INVESTIGATION

Response of Polygenic Traits Under Stabilizing Selection and Mutation When Loci Have Unequal Effects

Kavita Jain*,[1] and Wolfgang Stephan[†]

*Theoretical Sciences Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur P.O., Bangalore, India, and
[†]Section of Evolutionary Biology, Department of Biology, Ludwig-Maximilians University of Munich, Planegg-Martinsried, Germany

ABSTRACT We consider an infinitely large population under stabilizing selection and mutation in which the allelic effects determining a polygenic trait vary between loci. We obtain analytical expressions for the stationary genetic variance as a function of the distribution of effects, mutation rate, and selection coefficient. We also study the dynamics of the allele frequencies, focusing on short-term evolution of the phenotypic mean as it approaches the optimum after an environmental change. We find that when most effects are small, the genetic variance does not change appreciably during adaptation, and the time until the phenotypic mean reaches the optimum is short if the number of loci is large. However, when most effects are large, the change of the variance during the adaptive process cannot be neglected. In this case, the short-term dynamics may be described by those of a few loci of large effect. Our results may be used to understand polygenic selection driving rapid adaptation.

In the study of fast adaptation, polygenic selection may be more important than selection on single genes. At single genes, strong selection driving fast adaptation generally leads to the rapid fixation of beneficial alleles or at least large allele frequency shifts between populations. Classical examples of this type of adaptation are the case of industrial melanism in moths (van't Hof *et al.* 2011) and insecticide resistance in *Drosophila* (Daborn *et al.* 2002). However, in many (if not most) cases of fast adaptation, such as in island lizards that are able to adapt very quickly to a changing vegetation [*e.g.*, revealed in an experimental evolution study by Kolbe *et al.* (2012)], small shifts of allele frequencies at many loci may be sufficient to move a phenotype toward a new optimum under changed environmental conditions.

There is a large and growing body of literature on the detection of adaptive signatures in molecular population genetics. Following pioneering work of Maynard Smith and Haigh (1974), the impact of positive selection on neutral DNA variability (selective sweeps) has attracted much interest. This theory has been applied to huge datasets that emerge from modern high-throughput sequencing. A large number of statistical tests have been developed to detect sweep signals and estimate the frequency and strength of selection (Kim and Stephan 2002; Nielsen *et al.* 2005; Pavlidis *et al.* 2010). However, most theory so far excludes the phenotypic side of the adaptive process (except for fitness). Usually, selection is simply modeled as a constant force that acts on a new allele at a single locus. This is in striking contrast to the classical phenotype-based models of adaptation that are successfully used in quantitative genetics (Barton and Keightley 2002). These models typically assume that adaptations are based on allele frequency shifts of small or moderate size at a large number of loci. Also, adaptation does not require new mutations, at least in the short term. Instead, selection uses alleles that are found in the standing genetic variation. Genome-wide data of the past few years show that this quantitative genetic view is relevant. In particular, association studies confirm that quantitative traits are typically highly polygenic. High heritabilities most probably result from standing genetic variation at a large number of loci with small individual effect. Also, local adaptation to environmental clines involves moderate frequency shifts at multiple loci (Hancock *et al.* 2010). As a consequence, there is

growing evidence that the molecular scenario of sweeps only covers part of the adaptive process and needs to be revised to include polygenic selection.

Because genome-wide association studies (GWAS) yield information about the distribution of single-nucleotide polymorphisms relevant to quantitative traits (Visscher *et al.* 2012), it is important to understand the models of polygenic selection in terms of the frequency changes of molecular variants, *i.e.*, in terms of population genetics. So far, however, the dynamics of only very simple polygenic models have been studied and applied to data [*e.g.*, Turchin *et al.* (2012)]. In this article, we analyze the dynamics of a quantitative trait given by a much more general model that was originally proposed by Wright (1935) and recently re-visited by de Vladar and Barton (2014). These authors consider an infinitely large population evolving under stabilizing selection and mutation. Following the observations from many empirical studies [particularly from biomedicine, see Visscher *et al.* (2012)], they assume that the effects are locus-dependent. Furthermore, they consider additivity of the effects and linkage equilibrium between loci.

Here we study this model and obtain some analytical results on the stationary genetic variance and also the dynamics of the phenotypic mean. We show that the stationary genetic variance may exhibit nonmonotonic dependence on the shape of the distribution of effects. We also study how the trajectories of the allele frequencies and the mean trait respond to a sudden environmental shift. When most effects are small, as is the case in experiments on *Drosophila* (Mackay 2004), in livestock (Hayes and Goddard 2001; Goddard and Hayes 2009), and for human height (Visscher 2008), a simple analysis shows that the magnitude of the deviation of the phenotypic mean from the optimum decays roughly exponentially with time and approaches zero over a time scale that is inversely proportional to the initial genetic variance. When most effects are large, the short-term dynamics of the mean and variance can be understood by considering a few loci with large effects.

MODEL WITH LOCUS-DEPENDENT EFFECTS

We consider the ℓ-locus model recently analyzed by de Vladar and Barton (2014) where each locus is biallelic. The $+$ allele at site i has frequency p_i whereas the $-$ allele occurs with frequency $q_i = 1 - p_i$. The effects are assumed to be additive so that the trait value is $z = \sum_{i=1}^{\ell} \text{sgn}(i) \, \gamma_i$, where $\text{sgn}(i) = \pm 1$ denotes the sign of the genotypic value of locus i and $\gamma_i > 0$ is the effect of the allele at the ith locus. The loci are assumed to be in linkage equilibrium so that the joint distribution of effects at the loci factorises. As a result, the nth cumulant c_n of the phenotypic effect, obtained on averaging over the population distribution, can be written as the sum over the corresponding quantities at individual loci. The first three cumulants *viz.* mean c_1, variance c_2, and skewness c_3, are given by the following (Bürger 1991):

$$c_1 = \sum_{i=1}^{\ell} \gamma_i (p_i - q_i) \tag{1a}$$

$$c_2 = 2 \sum_{i=1}^{\ell} \gamma_i^2 \, p_i q_i \tag{1b}$$

$$c_3 = 2 \sum_{i=1}^{\ell} \gamma_i^3 \, (q_i - p_i) p_i q_i. \tag{1c}$$

The allele frequency evolves in time under selective pressure and is given by the following (Barton 1986):

$$\frac{\partial p_i}{\partial t} \approx p_i(t+1) - p_i(t) = \frac{p_i q_i}{2\bar{w}} \frac{\partial \bar{w}}{\partial p_i}, \tag{2}$$

where $\bar{w}$ is the average fitness of the population. For large ℓ, as the trait value z of an individual can be treated as a continuous variable, from (1a) and (1b), we obtain

$$\bar{w} = \int_{-\infty}^{\infty} dz \, p(z) w(z) = 1 - \frac{s}{2} \left(c_2 + (\Delta c_1)^2 \right) \approx e^{-\frac{s}{2} \left(c_2 + (\Delta c_1)^2 \right)}, \tag{3}$$

where the approximate equality sign holds because s is assumed to be small. In the above expression, $w(z) = 1 - (s/2)(z - z_o)^2$ is the fitness distribution of the phenotypic trait under stabilizing selection, z_0 the phenotypic optimum and $\Delta c_1 = c_1 - z_o$ the mean deviation from z_0. Thus the maximum fitness (namely, one) is obtained when the population is at the phenotypic optimum and has no genetic variance. Inserting equations (1) and (3) in (2) and accounting for mutations, we obtain the following basic equation for the evolution of allele frequencies,

$$\frac{\partial p_i}{\partial t} = -\frac{s\gamma_i^2}{2} p_i q_i \left(2 \frac{\Delta c_1}{\gamma_i} + q_i - p_i \right) + \mu(q_i - p_i), i = 1, \ldots, \ell, \tag{4}$$

where μ is the probability of (symmetric) mutation between the $+$ and $-$ allele at locus i. Note that the equation (1) of de Vladar and Barton (2014) is obtained by replacing s by $2s$ in the above equation.

