The Genetic Basis of Phenotypic Adaptation I: Fixation of Beneficial Mutations in the Moving Optimum Model

Michael Kopp¹ and Joachim Hermisson

Mathematics and Biosciences Group, Max F. Perutz Laboratories and Faculty of Mathematics, University of Vienna, A-1030 Vienna, Austria

> Manuscript received December 16, 2008 Accepted for publication February 24, 2009

ABSTRACT

We study the genetic basis of adaptation in a moving optimum model, in which the optimal value for a quantitative trait increases over time at a constant rate. We first analyze a one-locus two-allele model with recurrent mutation, for which we derive accurate analytical approximations for (i) the time at which a previously deleterious allele becomes beneficial, (ii) the waiting time for a successful new mutation, and (iii) the time the mutant allele needs to reach fixation. On the basis of these results, we show that the shortest total time to fixation is for alleles with intermediate phenotypic effect. We derive an approximation for this "optimal" effect, and we show that it depends in a simple way on a composite parameter, which integrates the ecological parameters and the genetic architecture of the trait. In a second step, we use stochastic computer simulations of a multilocus model to study the order in which mutant alleles with different effects go to fixation. In agreement with the one-locus results, alleles with intermediate effect tend to become fixed earlier than those with either small or large effects. However, the effect size of the fastest mutations differs from the one predicted in the one-locus model. We show how these differences can be explained by two specific effects of multilocus genetics. Finally, we discuss our results in the light of three relevant timescales acting in the system—the environmental, mutation, and fixation timescales—which define three parameter regimes leading to qualitative differences in the adaptive substitution pattern.

WHEN a population adapts to a changing environment, what is the genetic basis of this process? For example, does adaptation occur in small or large steps? These questions have been asked ever since the debate on micro- vs. macromutationalism in the early days of evolutionary theory (PROVINE 2001). Subsequently, the genetics of adaptation have become the subject of different modeling approaches (see ORR 2005a,b for review). Considerable work exists, in particular, in the context of Fisher's geometric model (e.g., FISHER 1930; KIMURA 1983; Orr 1998; Welch and Waxman 2005; Martin and LENORMAND 2006), Gillespie's mutational landscape model (*e.g.*, GILLESPIE 1983, 1984; ORR 2002), and various models of so-called "adaptive walks" on rugged fitness landscapes (e.g., KAUFFMAN and LEVIN 1987; KAUFFMAN 1993). These approaches treat organismal adaptation as a process in a high-dimensional phenotype or genotype space characterized by pleiotropy and epistasis. Another common feature is that all these models assume a constant selection pressure. Yet, many organisms live in environments that are continually changing (e.g., HAIRSTON et al. 2005; THOMPSON 2005; PARMESAN 2006; PERRON et al. 2008). This fact is acknowledged in the so-called moving optimum model, which assumes that the selectively favored phenotype

changes over time and, hence, experiences a mixture of stabilizing and directional selection.

The moving optimum model was originally devised in the field of quantitative genetics, where its primary aim was to study adaptation at the phenotypic level, without an explicit focus on the underlying genetics. In this context, the model has been used, for example, to study the ability of populations to avoid extinction in the face of global change (e.g., LYNCH et al. 1991; LYNCH and LANDE 1993; BÜRGER and LYNCH 1995; NUNNEY 2003), to study the role of temporal variation for the maintenance of genetic variation (BÜRGER 2000; BÜRGER and GIMELFARB 2002), and to understand the evolution of sex and recombination (BÜRGER 1999; WAXMAN and PECK 1999). Only recently, several studies have started to investigate how a moving optimum affects the dynamics of adaptive substitutions (BELLO and WAXMAN 2006; COLLINS et al. 2007; KOPP and HERMISSON 2007; SATO and WAXMAN 2008). A key result is that, under conditions of slow environmental change, mutations with small effect tend to become fixed earlier than those with large effect (BELLO and WAXMAN 2006; COLLINS et al. 2007; KOPP and HERMISSON 2007). This is in direct contrast to what is found under constant selection after a sudden change in the environment (e.g., ORR 1998, 2002, 2005a; KIM and ORR 2005).

In a previous note (KOPP and HERMISSON 2007), we identified three parameter regimes in the moving opti-

¹Corresponding author: University of Vienna, Dr. Bohr-Gasse 9, A-1030 Vienna, Austria. E-mail: michael.kopp@univie.ac.at

234

mum model that lead to qualitative differences in the dynamics of adaptive substitutions. These regimes are defined by three timescales that may dominate the adaptive process: the ecological timescale, which is given by the (inverse) speed of the optimum; the mutation timescale, which determines the waiting time for a new beneficial allele; and the time that it takes for this allele to rise to fixation. In particular, we found that small-effect mutations are fixed before large ones if the ecological timescale is dominating, that is, if the speed of adaptation is limited primarily by the speed of environmental change (see also COLLINS et al. 2007 for similar findings). In KOPP and HERMISSON (2007), this result was obtained from an approximate calculation in a minimal model with two haploid biallelic loci. Here, we extend the previous analysis in two main ways. First, we significantly improve our approximation for the total fixation time of a given allele under moving-optimum selection. Second, we analyze the order of adaptive substitutions in a full haploid or diploid multilocus model. We confirm the preliminary results from the two-locus study, but show that regime boundaries depend on both the ecological parameters and the genetic architecture of the trait. The current article focuses on the order of substitutions over relatively short timescales. The complementary question about which mutations dominate long-term evolution will be treated in a separate study.

THE MODEL

Assumptions on fitness: Consider the evolution of a quantitative trait *z* that is under stabilizing selection with respect to a moving optimum $z_{opt} = z_{opt}(t)$. We assume that the optimum increases over time at a constant rate *v*; that is,

$$z_{\rm opt}(t) = vt. \tag{1}$$

Note that, unlike in KOPP and HERMISSON (2007), the optimum increases indefinitely. Assuming Gaussian stabilizing selection, the fitness of an individual with phenotype z at time t is given by

$$w(z, t) = \exp(-\sigma(z - z_{\text{opt}}(t))^2), \qquad (2)$$

where $\sigma > 0$ determines the strength of selection. The selection coefficient of a mutant with phenotype α in a wild-type population with phenotype 0 at time *t* is

$$s(t) = \frac{w(\alpha, t)}{w(0, t)} - 1 = \exp(-\sigma\alpha(\alpha - 2\nu t)) - 1.$$
(3)

As long as selection is weak (σ or $|z - z_{opt}|$ small), Equation 3 can be approximated by

$$s(t) \approx \lambda(t - \frac{\alpha}{2v}),$$
 (4a)

where

$$\lambda = 2\sigma v\alpha. \tag{4b}$$

Thus, selection for α increases approximately linearly over time.

