE(\bar{z} - \theta)\,, \quad \text{(A16c)}$$

where we have used the fact that $\bar{e} = 0$ and thus $\bar{z} = \bar{x}$ prior to any selection. Note that $\bar{x}^* - \bar{x} = h^2(\bar{z}^* - \bar{z})$ and $\bar{e}^* - \bar{e} = (1 - h^2)(\bar{z}^* - \bar{z})$, as required. The variances after selection are

$$P^* = \left(\frac{1}{\omega^2} + \frac{1}{P}\right)^{-1} = (1 - SP)P \quad \text{(A17a)}$$

$$G^* = \left(\frac{1}{\omega^2+E} + \frac{1}{G}\right)^{-1} = (1 - SG)G \quad \text{(A17b)}$$

$$E^* = \left(\frac{1}{\omega^2+G} + \frac{1}{E}\right)^{-1} = (1 - SE)E, \quad \text{(A17c)}$$

which satisfy the relationships $G^* - G = h^4(P^* - P)$ and $E^* - E = (1 - h^2)^2(P^* - P)$. The phenotypic variance after selection must also satisfy

$$P^* = G^* + E^* + 2\,C^*_{x,e} \quad \text{(A18)}$$

so the covariance of breeding values and environmental effects after selection is

$$C^*_{x,e} = \frac{P^* - G^* - E^*}{2}. \quad \text{(A19)}$$

Using $P = G + E$ this yields

$$C^*_{x,e} = \frac{(1-SP)P - (1-SG)G - (1-SE)E}{2} = -S\frac{P^2 - G^2 - E^2}{2} = -SGE, \quad \text{(A20)}$$

which satisfies $C^*_{x,e} = C^*_{x,e} - C_{x,e} = h^2(1 - h^2)(P^* - P)$.

These formulas apply to the first selection episode, but in further episodes the change in the mean breeding value and environmental effects also depend on their covariance. More general formulas for change at any selection episode can be derived through a multivariate approach. Let us denote the vector of breeding values and environmental effects as $\mathbf{y} = (x, e)^{\mathrm{T}}$ (where superscript T denotes a transposition), with mean $\bar{\mathbf{y}}$ and covariance matrix $\mathbf{Y}$. Substituting $x + e$ for z in the fitness function $W(z)$, and integrating over the joint distribution of x and e, the mean fitness is

$$\bar{W} \propto \exp\left(-\frac{S_C\,(\bar{x} + \bar{e} - \theta)^2}{2}\right) \quad \text{(A21)}$$

$$S_C = \frac{1}{\omega^2 + G + E + 2C_{x,e}} \quad \text{(A22)}$$

927 Then using standard results from evolutionary quantitative genetics (Lande 1979, 1980, Lande and
928 Arnold 1983) and properties of the Gaussian function, the change in the vector of mean breeding
929 values and environmental effects under selection within a generation is

930 $$\Delta^*\bar{\mathbf{y}} = \mathbf{Y}\partial_{\bar{\mathbf{y}}}\ln\bar{W} \quad \text{(A23)}$$

931 where $\partial_{\bar{\mathbf{y}}}$ is the vector of partial derivatives with respect to each element of $\bar{\mathbf{y}}$, yielding

932 $$\Delta^*\bar{x} = -S_C(\bar{x} + \bar{e} - \theta)(G + C_{x,e}) \quad \text{(A24a)}$$

933 $$\Delta^*\bar{e} = -S_C(\bar{x} + \bar{e} - \theta)(E + C_{x,e}) \quad \text{(A24b)}$$

934 $$\Delta^*\bar{z} = \Delta^*\bar{x} + \Delta^*\bar{e} = -S_C(\bar{x} + \bar{e} - \theta)(G + E + 2C_{x,e}) \quad \text{(A24c)}$$

935 When $\bar{e} = C_{x,e} = 0$, these simplify to $\Delta^*\bar{x} = -SG(\bar{x} - \theta)$ and $\Delta^*\bar{e} = -SE(\bar{x} - \theta)$, consistent
936 with the formulas above for the first episode of selection. The recursion for the covariance matrix
937 is

938 $$\Delta^*\mathbf{Y} = \mathbf{Y}(\partial^2_{\bar{\mathbf{y}}}\ln\bar{W})\mathbf{Y} \quad \text{(A25)}$$

939 where $\partial^2_{\bar{\mathbf{y}}}$ is the Hessian matrix of second partial derivatives with respect to each element of $\bar{\mathbf{y}}$,
940 yielding

941 $$\Delta^*G = -S_C(G + C_{x,e})^2 \quad \text{(A26)}$$

942 $$\Delta^*E = -S_C(E + C_{x,e})^2 \quad \text{(A27)}$$

943 $$\Delta^*C_{x,e} = -S_C(G + C_{x,e})(E + C_{x,e}) \quad \text{(A28)}$$

944 When $\bar{e} = C_{x,e} = 0$, these simplify to $\Delta^*G = -SG^2$, $\Delta^*E = -SE^2$, and $C_{x,e} = -SGE$, again
945 consistent with the formulas above for the first episode of selection.

In the general case where $\bar{e}$ and $C_{x,e}$ are non-zero, as is expected if the population has previously undergone any episode of selection in the ongoing generation, then the relationship between phenotypic change and change in breeding values cannot be inferred directly from current components of phenotypic variance. Indeed we have

$$\frac{\Delta^*\bar{x}}{\Delta^*\bar{z}} = \frac{G+C_{x,e}}{G+E+2C_{x,e}} = \frac{\mathrm{Cov}(x,x+e)}{\mathrm{Var}(x+e)} \qquad \text{(A29)}$$

which is still a regression slope of breeding values on phenotypes, but this no longer translates into a ratio of additive genetic variance on phenotypic variance, because of the covariance between breeding values and environmental values. However, it can be shown that this ratio always equals the ratio $\frac{G}{G+E} = h^2$ at the beginning of a generation.

References

Charlesworth, B. 1993. Directional selection and the evolution of sex and recombination. Genetics Research **61**:205-224.

Chevin, L. M., and B. C. Haller. 2014. The temporal distribution of directional gradients under selection for an optimum. Evolution **68**:3381-3394.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. Evolution **33**:402-416.

Lande, R. 1980. Sexual dimorphism, sexual selection, and adaptation in polygenic characters. Evolution **34**:292-305.

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

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

969

970