On the right-hand side (RHS) of (4), the first term (in the first parenthesis) expressing the mean deviation from the optimum corresponds to directional selection toward the phenotypic optimum: if the mean is above (below) the optimum, the allele frequencies decrease (increase). However, once the phenotypic mean is sufficiently close to the optimum, stabilizing selection (described by the second term) takes over.

One of the difficulties in solving (4) is that it involves the mean c_1, which depends on all the allele frequencies. Moreover, it has been shown that the differential equations for the cumulants do not close: each one not only involves two higher cumulants but also contains terms that cannot be written in terms of other cumulants (Barton and Turelli 1987; Bürger 1991).

GENETIC VARIANCE IN THE STATIONARY STATE

In the stationary state in which the left-hand side of (4) vanishes, if the mean $c_1^* = z_o$, the allele frequency p_i^* has three solutions, namely $1/2$ and $\left(1 \pm \sqrt{1 - (\hat{\gamma}/\gamma_i)^2} \right) / 2$, where $\hat{\gamma} = 2\sqrt{2\mu/s}$. The latter two solutions are stable for $\gamma_i > \hat{\gamma}$, and therefore the allele frequency is close to fixation when the effects are large. For $\gamma_i < \hat{\gamma}$, the effects are small and the stationary state solution $p_i^* = 1/2$ is the only stable solution for the allele frequency (de Vladar and Barton 2014). From these results, the stationary genetic variance (1b) is easily seen to be (de Vladar and Barton 2014)

$$c_2^* = \frac{4\mu}{s} n_l + \frac{1}{2} \sum_{\gamma_i < \hat{\gamma}} \gamma_i^2, \tag{5}$$

where, for large ℓ, the number of effects larger than $\hat{\gamma}$ is given by $n_l = \ell \int_{\hat{\gamma}}^{\infty} d\gamma \, p(\gamma)$ with $p(\gamma)$ being the distribution of effects. Thus the genetic variance in the stationary state can be neatly written as

$$c_2^* = \frac{\ell}{2}\left[\hat{\gamma}^2 \int_{\hat{\gamma}}^{\infty} d\gamma\, p(\gamma) + \int_0^{\hat{\gamma}} d\gamma\, \gamma^2\, p(\gamma)\right], \qquad (6)$$

where the first (second) term is the contribution from loci with large (small) effects.

The gamma distribution $p(\gamma) \sim \gamma^{k-1} e^{-k\gamma/\bar{\gamma}}$ with shape parameter $k > 0$ and mean $\bar{\gamma}$ has been used to fit the distribution of quantitative trait loci effects (Hayes and Goddard 2001). This distribution is L-shaped for $k < 1$ and bell-shaped for $k > 1$, whereas for $k = 1$, it is an exponential function. For the gamma distribution, the stationary genetic variance (6) is given by

$$c_2^* = \frac{\ell\hat{\gamma}^2}{2}\left[\frac{\Gamma(k,k\gamma_r)}{\Gamma(k)} + \frac{\Gamma(2+k) - \Gamma(2+k,k\gamma_r)}{\gamma_r^2 k^2 \Gamma(k)}\right], \qquad (7)$$

where $\gamma_r = \hat{\gamma}/\bar{\gamma}$ and $\Gamma(a,b) = \int_b^{\infty} dt\, t^{a-1} e^{-t}$ is the incomplete gamma function. For the special case of exponentially distributed effects ($k = 1$), (7) simplifies to give $c_2^* = \ell\bar{\gamma}^2(1 - e^{-\gamma_r}(1 + \gamma_r))$.

For $\bar{\gamma} \gg \hat{\gamma}$, the House of Cards (HoC) variance, namely $c_2^* = \ell\hat{\gamma}^2/2 = 4\mu\ell/s$ (Turelli 1984), is obtained when $\hat{\gamma}$ is finite but $\bar{\gamma} \to \infty$ for all k. In the opposite case ($\bar{\gamma} \ll \hat{\gamma}$), we have $c_2^* = \ell\bar{\gamma}^2(k+1)/(2k)$, which depends on the shape of the distribution of effects. For fixed mean $\bar{\gamma}$, the genetic variance increases monotonically with the scale $\hat{\gamma}$ because a larger mutation probability increases the variance. If instead the distribution mean $\bar{\gamma}$ is increased keeping $\hat{\gamma}$ fixed, the variance increases with $\bar{\gamma}$ toward the HoC value. This is because for fixed k, the width of the distribution increases with $\bar{\gamma}$ and therefore larger effects can be accessed.

The stationary genetic variance has been computed numerically for various shape parameters when $\hat{\gamma} = 0.063$, $\bar{\gamma} = 0.1$ and $\ell = 1000$ in de Vladar and Barton (2014). From (7), the variance for $k = 1, 2, 10$, and 100 is found to be 1.32, 1.57, 1.94, and 2, respectively, which agrees well with the numerical data in their Figure 5. To understand how the variance depends on the shape parameter, we first note that with increasing k (and fixed $\bar{\gamma}$), the width of the gamma distribution decreases. For large k, if $\bar{\gamma} > \hat{\gamma}$, the variance saturates to the HoC variance since almost all loci have large effects with narrow distributions whereas in the opposite case, most effects are small and the variance tends to $\ell\bar{\gamma}^2/2$ (see Figure 1). For small k, irrespective of whether $\bar{\gamma}$ is above or below $\hat{\gamma}$, we find that most effects are small. To see this, consider the fraction $f_s = 1 - (n_l/\ell)$ of loci with small effects which is given by

$$f_s = \frac{(k\gamma_r)^k}{(k-1)!}\int_0^1 dx\, x^{k-1} e^{-k\gamma_r x} = 1 - \frac{\Gamma(k, k\gamma_r)}{\Gamma(k)}. \qquad (8)$$

If $k < (\gamma_r)^{-1}$, the aforementioned equation yields $f_s \approx (k\gamma_r)^k/k!$. Then for finite γ_r, when $k \to 0$, we find that $n_s \to \ell$ for any γ_r, as claimed previously. To summarize, as shown in Figure 1, for $\hat{\gamma} < \bar{\gamma}$, the variance increases with k toward the HoC variance, whereas for $\hat{\gamma} > \bar{\gamma}$, both c_2^* and n_s are nonmonotonic functions of k.