Assumptions on genetics: The aim of this article is to model the first steps of the adaptive process in the moving optimum model. We assume that the trait z is influenced by L additive, unlinked, haploid or diploid loci. Each locus has two alleles, which we refer to as the wild-type and the mutant allele, respectively. The contribution of the wild-type allele to the phenotype z is 0, and the contribution of the mutant allele is α_i (*i* = 1, ..., L). The α_i are referred to as (locus) mutational effects. In the one-locus case, we suppress the index i and simply write $\alpha_1 = \alpha$. Mutations from the wild-type to the mutant allele (and vice versa) occur recurrently at rate μ per locus. In our calculations, we usually use the population mutation parameter $\theta = 2N\mu$, where N is the population size. Environmental variation is not modeled explicitly, but is subsumed in the selection parameter σ (e.g., BÜRGER 2000). Our analytical results are compared to and extended by stochastic computer simulations, which are described in APPENDIX A.

RESULTS

In the following, we first derive an approximation for the expected time to fixation of a mutant allele at a single haploid locus. Consequences of diploidy are considered next. We then build on these results and complement them with computer simulations to determine the order of fixed substitutions in a model with multiple loci of unequal effect.

Expected time to fixation of a single mutant allele: Consider a monomorphic population with the wild-type phenotype z = 0 and a mutant allele (arising by recurrent mutation at rate μ) that alters the phenotype to $z = \alpha$. At time t = 0, the frequency of the mutant allele is p = 0. Our basic question is: How long does it take, on average, for the mutant allele to go to fixation?

Some care is needed to formulate a meaningful notion of "fixation." Problems may arise, in particular, from the final phase of the fixation process where mutant frequencies are close to p = 1. In fact, a mutant may never reach p = 1 in models with back mutation (and, of course, in natural populations). More generally, in the context of adaptation, the importance of the accidental time point when a mutation finally reaches full fixation at p = 1 is questionable. To avoid these problems, and following previous work (BELLO and WAXMAN 2006; KOPP and HERMISSON 2007), we instead calculate the time until the mutant allele reaches majority status $(p \ge \frac{1}{2})$ and, thus, dominates the locus genetics. For ease of terminology, we nevertheless refer to this state as fixation. (Additional simulations with a fixation criterion of $p \ge 0.9$ yielded results that are qualitatively and quantitatively similar to those reported below.)

Denote the total time to fixation (*i.e.*, to $p = \frac{1}{2}$) by T. Following KOPP and HERMISSON (2007), this time can be subdivided into three periods,

$$\mathbb{T} = T_\ell + T_\mathrm{w} + T_\mathrm{f},\tag{5}$$

where

$$T_{\ell} = \frac{\alpha}{2\nu} \tag{6}$$

is the *lag time* until the mutant allele becomes beneficial (*i.e.*, has a positive selection coefficient, see Equation 4a). $T_{\rm w}$ is the *waiting time* for a successful mutation (*i.e.*, for the appearance of a mutant allele that is not subsequently lost due to drift). Finally, $T_{\rm f}$ is the (narrow-sense) *fixation time* of this successful mutation (*i.e.*, the time needed for the frequency to increase from p = 1/(2N) to $p = \frac{1}{2}$). To calculate the waiting and fixation times, we need to distinguish two different cases. In the first case, fixation occurs from a mutant allele that appears after the end of the lag time, so that the waiting time is positive. We refer to this case as fixation occurs from a mutant allele that allele that allele that is already segregating in the population at the end of the lag time, and the waiting time is zero.

Case 1—fixation from a new mutation: In contrast to the lag time (Equation 6), the waiting time T_w is a random variable. It can be approximated as resulting from an inhomogeneous Poisson process with time-dependent rate $\theta_s(t)$. The probability density function of T_w is

$$f(T_{\rm w}) = \theta s(T_{\rm w} + T_{\ell}) \exp\left[-\frac{\theta}{2}s(T_{\rm w} + T_{\ell})^2\right]$$
(7a)

(KOPP and HERMISSON 2007), which has mean

$$\bar{T}_{\rm w} = \int_0^\infty T_{\rm w} f(T_{\rm w}) dT_{\rm w} = \sqrt{\frac{\pi}{2\theta\lambda}}$$
(7b)

and variance

$$\operatorname{Var}(T_{\mathrm{w}}) = \frac{4 - \pi}{2\theta\lambda}.$$
 (7c)

(Note that, here, we use the symbol T_w for both the random variable and its realization.) This approximation assumes that, during the short time period in which the fate of a new mutation is determined, the selection coefficient does not change significantly. This is the case if $s \gg \sqrt{\lambda}$ or, with $s = \lambda T_w$ (see below) and Equation 7b, if $\sqrt{\pi/(2\theta)} \gg 1$ (from diffusion theory; J. HERMISSON, unpublished results). The approximation is best if $\theta \ll 1$, such that mutations appear so late that selection is already strong. For $\theta \gg 1$, adaptation occurs from an already segregating allele, and $T_w = 0$ anyway. The approximation therefore performs reasonably well in the whole parameter space.

The fixation time T_f also is a random variable, but its variance is small relative to that of the waiting time (as is well known for constant selection, *cf.* ETHERIDGE *et al.* 2006). Therefore, we treat it as deterministic. However, for fixation from a new mutation, $T_f =: T_f^n$ still depends on the stochastic waiting time, because T_w determines

the selection pressure during the fixation phase. In APPENDIX B, we derive

$$T_{\rm f}^{\rm n}(T_{\rm w}) = \frac{\sqrt{(s^*)^2 + 2\lambda \ln(2Ns^*/(\theta+1) - 1)}}{\lambda} - T_{\rm w},$$
(8a)

where

$$s^* = \frac{\lambda}{2} \left(T_{\rm w} + \sqrt{T_{\rm w}^2 + \frac{4}{\lambda}} \right). \tag{8b}$$

The expected total time to fixation from a new mutation, \mathbb{T}^n , is given by

$$E(\mathbb{T}^{n}) = T_{\ell} + \int_{T_{w}=0}^{\infty} [T_{w} + T_{f}^{n}(T_{w})]f(T_{w})dT_{w}.$$
 (9)

Since $T_{\rm f}^{\rm n}$ (as a function of $T_{\rm w}$) is approximately linear around $T_{\rm w} = \overline{T}_{\rm w}$ (not shown), Equation 9 can be approximated by

$$E(\mathbb{T}^{n}) \approx T_{\ell} + \bar{T}_{w} + T_{f}^{n}(\bar{T}_{w}).$$
(10)

Inserting \overline{T}_{w} from (7b) into (8b), the value of *s** becomes

$$s^*(T_{\rm w} = \bar{T}_{\rm w}) = \sqrt{\lambda} \left(\sqrt{\frac{\pi}{8\theta} + 1} + \sqrt{\frac{\pi}{8\theta}} \right), \qquad (11)$$

which, for small θ is approximately equal to $\sqrt{\lambda\pi/2\theta} = \lambda \bar{T}_w$.