When the effects are chosen from an exponential distribution, the fraction $f_s = 1 - e^{-\gamma_r}$. On eliminating γ_r in favor of f_s in (7) for $k = 1$, we find the relative contribution of loci with small effects to the total variance to be

$$\frac{c_{2,small}^*}{c_2^*} = \frac{2f_s + (1-f_s)\ln(1-f_s)(2 - \ln(1-f_s))}{2f_s + 2(1-f_s)\ln(1-f_s)}, \qquad (9)$$

which increases as $f_s/3$ for small f_s and approaches unity as f_s increases toward one. The aforementioned expression shows that if 10% of the

effects are small, their contribution to the variance is merely 3%, which increases to 21% when f_s is one half. To obtain an equal contribution from small and large effects, a disproportionately large fraction (~83%) of small effects is required. We are unable to obtain an analytical expression analogous to (9) for arbitrary k since f_s is not a simple function of γ_r (see (8), shown previously). However, a numerical analysis using (7) and (8) shows that for the same value of f_s, small effects contribute more to the total genetic variance as the distribution of effects gets narrower. For $f_s = 0.1$, the relative contribution is found to be 2%, 3%, and 5% for $k = 1/2, 1$, and 2, respectively. To obtain an equal contribution from loci with small and large effects, $f_s = 0.89$, 0.83, and 0.77 is needed for the shape parameters $k = 1/2, 1$, and 2, respectively.

DYNAMICS OF THE ALLELE FREQUENCY

We now turn to a description of the allele frequency dynamics and will consider the situation when the phenotypic optimum is suddenly shifted. As mentioned previously, due to the term Δc_1 on the RHS of (4), all the frequencies are coupled, which makes it hard to obtain an exact analytical solution of the allele frequency dynamics. However, under certain conditions, it is a good approximation to consider only the c_1 term in (4) for the initial dynamics and the rest of the terms for long-term evolution.

To see this, we first note that because $0 < p_i < 1$, the mean $|c_1(t)| < \sum_i \gamma_i \approx \bar{\gamma}\ell$. For independent and uniformly distributed initial frequencies, as the average initial frequency is one half, the leading order contribution (in ℓ) to the initial mean is zero. The initial variance is, however, nonzero which gives the typical initial mean $|c_1(0)| \sim \bar{\gamma}\sqrt{\ell}$. When the phenotypic optimum $z_o \lesssim \bar{\gamma}\sqrt{\ell}$ and the number of loci is large, the initial value $|\Delta c_1(0)|/\gamma_i \sim \sqrt{\ell} \gg 1$. Thus at short times, we can neglect $|2p_i - 1|$ (which is bounded above by one) and the mutation term in comparison to the term $2\Delta c_1/\gamma_i$ in (4). At large enough crossover time $t_\times$, as explained in the following section, the mean deviation is close to zero and the reverse condition holds, i.e., $2|\Delta c_1(t)|/\gamma_i \ll |2p_i - 1|$ in (4), and we may set $\Delta c_1 \approx 0$ for later evolution. Biologically, these considerations mean that initially the effects are weaker than the mean trait deviation, but as the population adapts due to directional selection, the deviation of the mean from the phenotypic optimum becomes smaller than the effects.

The aforementioned argument applies not only to uniformly distributed initial frequencies but in more general settings as well where $|c_1(0)| \sim \bar{\gamma}\ell$ by replacing $\sqrt{\ell}$ by ℓ. Here we will focus on the dynamics of the allele frequency when the optimum is suddenly shifted to a new value $z_f (< \ell\bar{\gamma})$, starting from the population which is equilibrated to a phenotypic optimum value z_o. In this situation, as the initial frequency is close to one half when $\gamma_i < \hat{\gamma}$, the frequency $|2p_i(0) - 1|$ is obviously negligible compared with $\Delta c_1(0)/\gamma_i$, whereas for $\gamma_i > \hat{\gamma}$, $|2p_i(0) - 1|$ is close to one because the initial frequency is close to either zero or one (de Vladar and Barton 2014).

When most effects are small

The effects at most of the loci can be smaller than the scale $\hat{\gamma}$ either if k is large and $\bar{\gamma} < \hat{\gamma}$, or if k is small. Then for most loci, at short times, the full model defined by (4) can be approximated by

$$\frac{\partial p_i}{\partial t} = -s\gamma_i p_i q_i \Delta c_1 \qquad (10a)$$

$$\frac{\partial c_n}{\partial t} = -s\Delta c_1 c_{n+1}, \quad n \geq 1, \qquad (10b)$$

where the last equation for cumulants is obtained from the results of Bürger (1991). Equation (10b) for $n = 1$ shows that the magnitude of
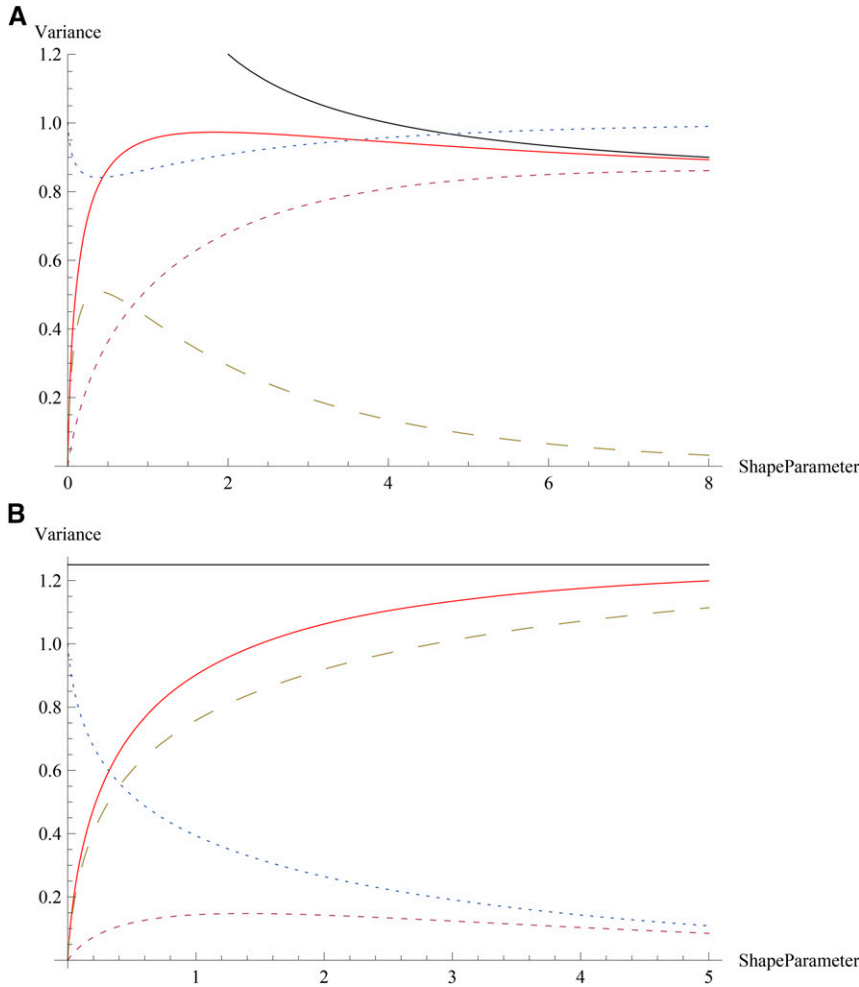
Figure 1 Genetic variance in the stationary state as a function of the shape parameter k when the effects are distributed according to the gamma function. The plot shows the total genetic variance (solid), variance caused by small effects (small dashes), large effects (large dashes), and the fraction of small effects (dotted) for (a) $\bar{\gamma} = 0.04, \hat{\gamma} = 0.08$ and (b) $\bar{\gamma} = 0.1, \hat{\gamma} = 0.05$ for $\ell = 1000$. The asymptotic values $\ell \bar{\gamma}^2 (k+1)/(2k)$ when $\hat{\gamma} > \bar{\gamma}$ and $\ell \hat{\gamma}^2/2$ when $\hat{\gamma} < \bar{\gamma}$ are also shown (top solid curves).