Case 2—fixation from an already segregating allele: If θ is large, there is a high probability that fixation occurs from a mutant allele that has appeared already during the lag time. In this case, $T_{\rm w} = 0$ and the fixation time $T_{\rm f} =: T_{\rm f}^{\rm s}$ is given by the same expression as $T_{\rm f}^{\rm n}$ (8a), but with

$$s^* = \sqrt{\lambda} \left(\sqrt{\frac{\pi}{16} + 1} - \sqrt{\frac{\pi}{16}} \right) \tag{12}$$

(see APPENDIX B). The expected total time to fixation from an already segregating allele, \mathbb{T}^s , is simply

$$E(\mathbb{T}^{\mathrm{s}}) = T_{\ell} + T_{\mathrm{f}}^{\mathrm{s}}.$$
(13)

In summary, combining Equations 8a, 10, 13, A7b, and A9, the total time to fixation can be estimated as

$$E(\mathbb{T}^{i}) \approx \frac{\alpha}{2v} + \frac{1}{\lambda} \sqrt{(s^{*})^{2} + 2\lambda \ln\left(\frac{2Ns^{*}}{\theta + 1} - 1\right)}, \quad (14a)$$

where

$$s^{*} = \begin{cases} \sqrt{\lambda} \left(\sqrt{\frac{\pi}{8\theta} + 1} + \sqrt{\frac{\pi}{8\theta}} \right) \\ \text{for fixation from a new mutation } (i = n), \\ \sqrt{\lambda} \left(\sqrt{\frac{\pi}{16} + 1} - \sqrt{\frac{\pi}{16}} \right) \\ \text{for fixation from a segregating allele } (i = s). \end{cases}$$
(14b)

A necessary condition for the validity of this approximation is $s^* > 2\mu + 1/N$. In APPENDIX B, we show that the probability of fixation from an already segregating allele is



FIGURE 1.—The total time to fixation (*i.e.*, time until frequency $p > \frac{1}{2}$) of a mutant allele in the biallelic one-locus model as a function of the mutational effect α . Each plot shows means and standard deviations from 1000 simulation runs, together with the prediction from the analytical approximation in Equation 16 (solid line). The dotted line indicates α^* , the value of α with the shortest time to fixation (Equation C2). The area with dark shading indicates the lag time T_{ℓ} (Equation 6) and the area with light shading the mean waiting time (*i.e.*, \bar{T}_w from Equation 7b multiplied by the probability for fixation from a new mutation from Equation 15). $\theta = 2N\mu$ with $\mu = 10^{-5}$.

$$\mathcal{P}^{\rm s} = 1 - \exp\left(-\frac{\theta}{\sqrt{2}}\right). \tag{15}$$

Thus, the expected total time to fixation from either a new or a segregating allele is

$$E(\mathbb{T}) = (1 - \mathcal{P}^{s})E(\mathbb{T}^{n}) + \mathcal{P}^{s}E(\mathbb{T}^{s}).$$
(16)

Numerical analysis of Equation 16 shows that the total time to fixation for a mutant with a given effect decreases with increasing v, σ , and θ . This means that adaptation is fastest if the environment changes quickly, selection is strong, and the populationwide mutation rate is high. The predictions from Equation 16 are in excellent agreement with results from stochastic simulations (Figure 1). Visible deviations occur only for the combination of large v and small θ (v = 0.01, $\theta = 0.02$ in Figure 1),

where the predicted time to fixation is slightly shorter than the observed one. The reason is that, for these parameters, the selection coefficient may become quite strong before a beneficial allele arrives. In this case, assuming a fixation probability of 2s(t) (as is done in the derivation of Equation 7a) is an overestimation. The true fixation probabilities are lower and, hence, the true waiting times are longer. With the more accurate approximation for the fixation probability, $1 - \exp(-2s(T_w + T_\ell))$, the waiting time distribution becomes

$$F(T_{\rm w}) = 1 - \exp\left[-\frac{\theta}{2}\left(T_{\rm w} - \frac{1 - \exp(-2\lambda T_{\rm w})}{2\lambda}\right)\right].$$

This equation provides a very good fit to the simulated waiting time distribution, but cannot easily be used in further analytical derivations.



FIGURE 2.— α^* , the allelic effect of the mutant allele with the shortest time to fixation in the biallelic one-locus model, as a function of the speed of the environmental change v for various values of the selection parameter σ and the mutation parameter $\theta = 2N\mu$. Lines are numerically calculated minima of Equation 16 (see Figure 1). Symbols show the approximation (C2).

For given, v, σ , and θ , the total time to fixation is a Ushaped function of α (Figure 1). That is, mutations with intermediate effect become fixed faster than mutations with either small or large effects. The reason is that mutations with small effects are only weakly selected for and, thus, have long waiting times (due to a low fixation probability) and long fixation times. In contrast, mutations with large effects are deleterious for a long time and, therefore, have long lag times.

The value of α that minimizes the total time to fixation is denoted by α^* . In APPENDIX C, we derive an analytical approximation (Equation C2), which elucidates how α^* depends on the model parameters (see Figure 2). To a good approximation,

$$\alpha^* \propto \sqrt[3]{\frac{v}{\sigma}} \tag{17a}$$

(where \propto denotes proportional) and, if θ is small,

$$\alpha^* \propto \sqrt[3]{\frac{v}{\sigma\theta}}.$$
(17b)

Thus, fast environmental change favors relatively large mutations, whereas a large populationwide mutation rate and strong selection both favor relatively small mutations. The reason is that fast environmental change leads to the buildup of a large lag between the wild-type phenotype and the optimal phenotype, which can then be bridged by a large mutation. In contrast, high mutation rates and strong selection lead to the fixation of small mutations before the lag becomes large. For a mutation of size α^* we show in APPENDIX c that the lag time is approximately one-third of the total time to fixation, suggesting that none of the three timescales is dominating the other two.

The diploid case: So far, we have assumed that the locus under study is haploid. Here, we show how the previous analysis can be extended to the diploid case. We assume that heterozygotes have phenotype $\alpha > 0$ and that mutant homozygotes have phenotype $\beta \ge \alpha$. This allows for various degrees of (partial) dominance, but not for over- or underdominance (nor for complete recessiveness). Furthermore, we define μ as the diploid (twice the haploid) mutation rate (such that the number of new mutations per generation remains $N\mu$). Since, initially, the mutant allele occurs only in heterozygotes, the lag time and the waiting time are identical to those in the haploid case (*i.e.*, Equations 6 and 7 remain valid). The same is not true for the fixation time, however, because the further spread of the mutant allele depends on both the heterozygous and the homozygous fitnesses. Nevertheless, there is a simple argument that allows us to derive a rough approximation based on the haploid result.

The mutant allele can invade the population once the heterozygotes have higher fitness than the wild type, which is the case for $t > t_1 = \alpha/(2v)$. However, the wildtype allele can be replaced only once the mutant homozygotes have higher fitness than the heterozygotes, that is, for $t > t_2 = (\alpha + \beta)/(2\nu)$. For t between t_1 and t_2 , heterozygotes have the highest fitness, and selection favors a stable polymorphism of the two alleles. During this time, the frequency of the mutant allele cannot exceed the equilibrium frequency of the polymorphism, which increases from 0 at t_1 to 1 at t_2 . At time $t_3 = \beta/(2v)$, in particular, the equilibrium frequency reaches $\frac{1}{2}$ (our criterion for fixation). Therefore, the total time to fixation can be estimated as the maximum of $\beta/(2v)$ and of the time predicted for the haploid model (Equation 16). As seen in Figure 3, this approximation works surprisingly well for most parameter values, and also the minimum α^* remains largely unchanged. Visible deviations from simulation results occur only for intermediate values of α that are close to the point where the line $\beta/(2v)$ intersects with the haploid prediction. Note that for $\beta = 0$ (full dominance) the diploid and the haploid result coincide (confirmed by simulations, not shown).

The order of fixations in the multilocus case: In the single-locus model, we have seen that a moving optimum favors fast fixation of mutations with an intermediate phenotypic effect, and that the size of the fastest mutation depends on the mutation rate and on the ratio of the speed of environmental change to the strength of selection. What does this result predict (qualitatively and quantitatively) for the order of fixed mutations in a full multilocus model, in which alleles with different



FIGURE 3.—The total time to fixation of a single mutant allele in the diploid model. The additive case is shown. Heterozygotes have phenotype α , and mutant homozygotes have phenotype $\beta = 2\alpha$. The diagonal dashed line is α/ν , the time at which the mutant homozygote has the same fitness as the wild type. The solid line is the maximum of the predicted time to fixation in the haploid model (Equation 16) and α/ν . The vertical dotted line is α^* , the effect of the fastest mutation in the haploid model (Equation C2). μ is the diploid mutation rate. For further details, see Figure 1.

effects "compete" against each other? Figure 4 shows simulation results for a model with 40 biallelic loci and uniformly distributed mutational effects α_i . Details of two example runs are shown in Figure 5. We see that the characteristic U shape from Figure 1 is indeed preserved (Figure 4, A–C). Mutant alleles with intermediate effects are fixed earlier than those with either small or large effects. Not surprisingly, the full pattern is seen only if the range of mutational effects is sufficiently large (Figure 4, D–F): If only small (large) alleles are available, then only the left (right) branch of the U appears, leading to a monotonic and almost linear decrease (increase) of fixed effects over time.

Despite the qualitative agreement with the singlelocus results (Figure 1), there are two quantitative differences that require explanation. First, the size of the fastest alleles is shifted to the left, that is, to effects less than α^* . And second, the pattern assumes a Y shape, rather than a U shape, if mutation rates are high (Figure 4C). These differences can be explained by two effects of multilocus genetics: a "sampling effect" for small θ and an "interaction effect" for large θ .

Sampling effect: The sampling effect arises from a combination of two factors: For small θ , the waiting time has a large variance (Equation 7c), and small mutations have a shorter lag time than large mutations. As a consequence, even though small mutations might have a high expected total time to fixation, their minimal possible time is quite low. If the number of loci is large, there is a high probability that one of the small mutations has passed. Thus, the sampling effect leads to an "advantage" for small mutations, which is strongest if the waiting time is the dominating timescale.

Since the sampling effect is a purely statistical phenomenon, it can still be estimated within a single-locus



FIGURE 4.—The order of fixations in the multilocus model. Symbols show the time to fixation of mutant alleles as a function of their effect α . Different symbols represent results from five replicated simulations with mutational effects drawn from a uniform distribution with a given range (between 0 and $2\alpha^*$ in A–C, between 0 and $\alpha^*/2$ in D, between 0 and α^* in E, and between $\alpha^*/2$ and $2\alpha^*$ in F). The vertical dashed line marks α^* , the effect of the mutant allele with the smallest expected time to fixation in the one-locus model (Equation C2). In A–C, shaded bars show the distribution of the effect of the first fixation as obtained from 1000 replicated simulations with fixed and evenly spaced α -values. The dashed line shows the predicted distribution of the first fixation according to Equation D5. Parameters: L = 40, $\mu = 10^{-5}$, $\sigma = 0.1$, $\theta = 2N\mu$, v = 0.001.

framework. In APPENDIX D, we use our results from the one-locus model to derive an approximation for the distribution of the first fixation in the multilocus model. This approximation neglects epistatic interactions between mutant alleles at different loci (see below). Nevertheless, it is in good agreement with simulation results, as long as θ is not too large (compare dashed lines and shaded histograms in Figure 4, A–C). The reason is that, for small θ , narrow sense fixation times are small, and different alleles rarely segregate simultaneously (see Figure 5A). Note also that the mode of the distribution of the first fixation coincides well with the base of the U in Figure 4, A and B.

Interaction effect: If θ is large, the probability of mutant alleles segregating simultaneously is high (see Figure 5B). These alleles interact with each other, because stabilizing selection entails epistasis for fitness: Each mutant allele that increases in frequency brings the population mean closer to the optimum and, thereby, decreases the selection coefficient of the other alleles. In particular, segregating small mutations can delay the fixation of large mutations by effectively prolonging their lag time.

The interaction effect has two main consequences: First, the size of the first fixation is smaller than predicted from the one-locus model, even if the sampling effect is negligible. This can be seen from Figure 4C, where the prediction for the first step according to Equation D5 (dashed line) is close to α^* , but the observed distribution (shaded bars) is considerably smaller. Second, the interaction effect causes a delay in the fixation of some intermediate and all large mutations, thus transforming the U-shaped into a Y-shaped pattern (Figure 4, C and E). Figure 5B gives a more detailed view of this process. The initial decrease in effect sizes (the stem of the Y) arises because, at the beginning of the simulation, several alleles with small and intermediate effect start to increase in frequency more or less simultaneously. During their increase, these alleles suppress other mutant alleles, since their combined effect keeps the population mean close to the optimum. At about the "branching point," the bulk of the initial alleles have become fixed, and only two classes of alleles remain: those with very small effect, which have a very long fixation time and slowly increase in frequency "in the background," and those with large effects and a long lag time, whose frequency up to this point was almost zero and who now, one by one, quickly go to fixation.