the mean deviation decreases with time, and for the phenotypic optimum smaller than the maximum attainable value of the mean ($z_o \ll \bar{\gamma}\ell$), the trait mean becomes close to the phenotypic optimum at large times (see Figure 2A). We now assume that the variance c_2 is independent of time and stays at its initial value $c_2(0)$ (Chevin and Hospital 2008). As explained in Appendix A, this approximation is good when a combination of the initial cumulants is small (see also Figure 2B). This allows us to solve (10a) and (10b), and we immediately find that

$$\Delta c_1(t) = \Delta c_1(0) e^{-c_2(0)st} \tag{11}$$

$$p_i(t) = \frac{p_i(0)}{p_i(0) + q_i(0)e^{\frac{\gamma_i \Delta c_1(0)}{c_2(0)}(1 - e^{-c_2(0)st})}}. \tag{12}$$

Equation (11) shows that the mean deviation approaches zero over a time scale $t_\times \sim (sc_2(0))^{-1}$.

Next we analyze the long-term evolution of the allele frequencies. As Figure 2A shows, there is a small but nonzero mean deviation Δc_1^* in the stationary state. Taking this into consideration and accounting for the other terms in (4), for $t > t_\times$, we can write

$$\frac{\partial p_i}{\partial t} = -\frac{s}{2}\gamma_i^2 p_i q_i \left(1 - 2p_i + \frac{2\Delta c_1^*}{\gamma_i}\right) + \mu(1 - 2p_i). \tag{13}$$

For $\Delta c_1^* = 0$, the aforementioned equation can be easily solved to give

$$p_i^{(\pm)}(t) = \frac{1}{2}\left(1 \pm \sqrt{\frac{1 - m_i}{1 - M_i(t)}}\right), t > t_\times, \tag{14}$$

where

$$M_i(t) = \frac{4p_i^2(t_\times) - 4p_i(t_\times) + m_i}{(2p_i(t_\times) - 1)^2} e^{-\frac{s\gamma_i^2(1 - m_i)(t - t_\times)}{2}}, t > t_\times, \tag{15}$$

$m_i = (\hat{\gamma}/\gamma_i)^2$ and $p_i(t_\times)$ is obtained from (12). We check that the stationary state solutions $(1 \pm \sqrt{1 - m_i})/2$ and $1/2$ are obtained from the above result for $m_i < 1$ and > 1, respectively. Furthermore, the solution $p_i^{(+)}(t)$ is obtained for $p_i(t_\times) > 1/2$ and $p_i^{(-)}(t)$ for $p_i(t_\times) < 1/2$.

Figure 2 and Figure 3 show a comparison between the numerical solution of (4) and the approximation described previously, when the initial condition is the stationary state of the population equilibrated to a phenotypic optimum z_0. The initial mean deviation $\Delta c_1(0)$ is seen to be close to $-z_f$, and the initial variance $c_2(0) = c_2^*$ for the zero mean deviation is 0.0967, which is close to the value 0.131 obtained from the set of effects used in Figure 2 and Figure 3. As Figure 2 shows, the dynamics of the mean deviation are captured well by (11) and approach a stationary value close to zero ($\Delta c_1^* \approx -0.016$) in about 1500 generations. The variance also evolves with time, but the change is not substantial and the approximation $c_2(t) \approx c_2(0)$ is good over the time scale directional selection toward the phenotypic