The interaction effect depends on the cosegregation of beneficial alleles at multiple loci and is strongest if both small and large mutations are competing for fixation. This is demonstrated in Figure 4, C–F. The Yshaped pattern is seen only if the effects of available mutations cover a large range (Figure 4, C and E). In contrast, if all mutations are large, they rarely cosegregate at higher frequencies. No delay in fixation is observed, and the size of fixed mutations increases almost



FIGURE 5.—Adaptation in the multilocus model. A1 is an example of a single simulation run with 20 loci ($\alpha_i = 0.02, 0.04, 0.06, 0.08,$ 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.5, 0.6, 0.7, 0.8, 1, 1.2, 1.4, 1.6, 1.8) in a small population (parameters as in Figure 4A). Each symbol shows the time to fixation of the allele with the respective phenotypic effect. Note that, in contrast to Figure 4, time is on the x-axis and α on the y-axis. (This is to allow for comparison with A2, see below.) The dotted line marks α^* , the effect of the mutant allele with smallest expected time to fixation in the one-locus model (Equation C2). A2 shows the corresponding allele frequency dynamics. Colors indicate the allelic effects size α_i (light red, small effect; dark blue, large effect). Each symbol in A1 corresponds to the line in A2 with the

same color, and the time coordinate at which the symbol is placed in A1 equals the time the corresponding line in A2 crosses frequency 0.5 (dotted line). B1 and B2 show a similar example for a large θ (parameters as in Figure 4C) and α_r values half the size of those in A. Populations were sampled every 100 generations.

linearly over time (Figure 4F). Finally, if all mutations are small, epistatic interactions between them are even smaller. Cosegregating alleles hardly interfere, and the size of fixed mutations decreases linearly over time (Figure 4D).

Robustness of results: To test the robustness of the above results, we conducted a large number of supplementary simulations that consider additional factors (Figure 6).

- i. Distribution of locus effects: Both theoretical and empirical evidence suggests that small beneficial mutations are more common than large ones (*e.g.*, BÜRGER 2000; BARTON and KEIGHTLY 2002; ORR 2005a; EYRE-WALKER and KEIGHTLEY 2007). As seen in Figure 4, a limited range of mutational effects can have a large influence on the order of fixations. However, if the range is sufficiently large, the exact shape of the distribution from which mutational effects are drawn seems to matter little. For example, results qualitatively similar to those in Figure 4 are obtained if the distribution of mutational effects is exponential instead of uniform (Figure 6A).
- ii. Diploidy: As in the single-locus case, haploid and diploid genetics lead to qualitatively identical results (Figure 6B).
- iii. Variable mutation rate: So far, we have assumed that the per-locus mutation rate $\mu = 10^{-5}$, and we have varied the population size *N* to obtain various values of $\theta = 2N\mu$. Choosing a different value for μ has little effect in most cases (Figure 6C), showing that the crucial parameter is θ , not *N* or μ individually. Substantial differences occur only if μ is extremely high ($\geq 10^{-3}$ in Figure 6C). In this case, recurrent

mutation may become more important than selection in determining the dynamics of allele frequencies ($\mu > s^*/2$, see remarks regarding Equations 14 and B5). This leads to a "boost" for small mutations, which are under the weakest selection and have the shortest lag times. As a consequence, the U shape disappears, and the size of fixed mutations increases monotonically over time (Figure 6C3).

- iv. Linkage: The multilocus simulations presented above assume linkage equilibrium (see APPENDIX A). However, additional simulations using individual-based modeling (Figure 6D; M. KOPP and J. HERMISSON, unpublished results) show that the results are robust in the presence of moderate linkage (recombination rates ≥ 0.01 , Figure 6D1). Strong linkage in combination with large θ , in contrast, leads to an increase in stochasticity and an increase in the size of the first fixation (Figure 6, D2 and D3). Such a shift is also known from other models of adaptation and has been described in the context of clonal interference (e.g., GERRISH and LENSKI 1998; PARK and KRUG 2007) and the Hill-Robertson effect (HILL and ROBERTSON 1966; BARTON 1995). Individual-based simulations also show that our results are not biased by the quadratic approximation (A2).
- v. Standing genetic variation: The simulations above were started with frequency p = 0 for all mutant alleles at time t=0, but genetic variation was allowed to accumulate during the lag time (see fixation from "already segregating alleles" above). In another set of simulations, we let the population equilibrate to

240



FIGURE 6.—The order of fixations in variants of the multilocus model. (A) Mutational effects were drawn from an exponential distribution with mean α^* . (B) The diploid case (no dominance). Here, α is the effect of a single mutant allele. (C) Various combinations of *N* and μ leading to the same value of θ . (D) Various degrees of linkage. *r* is the recombination rate between adjacent loci. Symbols show the time of fixation of mutant alleles as a function of their effect α . Different symbols represent results from five replicated simulations with randomly drawn mutational effects. The vertical dashed line marks α^* , the effect of the mutant allele with the smallest expected time to fixation in the one-locus model (Equation C2). In C3, α^* could not be calculated because, for small α , the approximation (14a) failed ($\mu > s^*/2$). In B–D, mutational effects were drawn from a uniform distribution with mean α^* . See Figure 4 for more details. Parameters: L = 40, $\sigma = 0.1$, v = 0.001, and $\theta = 2N\mu = 2$ (B and C), and $\mu = 10^{-5}$ (A, B, and D).

the initial conditions before starting to move the optimum. This allows for the buildup of standing genetic variation (HERMISSON and PENNINGS 2005; BARRETT and SCHLUTER 2008). It turns out that the

order of fixations is only slightly affected, with small loci benefiting more from standing variation than large ones (which are more deleterious initially). The reason for the small effect is that, for large Θ ,

most fixations occur from mutations that arise during the lag time. However, this is already accounted for in our model (Equation 15), and adding genetic variation that is present even before the lag time does not change the qualitative picture.

vi. Speed of environmental change: As in the one-locus case, we performed simulations with v = 0.01 and v = 0.0001 and obtained qualitatively identical results (not shown). For v > 0.01, small populations sometimes are unable to follow the moving optimum and, instead, go to extinction (Bürger and LYNCH 1995). While the persistence or extinction of populations in the face of environmental change is an important topic (LYNCH *et al.* 1991; LYNCH and LANDE 1993; Bürger and LYNCH 1995; NUNNEY 2003) with obvious implications for conservation, it is not the focus of this study.

DISCUSSION

The speed and pattern of phenotypic adaptation depend on a combination of factors that are external or internal to an organism. While new selection pressures on a phenotypic trait result from changes in the external environment, the response to selection depends on the details of the internal genetics, that is, on the rates and effects of new mutations. The moving optimum model with an explicit genetics is a first attempt to combine these various factors in a single modeling framework and to analyze their relative roles in the adaptive process.

In this article, we have focused on a specific issue: Assume that mutations with various effects on the trait compete to be recruited in the adaptive process: What is the expected order of adaptive fixations? Which size of mutations—small, intermediate, or large—is fixed first under the given environmental and genetic conditions? We have taken a two-step approach to address this question. In the first step, we consider a single mutant allele. We have derived a highly accurate approximation for how the expected total time to fixation of this allele depends on the internal (genetic) and external (environmental) factors. In the second step, we examine how these results are altered by multilocus effects in the context of a polygenic trait.

In both cases, we find that the fastest mutations are those with intermediate size. They have the shortest time to fixation in the one-locus model (Figures 1 and 3), and they are the first ones to reach fixation in the multilocus case (Figures 4 and 5). But what exactly does "intermediate" mean in this context? In the following, we give a qualitative and quantitative answer to this question by applying the notion of timescales and evolutionary regimes.

Fixation of a single mutation: Our analysis for the total time to fixation of a single allele is based on the

insight that fixation in the moving optimum model is controlled by three different timescales (KOPP and HERMISSON 2007). The first timescale is ecological. It is given by the (inverse) speed of the trait optimum and governs the strength of selection. The other two timescales are genetic and measure the waiting time for a successful beneficial mutation and the time for the mutant allele to rise to fixation. Accordingly, the total time to fixation can be subdivided into three components that correspond to these timescales: The lag time until the allele becomes beneficial as a result of environmental change, the waiting time for a successful mutation (which may be zero if fixation occurs from an already segregating allele), and the (narrow-sense) fixation time.

The three timescales partition the genetic and environmental parameter space into three evolutionary regimes (KOPP and HERMISSON 2007): In the environmentally limited regime, the total time to fixation of an allele is dominated by the lag time; in the mutation*limited regime*, it is dominated by the waiting time; and in the fixation time-limited regime, it is dominated by the fixation time. The latter two regimes may be summarized as genetically limited. The principal distinction between them is the degree to which adaptation is rather stochastic (mutation-limited regime, populationwide mutation rate $\theta \leq 1$) or rather deterministic (fixation time-limited regime, $\theta \gtrsim 1$). For our present issue, however (that is, the order of adaptive substitutions), what matters most is the distinction between the two genetically limited regimes and the environmentally limited regime.

As seen in Figure 1, the total time to fixation is shortest for mutations with intermediate effect $\alpha = \alpha^*$. It increases with α for $\alpha > \alpha^*$ and decreases with α for $\alpha < \alpha^*$. The limiting factor is different for small and large mutations, respectively. Mutations with small effects are under weak selection and, therefore, have low fixation probabilities and long fixation times. They are thus genetically limited (either by mutation or by fixation time, depending on θ). In contrast, mutations with large effects have increasingly long lag times before they become beneficial and are environmentally limited. For mutations of size α^* , the lag time is approximately onethird of the total time to fixation (*i.e.*, it is equal to the mean of the waiting time and the fixation time), suggesting that fixation is fast because the genetic and environmental timescales are balanced. Because of all this, it makes sense to say that α^* parameterizes the boundary between the two genetically limited regimes and the environmentally limited regime. The fastest mutations are located at this boundary.

In consequence, α^* provides a natural scale for measuring allelic effects (see APPENDIX c). Its numerical value decides which effect sizes are "small" or "large" in the context of the adaptive process. Of course, this classification is a purely relative one. Whether a mutation of a given absolute effect will be counted as small or large depends on the value of α^* and, hence, on the ecological and genetic parameters.

Indeed, the relationship between α^* and the model parameters is very simple (Equation 17). To a good approximation, α^* depends on the speed of environmental change relative to the strength of selection, v/σ , and on the populationwide mutation rate θ . If θ is small, α^* depends only on a single composite parameter

$$\gamma = \frac{v}{\sigma \theta} \tag{18}$$

(see Equation 17b), which summarizes the genetic and ecological factors and should be amenable to empirical study. In nature, γ is expected to vary over several orders of magnitude. The speed of environmental change vexperienced by a population depends on the ecological scenario and on the generation time, and it may vary from virtually imperceptible to extremely rapid. Similarly, the effective population size (and hence θ) differs greatly among species. Thus, for a given ecological change, α^* will usually be larger for microbes (with large populations and short generation times) than, for example, mammals. On the other hand, the dependence of α^* on γ as a third root is rather weak (Equation 17). For example, a 10-fold increase in the speed of environmental change leads to only a 2.14-fold increase in α^* . For the parameter values covered in Figure 2, α^* is on the order of 0.1 or 1. If all phenotype-related variables (α, σ, v) are measured in units of the environmental standard deviation, this is well within the range of mutational effects estimated from empirical data, which vary between 10⁻² and 1 (Bürger 2000, p. 264). In addition, the distribution of mutational effects is thought to be highly leptokurtic; that is, mutations with large effects are overrepresented relative to a Gaussian distribution (BÜRGER 2000, p. 264). Therefore, it is likely that the majority of mutations will be smaller than α^* , but a nonnegligible part is bound to be larger. This means, in particular, that the fastest mutations are likely to be available in the population.

The order of fixations at multiple loci: In the singlelocus model, we have argued that the fastest mutations are those at the boundary of the environmentally limited regime. A similar analysis can also be applied to the order of fixations in the multilocus model. However, to do so, we first need to define the notion of evolutionary regimes in a multilocus context. We will do so on a per-locus basis. Thus, in a multilocus model with unequal locus effects, one part of the loci may belong to one regime (*e.g.*, environmentally limited), whereas others fall into a different regime (*e.g.*, mutation limited). Note that this definition deviates slightly from the one used in KOPP and HERMISSON (2007), where regimes were defined for the whole model, not for single loci. Using only the timescale arguments from the singlelocus case, the results shown in Figure 4 can be interpreted in the following way:

- i. If all loci are in one of the genetically limited regimes, large-effect loci adapt before small-effect loci. This corresponds to the classical pattern known for adaptation under constant selection (*e.g.*, ORR 1998, 2002; KIM and ORR 2005).
- ii. If all loci are in the environmentally limited regime, loci with small effect are favored over ones with larger effect. This confirms the pattern described by BELLO and WAXMAN (2006), KOPP and HERMISSON (2007), and COLLINS *et al.* (2007).
- iii. In a general model with loci in both regimes, we see a branching pattern in the order of fixations, where loci with an intermediate effect fix first and mutations with either small or large effects follow up later. Thus, the linear patterns described under i and ii can be seen as special cases, which only arise if the trait has a genetic architecture with a restricted range.

Despite the qualitative agreement with the single-locus results, however, Figure 4 suggests that the boundary between the genetically limited and the environmentally limited regimes is shifted to the left, that is, to locus effects less than α^* . This can be explained by two genuine multilocus effects that are acting at small or large mutation rates, respectively:

- 1. For small θ , a *sampling effect* causes the average size of the first fixations to be smaller than α^* (the fastest mutation in the one-locus model). The reason is that the waiting time for the first successful mutation in a "race" at multiple loci is shorter than the average waiting time at a single locus. This advantage is larger for loci with small effect, for which the waiting time is the dominating timescale. The sampling effect can still be captured analytically by using results from the one-locus model and ignoring interactions between loci (APPENDIX D).
- 2. For large mutation rates ($\theta \ge 1$), the fixation pattern is affected by interactions between cosegregating alleles. Figure 4, C and E, shows that the split into two branches is again shifted to smaller locus effects and sets in only after an initial phase (Y shape). The reason is epistatic interactions: For large θ , mutations with small effect are dominated by the (narrow sense) fixation time. At any given time, therefore, there will be multiple segregating mutations that have not yet reached fixation. These mutations contribute to a narrowing of the gap between the mean phenotype and the optimum phenotype and reduce the selection pressure on mutations at other loci. In particular, they increase the lag time for mutations with larger effect, which leads to a shift in the pattern of early fixations toward smaller loci.