A

MeanDeviation

B
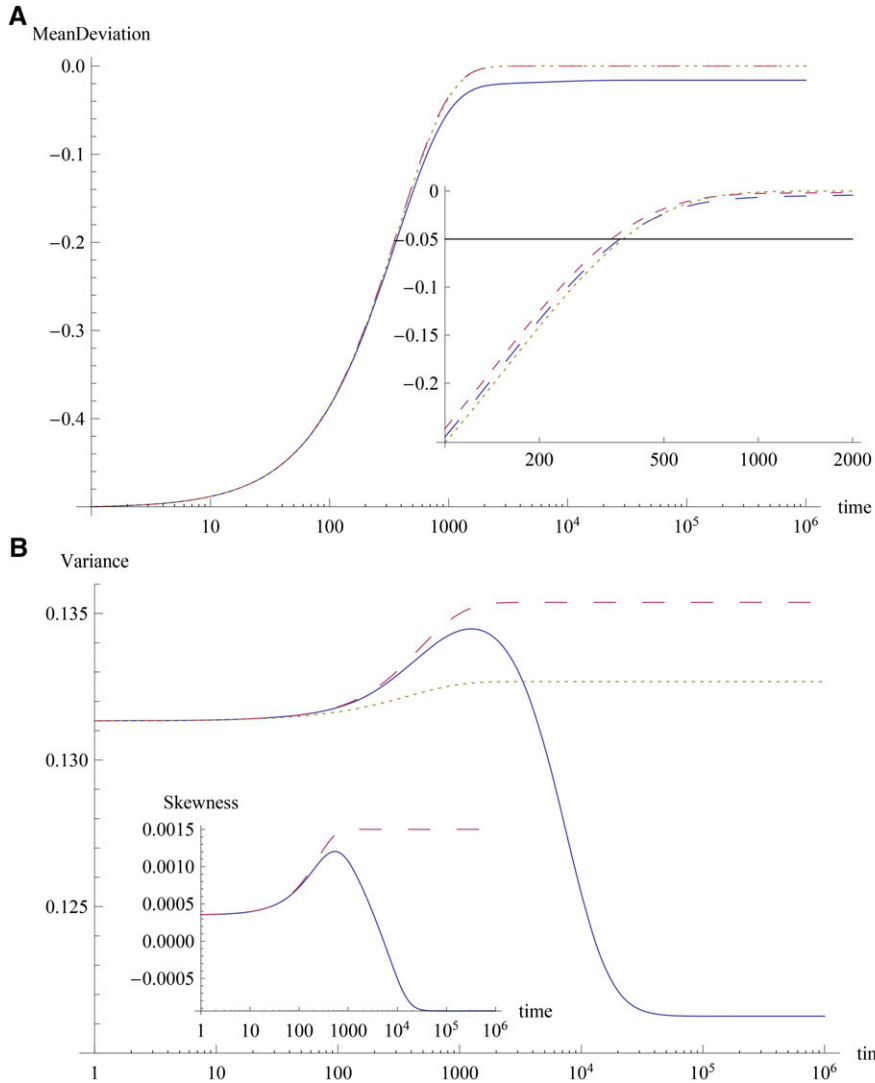
Variance

Skewness

Figure 2 Response to change in optimum when most effects are small. The plot shows the results for (A) mean deviation $\Delta c_1(t)$ and (B) variance $c_2(t)$ and skewness $c_3(t)$ obtained using the exact numerical solution of the full model (solid) and the short-term dynamics model (large dashes). The dotted curves show the time-dependent solution (11) for mean and (A.1) for variance. The parameters are $\ell = 50$, $s = 0.02$, $\mu = 5 \times 10^{-5}$, $\hat{\gamma} = 0.14 > \bar{\gamma} = 0.05$, $z_o = -0.0012$, $z_f = 0.5$, $n_l = 5$. The effects are chosen from an exponential distribution, and the parameter $\mathcal{C} = -0.01$ (see Appendix A). The inset in the top figure shows the difference $\Delta c_1(t) - \Delta c_1^*$ as a function of time for the full model when the effects are gamma-distributed with shape parameter $k = 1$ (large dashes), 5 (small dashes), and 20 (dotted).

optimum operates. Equation (11) also indicates that directional selection toward the optimum will occur faster when the initial variance is large since $t_\times \sim 1/c_2^*$.

Because the stationary genetic variance displays a nonmonotonic dependence on the shape parameter k of the gamma distribution (see Figure 1), the relaxation time for the mean deviation is expected to decrease and then increase with increasing k. Indeed, as the inset of Figure 2A shows, the difference $c_1(t) - c_1^*$ (which, by definition, is zero in the equilibrium state for all k) equals a reference value -0.05 at time 360, 340, and 370 for $k = 1$, 5, and 20, respectively.

The allele frequency dynamics are shown in Figure 3. We see that although the short-term dynamics can be accurately described by (12) for loci with effects smaller than or close to the distribution mean, there is a substantial difference when the effects are larger than the mean. This is because for such loci, the initial frequency is not close to half and the term involving $q_i - p_i$ on the RHS of (4) cannot be neglected. For $t > t_\times$, the long-term behavior described by (13) is shown with $\Delta c_1^* = 0$ and the actual mean deviation.

When most effects are large

When $\bar{\gamma} > \hat{\gamma}$ and k is large, the number of loci with large effects is also large, and the initial allele frequencies are close to either zero or one.

In this parameter regime, both the variance and the skewness may change appreciably during directional selection toward the optimum, and the constant-variance approximation discussed above is not suitable. However, at very short times when $\Delta c_1(t)$ is close to its initial value, the solution (12) for the allele frequency gives

$$p_i(t) \approx \frac{1}{1 + \frac{q_i(0)}{p_i(0)} e^{\gamma_i \Delta c_1(0) st}} \ . \tag{16}$$

From the aforementioned equation, we first note that the allele frequency at large-effect loci changes fast as expected intuitively. Equation (16) also shows that for $\Delta c_1(0) < 0$, the allele frequency quickly increases toward unity, if the initial frequency is close to unity and therefore does not contribute to the dynamics of the variance or skewness. Thus, to understand the short-term dynamics, we need to focus our attention on large-effect loci with low initial allele frequency for negative initial mean deviation. Similar remarks apply to the situation when $\Delta c_1(0)$ is positive where the large-effect loci with high initial frequency determine the dynamics.

In the following, we assume that $\Delta c_1(0) < 0$ and consider the time evolution of the allele frequency P of the largest effect locus with lowest initial frequency. Figure 4 shows that the allele frequency
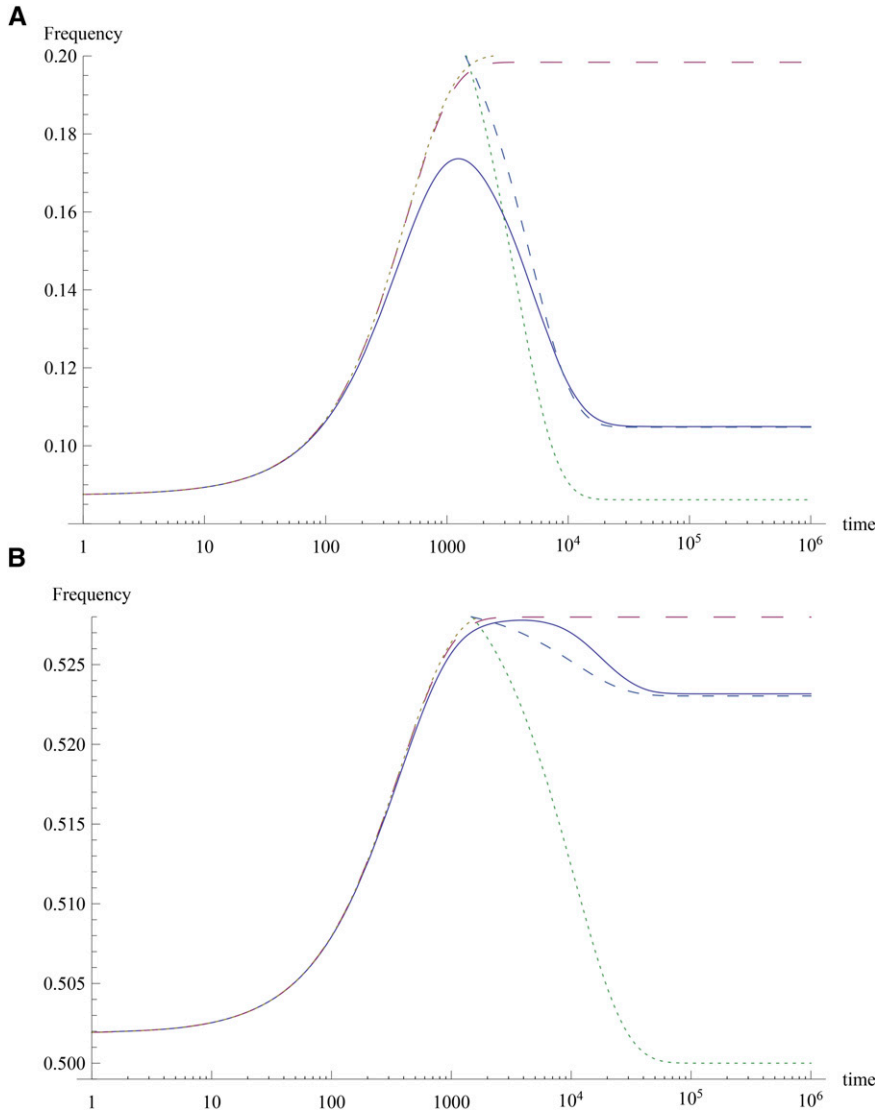
Figure 3 Response to change in optimum when most effects are small. The plot shows the allele frequencies for two representative loci with (A) $\gamma_i = 0.252$ and (B) $\gamma_i = 0.028$ for the full model (solid) and short-term dynamics model (large dashes). The dotted curves show the time-dependent solution (12) for $t < t_\times$ and (14) for $t > t_\times$ where $t_\times = 1500$. The dashed curve for $t > t_\times$ is the solution of (13) with $\Delta c_1^* = -0.016$. The other parameter values are the same as in Figure 2.