Discussion of the modeling approach: Previous modelers have used different approaches for studying the genetics of adaptation (ORR 2005a). Given the complex nature of the adaptive process, it is not surprising that different models have focused on different aspects. The adaptive walk and mutational landscape models view evolution as a search process on a high-dimensional fitness surface (GILLESPIE 1983, 1984; KAUFFMAN and LEVIN 1987; KAUFFMAN 1993; ORR 2002). Here, the key factor is epistasis, which makes the landscape rugged. Models of clonal interference focus on the effects of linkage and recombination on the fixation process (BARTON 1995; GERRISH and LENSKI 1998; KIM and ORR 2005; PARK and KRUG 2007). None of the above models contains an explicit phenotype. In contrast, Fisher's geometric model explores the consequences of pleiotropy in a high-dimensional genotype-phenotype map (FISHER 1930; ORR 1998; WELCH and WAXMAN 2005; MARTIN and LENORMAND 2006). In our moving optimum model, finally, the key point is the inclusion of a changing environment, that is, of ecology.

So far, the specifics of the alternative approaches epistasis, pleiotropy, and linkage—are largely ignored in our model. However, all these factors are potentially important for a full understanding of the genetic basis of adaptation. Since the moving optimum model can be readily extended to include all these aspects, it provides a natural basis for an integrative approach toward the genetics of phenotypic adaptation.

Extending our model in this way might well lead to changes in some of our predictions. For example, many of our results depend on the fact that large mutations initially overshoot the optimum and only have a chance at fixation after a certain amount of time. This conclusion is inevitable in our model, where the phenotype space has only a single dimension. However, starting with FISHER (1930), many authors have stressed the necessity to view adaptation as an inherently multidimensional problem. In a high-dimensional phenotype space, it is conceivable that a mutation overshooting the optimum in a particular dimension is nevertheless selected for due to beneficial effects in other dimensions (i.e., due to pleiotropic side effects). Thus, large mutations might become fixed earlier than it is possible in the present model.

Another way to compare the various models is with respect to the included timescales. In the models of Kauffman, Gillespie, and Orr, only the mutation timescale is relevant. KIM and ORR (2005) allow for an increased mutation rate, which introduces the fixation time as a second important scale. The interaction of the mutation and fixation timescales also determines the outcome of clonal interference models. Our approach adds as a third component the ecological timescale. As we have shown, the inclusion of this scale can change the predicted pattern of fixations substantially. Models of constant selection predict that large mutations fix before small mutations (regardless of whether mutational effects are measured in terms of phenotype or fitness; ORR 2005a,b). Our model contains this result as a special case if all alleles are in the mutation-limited regime (*i.e.*, if their effects are sufficiently small). However, in the general case, it is mutations of intermediate size that fix first. If all mutations are in the environmentally limited regime, we can even observe a complete reversal of the pattern, with small mutations fixing before large ones.

So far, the classical prediction of large mutations fixing first has largely been confirmed by studies of microbial experimental evolution (ELENA and LENSKI 2003). However, almost all of these studies have subjected microbes to selection toward a fixed new optimum. One exception is the recent work by PERRON et al. (2008), who exposed bacteria to increasing concentrations of an antibiotic. However, these authors did not resolve the size and order of individual substitutions. Also, even though the selection pressure increased over time, the selection target itself (resistance) was fixed. In other words, selection was directional with increasing intensity, not stabilizing with a moving optimum, as assumed in our study. This probably leads to predictions different from the ones derived here, as large mutations (conferring a high degree of resistance) are likely to be selected for from the very beginning (unless there are strong trade-offs involved). The above example shows that gradual environmental change can imply different scenarios, which call for different modeling approaches.

In this article, we have focused on the order of adaptive substitutions. Another important question is the distribution of fixed mutations over a longer bout of adaptation (ORR 1998, 2002, 2005a,b). However, while the present biallelic model can successfully predict the size of the first fixation (Equation D5), it is less suited for studying the distribution of subsequent adaptive steps. The reason is that we assume only a limited number of mutant alleles, all of which eventually go to fixation. As a consequence, these mutations are no longer available for future steps. For long-term adaptation, a more realistic approach is to use a continuum-of-alleles model (e.g., BÜRGER 1999), in which the effects of new mutations at each step are picked from a (fixed) distribution. The main question in such a model is how the (longterm) distribution of fixed substitutions is related to the distribution of new mutations. This will be the subject of a separate study.

Key concepts and findings of the present study include the role of timescales and evolutionary regimes and the dependence of regime boundaries on the ecological parameters and the genetic architecture. They are not the results of specific model assumptions, but were developed and refined in a stepwise process from a single haploid locus to the full diploid multilocus case. We therefore expect that these findings also remain valid in a general context and can serve as guide for future model extensions.

We thank N. Barton, R. Bürger, C. Rueffler, A. de Visser, and two anonymous reviewers for helpful comments on the manuscript. M.K. and J.H. are members of the Mathematics and BioSciences Group at the University of Vienna, which is funded by a grant from the Vienna Science and Technology Fund to J.H.

LITERATURE CITED

- BARRETT, R. D. H., and D. SCHLUTER, 2008 Adaptation from standing genetic variation. Trends Ecol. Evol. 23: 38–44.
- BARTON, N. H., 1995 Linkage and the limits to natural selection. Genetics 140: 821–841.
- BARTON, N. H., and D. KEIGHTLY, 2002 Understanding quantitative genetic variation. Nat. Rev. Genet. 3: 11–21.
- BELLO, Y., and D. WAXMAN, 2006 Near-periodic substitutions and the genetic variance induced by environmental change. J. Theor. Biol. 239: 152–160.
- BÜRGER, R., 1999 Evolution of genetic variability and the advantage of sex and recombination in a changing environment. Genetics 153: 1055–1069.
- BÜRGER, R., 2000 The Mathematical Theory of Selection, Recombination, and Mutation. Wiley, Chichester, UK.
- BÜRGER, R., 2005 A multilocus analysis of intraspecific competition and stabilizing selection on a quantitative trait. J. Math. Biol. 50: 355–396.
- BÜRGER, R., and A. GIMELFARB, 2002 Fluctuating environments and the role of mutation in maintaining quantitative genetic variation. Genet. Res. 80: 31–46.
- BÜRGER, R., and M. LYNCH, 1995 Evolution and extinction in a changing environment: a quantitative genetic analysis. Evolution 49: 151–163.
- COLLINS, S., J. DE MEAUX and C. ACQUISTI, 2007 Adaptive walks toward a moving optimum. Genetics 176: 1089–1099.
- ELENA, S. F., and R. E. LENSKI, 2003 Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. Nat. Rev. Genet. 4: 457–469.
- ETHERIDGE, A., P. PFAFFELHUBER and A. WAKOLBINGER, 2006 An approximate sampling formula under genetic hitchhiking. Ann. Appl. Probab. 16: 685–729.
- EWENS, W. J., 2004 Mathematical Population Genetics, Ed. 2. Springer-Verlag, Berlin.
- EYRE-WALKER, A., and P. D. KEIGHTLEY, 2007 The distribution of fitness effects of new mutations. Nat. Rev. Genet. 8: 610–618.
- FISHER, R., 1930 The Genetical Theory of Natural Selection. Clarendon Press, Oxford.
- GERRISH, P. J., and R. E. LENSKI, 1998 The fate of competing beneficial mutations in an asexual population. Genetica 102/103: 127–144.
- GILLESPIE, J. H., 1983 A simple stochastic gene substitution model. Theor. Popul. Biol. 23: 202–215.
- GILLESPIE, J. H., 1984 Molecular evolution over the mutational landscape. Evolution 38: 1116–1129.
- GILLESPIE, J. H., 1993 Substitution processes in molecular evolution. I. Uniform and clustered substitutions in a haploid model. Genetics 134: 971–981.