P sweeps to fixation, but the frequency of the next relevant locus (*i.e.*, the next largest effect locus with low initial frequency) does not. In such a case, we can approximate the mean c_1 and variance c_2 by the contribution from the frequency P with effect Γ, and obtain

$$c_1(t) \approx 2\Gamma(P - P_0) + c_1(0) \tag{17a}$$

$$c_2(t) \approx 2\Gamma^2(PQ - P_0Q_0) + c_2(0), \tag{17b}$$

where $P_0 \equiv P(0)$. Then using the aforementioned expression for the phenotypic mean in (4) and neglecting mutations (since most effects are large), we get

$$\frac{\partial P}{\partial t} = -s\Gamma^2 P(1 - P)(P + \alpha), \tag{18}$$

where

$$\alpha = \frac{\Gamma + 2\Delta c_1(0)}{2\Gamma} - 2P_0. \tag{19}$$

We thus find that the allele frequency P is a solution of the following equation:

$$\frac{(P/P_0)^{1+\alpha}}{(Q/Q_0)^\alpha} = e^{-s\Gamma^2\alpha(1+\alpha)t}\frac{P+\alpha}{P_0+\alpha}. \tag{20}$$

An explicit solution of (20) seems hard to obtain since α is in general not an integer. However, for large and negative α, the aforementioned equation yields

$$P = \frac{1}{1 + \frac{Q_0}{P_0}e^{-s\Gamma^2|\alpha|t}}. \tag{21}$$

Thus for $z_f \gg \Gamma$, the frequency P sweeps to fixation in a time of order $(s\Gamma z_f)^{-1}$.

Figure 4 shows the allele frequency of the largest effect locus with lowest initial frequency obtained using (4). It agrees reasonably well with the solution of (20) and the expression (21) where $\alpha = -1.43$. In Figure 5, the dynamics of the first two cumulants given by (1a) and (1b) are compared with the approximate expressions (17a) and (17b), respectively, and we see a good agreement.

A detailed numerical analysis of the set of parameter values of Figure 4 suggests that the dynamics of this example can be understood by considering one, two, or three of the largest-effect loci. When only the largest-effect locus was required, the second largest effect was much lower than the largest one.
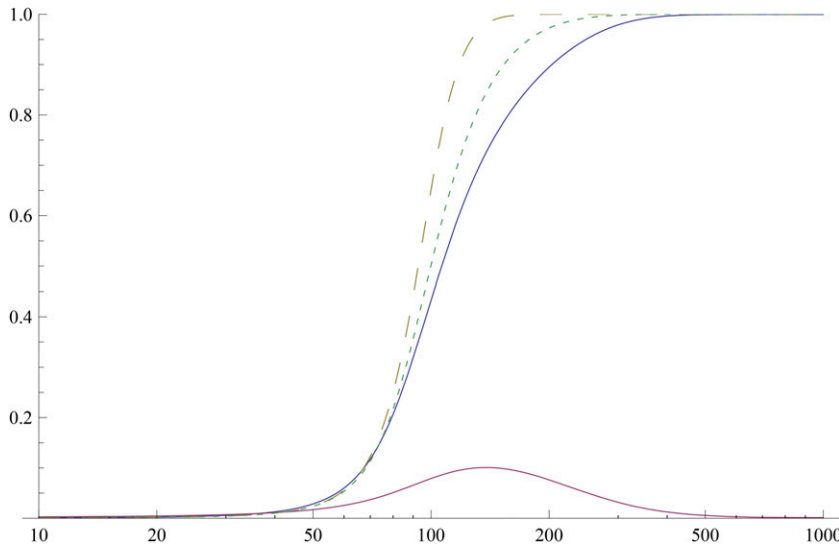
Figure 4 Response to change in optimum when most effects are large. The plot shows the exact numerical solution of the full model (solid) and the equations (20) (large dashes) and (21) (small dashes) for the dynamics of the allele frequency P with the largest effect and lowest initial frequency ($\Gamma = 0.776, P_0 = 3.3 \times 10^{-4}$). The solid curve at the bottom shows the numerical solution of the full model for the frequency of the next relevant locus with effect size 0.319 and initial frequency 1.9×10^{-3}. The parameters are $\ell = 20, s = 0.1, \mu = 10^{-5}, \hat{\gamma} \approx 0.028 \ll \bar{\gamma} = 0.2, z_o = 7.8 \times 10^{-5}, z_f = 1.5, n_l = 19$.

DISCUSSION

One of the fundamental questions in adaptation is whether the adaptive process is governed by many loci of small effect or few loci of large effect (Orr, 2005). However, which effects are small, and which large? de Vladar and Barton (2014) have provided a scale $\hat{\gamma} \sim \sqrt{\mu/s}$ for the size of effects, which is a function of basic population genetic parameters, namely mutation probability μ and selection coefficient s, relative to which an effect is defined as large or small. For a given distribution of effects, assumed here to be a gamma distribution with mean $\bar{\gamma}$ and shape parameter k, an effect is small (large) if it is below (above) $\hat{\gamma}$. But for fixed $\bar{\gamma}$ and $\hat{\gamma}$, whether most or a few effects are small depends on the shape parameter k: for small k, most effects are small, but for large k, the number of small effects depends on the ratio $\gamma_r = \hat{\gamma}/\bar{\gamma}$.

Genetic variance in stationary state

Here we have provided analytical expressions for the stationary genetic variance c_2^* when the effects are locus-dependent. We find that when most effects are small, c_2^* is a nonmonotonic function of the shape parameter of the gamma distribution (going through a maximum for intermediate values of k; see Figure 1). In contrast, it increases monotonically when most effects are large. As Figure 1 shows, when the shape distribution is narrow, large (small) effects contribute most to the variance when $\gamma_r > 1 (<1)$. However, for broad distributions, although the number of small effects is large, small effects do not contribute much when $\gamma_r < 1$. The HoC variance is obtained irrespective of k when $\gamma_r \rightarrow 0$ because all loci have large effect in this limit. As noted previously (de Vladar and Barton, 2014), HoC provides an upper bound on the genetic variance.

Dynamics when most effects are small

As the distribution of quantitative trait loci measured in experimental and natural populations (Hayes and Goddard, 2001; Mackay, 2004; Visscher, 2008; Goddard and Hayes, 2009) find most effects to be small, it is important to study this situation in detail. Here we have obtained analytical expressions for the dynamics by assuming the genetic variance to be constant. Although the fact that the variance does not change much in time when most effects are small was observed numerically in de Vladar and Barton (2014), an explanation of this behavior was not provided. Here, as explained in Appendix A, it is

a good approximation to assume the variance to be time-independent provided the product of the initial values of the mean deviation and skewness is small.

In the absence of mutations, Chevin and Hospital (2008) have considered the effect of background with a time-independent genetic variance on the frequency at a single focal locus. Their results match the ones obtained here using the short-term dynamics model with directional selection only; in particular, (11) and (12) match the results (21) and (25) of Chevin and Hospital (2008), respectively, on identifying their parameters ω^2 and a with $1/s$ and γ from this study.

Our basic result concerning the dynamics of the phenotypic mean is that it relaxes over a time scale that is inversely proportional to the initial variance. Because the variance is of order ℓ, we thus have the important result that the mean approaches the optimum faster if a larger number of loci is involved. Moreover, this time depends nonmonotonically on the shape parameter of the gamma distribution. Note that the phenotypic mean deviation relaxes to zero when the phenotypic optimum is far below the upper bound $\bar{\gamma}\ell$ on the phenotypic mean. However, when the phenotypic optimum exceeds the maximum typical value of the mean, such that the mean deviation remains substantially different from zero at late times, (10b) shows that all higher cumulants vanish at the end of the phase of directional selection.

Dynamics when most effects are large

When the initial mean deviation is moderately large (and negative), the genetic variance changes by a large amount over the time scale directional selection occurs and the dynamics can be understood by considering a few loci whose effect is large but initial frequency is low.

However, for larger mean deviations (but smaller than $\ell\bar{\gamma}$), a few large-effect loci do not completely capture the dynamics of the mean and the variance. As Supporting Information, Figure S1 shows, the initial increase of the absolute mean deviation and the transient rise of the variance can be explained by considering the large-effect locus. At later times, however, as the change in variance is small, we can use the constant-variance approximation to understand the dynamics of the phenotypic mean deviation until it nearly vanishes. The constant-variance approximation also can be used when the initial mean deviation is sufficiently small (see Figure S2).

A

MeanDeviation

B
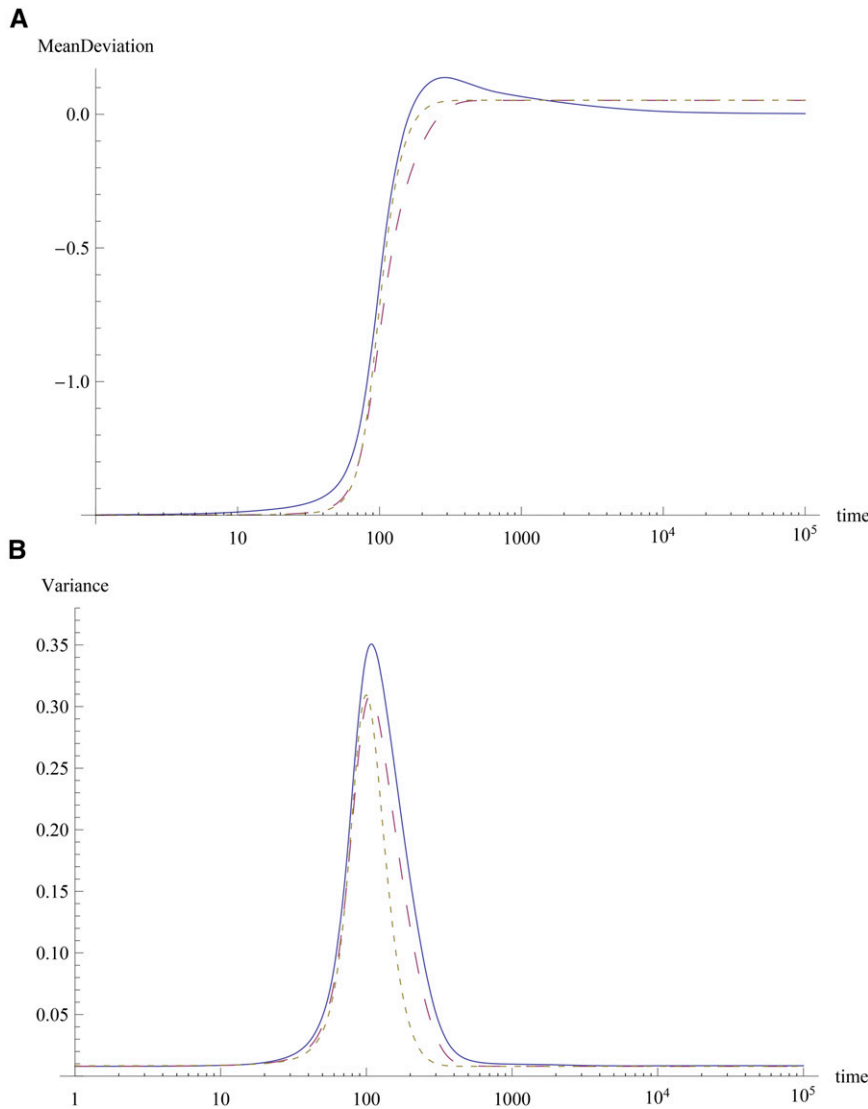
Variance

Figure 5 Response to change in optimum when most effects are large. Solid lines show the mean deviation (A) and variance (B), whereas the large dashed curves show the contribution to these cumulants from the locus with the largest effect and lowest initial frequency $(\Gamma = 0.776, P_0 = 3.3 \times 10^{-4})$. In both cases, the exact numerical solution of the full model is used. The numerical solution of (18) (small dashes) is also shown. The other parameter values are the same as in Figure 4.

Applications

The approximations presented here hold for the short-term evolution of phenotypic traits and allele frequencies. This means that our results may be used to understand polygenic selection driving rapid adaptation. In this respect, our most important result is that the mean of a phenotypic trait may respond faster to a sudden environmental change when the number of loci is large and most effects are small.

Evidence for rapid phenotypic evolution has been reported in recent years from several groups of organisms. For instance in *Drosophila subobscura*, latitudinal clines of wing size have been formed within 20 years since this species colonized America (Huey *et al.*, 2000). Similarly, in field experiments in which lizard populations were newly established on small islands in the Bahamas, the hindlimbs adapted very quickly to the different vegetations on the islands (Kolbe *et al.*, 2012). To our knowledge, however, data from GWAS are not yet available in these cases.

The theory presented here can also be applied to the large amounts of GWAS data that have been gathered in model species such as humans and *Drosophila*. To analyze the observed allele frequency shifts in single-nucleotide polymorphisms associated with quantitative traits, such as human height (Turchin *et al.* 2012) and cold tolerance

in *Drosophila* (Huang *et al.* 2012), the results derived in this study provide a more general theoretical basis than the dynamical equations used in previous analyses (*e.g.*, Turchin *et al.* 2012).

Open questions

The analytical calculations in this article work when the phenotypic mean at the equilibrium coincides exactly with the optimum. However, in the stationary state, there is a small but nonzero mean deviation due to which the long-term dynamics are not accurately captured, especially when effects are small, as shown in Figure 3B. An improved calculation of the dynamics that takes a nonzero mean deviation into account is certainly of interest, for instance to estimate the frequency of selective fixations (leading to selective sweeps) in this model (Chevin and Hospital 2008; Pavlidis *et al.* 2012; Wollstein and Stephan 2014).

Another open question concerns the generality of our results presented here. The current model perhaps oversimplifies biological reality in that it neglects genetic drift and assumes additive effects, symmetric mutations and free recombination between loci. It can be shown that the current model (neglecting mutation) can be derived from the classical symmetric viability model with arbitrary position of

the optimum under the quasi-linkage equilibrium assumption. In the latter model, the probability of selective fixation has been studied numerically for up to eight loci (Pavlidis *et al.* 2012; Wollstein and Stephan 2014). However, at present we are lacking an analytical understanding of the role of recombination in this model and how it relates to the high-recombination limit represented by our current model.

LITERATURE CITED

Barton, N., 1986   The maintenance of polygenic variation through a balance between mutation and stabilizing selection. Genet. Res. 47: 209–216.

Barton, N., and M. Turelli, 1987   Adaptive landscapes, genetic distance, and the evolution of quantitative characters. Genet. Res. 49: 157–174.

Barton, N. H., and P. D. Keightley, 2002   Understanding quantitative genetic variation. Nat. Rev. Genet. 3: 11–21.

Bürger, R., 1991   Moments, cumulants, and polygenic dynamics. J. Math. Biol. 30: 199–213.

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

Daborn, P. J., J. L. Yen, M. R. Bogwitz, G. Le Goff, E. Feil *et al.*, 2002   A single p450 allele associated with insecticide resistance in *Drosophila*. Science 27: 2253–2256.

de Vladar, H. P., and N. H. Barton, 2014   Stability and response of polygenic traits to stabilizing selection and mutation. Genetics 197: 749–767.

Goddard, M., and B. Hayes, 2009   Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat. Rev. Genet. 10: 381–391.

Hancock, A. M., D. B. Witonsky, E. Ehler, G. Alkorta-Aranburu, C. Beall *et al.*, 2010   Human adaptations to diet, subsistence, and ecoregion are due to subtle shifts in allele frequency. Proc. Natl. Acad. Sci. USA 107: 8924–8930.

Hayes, B., and M. Goddard, 2001   The distribution of the effects of genes affecting quantitative traits in livestock. Genet. Sel. Evol. 33: 209–229.

Huang, W., S. Richards, M. Carbone, D. Zhu, R. Anholt *et al.*, 2012   Epistasis dominates the genetic architecture of *Drosophila* quantitative traits. Proc. Natl. Acad. Sci. USA 109: 15553–15559.

Huey, R., G. Gilchrist, M. Carlson, D. Berrigan, and L. Serra, 2000   Rapid evolution of a geographic cline in size in an introduced fly. Science 287: 308–309.

Kim, Y., and W. Stephan, 2002   Detecting a local signature of genetic hitchhiking along a recombining chromosome. Genetics 160: 765–777.

Kolbe, J., M. Leal, T. Schoener, D. Spiller, and J. Losos, 2012   Founder effects persist despite adaptive differentiation: a field experiment with lizards. Science 335: 1086–1089.

Mackay, T., 2004   The genetic architecture of quantitative traits: lessons from *Drosophila*. Curr. Opin. Genet. Dev. 14: 253–257.

Maynard Smith, J., and J. Haigh, 1974   Hitchhiking effect of a favourable gene. Genet. Res. 23: 23–35.

Nielsen, R., S. Williamson, Y. Kim, M. J. Hubisz, A. G. Clark *et al.*, 2005   Genomic scans for selective sweeps using SNP data. Genome Res. 15: 1566–1575.

Orr, H. A., 2005   The genetic theory of adaptation: a brief history. Nat. Rev. Genet. 6: 119–127.

Pavlidis, P., J. D. Jensen, and W. Stephan, 2010   Searching for footprints of positive selection in whole-genome SNP data from nonequilibrium populations. Genetics 185: 907–922.

Pavlidis, P., D. Metzler, and W. Stephan, 2012   Selective sweeps in multilocus models of quantitative traits. Genetics 192: 225–239.

Turchin, M., C. Chiang, C. Palmer, S. Sankararaman, D. Reich *et al.*, 2012   Evidence of widespread selection on standing variation in Europe at height-associated SNPs. Nat. Genet. 44: 1015–1019.

Turelli, M., 1984   Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle. Theor. Popul. Biol. 25: 138–193.

van't Hof, A. E., N. Edmonds, M Daliková, F Marec, and I. J Saccheri, 2011   Industrial melanism in British peppered moths has a singular and recent mutational origin. Science 332: 958–960.

Visscher, P., 2008   Sizing up human height variation. Nat. Genet. 40: 489–490.

Visscher, P., M. Brown, M. McCarthy, and J. Yang, 2012   Five years of GWAS discovery. Am. J. Hum. Genet. 90: 7–24.

Wollstein, A., and W. Stephan, 2014   Adaptive fixation in two-locus models of stabilizing selection and genetic drift. Genetics 198: 685–697.

Wright, S., 1935   Evolution in populations in approximate equilibrium. J. Genet. 30: 257–266.

Communicating editor: S. I. Wright

A Validity of Constant-Variance Approximation

As explained in the main text, we obtain (11) when the variance is assumed to be constant in time. To see under what conditions this approximation holds, we find a correction to the variance by assuming that the skewness c_3 is nonzero and time-independent (see the inset of Figure 2B). Using (11) on the RHS of (10b) for $n = 2$, we immediately get

$$c_2(t) = c_2(0) \left[1 - \mathcal{C}\left(1 - e^{-c_2(0)st}\right)\right],$$
(A.1)

where $\mathcal{C} = \Delta c_1(0)c_3(0)/c_2^2(0)$. Plugging the above solution into (10b) for $n = 1$ gives the mean deviation as

$$\Delta c_1(t) = \Delta c_1(0) \, \exp\left[-stc_2(0)(1 - \mathcal{C}) - \mathcal{C}\left(1 - e^{-stc_2(0)}\right)\right].$$
(A.2)

The solution (11) is recovered if the constant $\mathcal{C}$, which depends on the initial value of the first three cumulants, is negligible.

If we start with the initial condition in which the population is equilibrated to an optimum and most effects are small, since most initial allele frequencies are close to one half, the initial variance is substantial and the skewness is close to zero. In this case, the constant-variance approximation is expected to work well. However, if most effects are large, since most initial allele frequencies are close to fixation, although the skewness remains small, the variance also becomes small, thus leading to an increase in $\mathcal{C}$. Then for sufficiently small mean deviations, we also may employ the constant-variance approximation when most effects are large.