- HAIRSTON, N. G., S. P. ELLNER, M. A. GEBER, T. YOSHIDA and J. A. FOX, 2005 Rapid evolution and the convergence of ecological and evolutionary time. Ecol. Lett. 8: 1114–1127.
- HERMISSON, J., and P. S. PENNINGS, 2005 Soft sweeps: molecular population genetics of adaptation from standing genetic variation. Genetics 169: 2335–2352.
- HILL, W. G., and A. ROBERTSON, 1966 The effect of linkage on limits to artificial selection. Genet. Res. 8: 269–294.
- KAUFFMAN, S. A. A., 1993 Origins of Order: Self-Organization and Selection in Evolution. Oxford University Press, Oxford.
- KAUFFMAN, S. A. A., and S. LEVIN, 1987 Towards a general theory of adaptive walks on rugged landscapes. J. Theor. Biol. 128: 11–45.
- KIM, Y., and H. A. ORR, 2005 Adaptation in sexuals vs. asexuals: clonal interference and the Fisher-Muller model. Genetics 171: 1377–1386.
- KIMURA, M., 1983 The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge, UK.
- KOPP, M., and J. HERMISSON, 2007 Adaptation of a quantitative trait to a moving optimum. Genetics 176: 715–718.
- LYNCH, M., and R. LANDE, 1993 Evolution and extinction in response to environmental change, pp. 43–53 in *Biotic Interactions* and *Global Change*, edited by P. KAREIVA, J. G. KINGSOLVER and R. B. JUEY. Sinauer Associates, Sunderland, MA.
- LYNCH, M., W. GABRIEL and A. M. WOOD, 1991 Adaptive and demographic responses of plankton populations to environmental change. Limnol. Oceanogr. 36: 1301–1312.
- 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.
- NUNNEY, L., 2003 The cost of natural selection revisited. Ann. Zool. Fenn. 40: 185–194.
- ORR, H. A., 1998 The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution 52: 935–949.
- ORR, H. A., 2002 The population genetics of adaptation: the adaptation of DNA sequences. Evolution 56: 1317–1330.
- ORR, H. A., 2005a The genetic theory of adaptation: a brief history. Nat. Rev. Genet. 6: 119–127.
- Orr, H. A., 2005b Theories of adaptation: what they do and don't say. Genetica **123**: 3–13.
- PARK, S.-C., and J. KRUG, 2007 Clonal interference in large populations. Proc. Natl. Acad. Sci. USA **104**: 18135–18140.
- PARMESAN, C., 2006 Ecological and evolutionary responses to recent climate change. Annu. Rev. Ecol. Evol. Syst. 37: 637–669.
- PERRON, G. G., A. GONZALEZ and A. BUCKLING, 2008 The rate of environmental change drives adaptation to an antibiotic sink. J. Evol. Biol. 21: 1724–1731.
- PROVINE, W. B., 2001 The Origins of Theoretical Population Genetics. University of Chicago Press, Chicago.
- SATO, M., and D. WAXMAN, 2008 Adaptation to slow environmental change, with apparent anticipation of selection. J. Theor. Biol. 252: 166–172.
- THOMPSON, J. N., 2005 The Geographic Mosaic of Coevolution. University of Chicago Press, Chicago.
- WAXMAN, D., and J. R. PECK, 1999 Sex and adaptation in a changing environment. Genetics 153: 1041–1053.
- WELCH, J. J., and D. WAXMAN, 2005 Fisher's microscope and Haldane's ellipse. Am. Nat. 166: 447–457.

Communicating editor: M. W. FELDMAN

APPENDIX A: SIMULATION METHODS

Here, we describe the computer simulations. We assume that the population has a constant size N, individuals are hermaphroditic, and generations are nonoverlapping. Simulations start at time t = 0, where the population is assumed to be monomorphic for the wild-type alleles at each locus. Each generation is modeled in two steps (see KOPP and HERMISSON 2007). First, we use deterministic equations to calculate the *expected* genotype frequencies after selection, (free) recombination, and mutation. Then, we use stochastic multinomial sampling to obtain the *actual* genotype frequencies in a finite population subject to genetic drift (*e.g.*, GILLESPIE 1993). This means that the next generation is obtained by drawing N individuals (with replacement) from the frequency distribution

calculated in step one. (In the diploid case, 2N haplotypes are drawn, assuming Hardy–Weinberg proportions before selection.)

As the number of loci (and hence the number of genotypes) increases, this model rapidly becomes computationally intractable. For the multilocus case, we therefore used a weak-selection approximation: As long as the population mean phenotype is sufficiently close to the optimum, linkage disequilibria can be neglected, and fitness can be approximated by the quadratic function

$$w(z, t) \approx 1 - \sigma [z - z_{\text{opt}}(t)]^2.$$
(A1)

This makes it possible to model the evolution of allele frequencies (instead of genotype frequencies) directly, using the equations

$$\Delta p_i = p_i + p_i (1 - p_i) \sigma \alpha^2 \left(2p_i - 1 - 2 \frac{\bar{x}_t - z_{\text{opt}}(t)}{\alpha_i} \right)$$
(A2)

(*e.g.*, BÜRGER 2005), where p_i is the frequency of the mutant allele at locus *i*, Δp_i is the change in p_i from one generation to the next, and $\bar{x}_t = \sum_{i=1}^{L} p_i \alpha_i$ is the population mean phenotype. When using this approximation, sampling is performed on alleles instead of genotypes.

APPENDIX B: DERIVATION OF THE FIXATION TIME $T_{\rm f}$

Here, we derive the narrow-sense fixation time $T_{\rm f}$. For ease of notation, we define an alternative timescale \tilde{t} , which measures time since the start of the fixation process (so that $\tilde{t} = t - T_{\ell} - T_{\rm w}$). The selection coefficient on this timescale is given by

$$s(\tilde{t}) = \lambda(T_{\rm w} + \tilde{t}). \tag{B1}$$

Because $s(\tilde{t})$ depends on the waiting time, so does T_f . Once the fixation process has started (at $\tilde{t} = 0$), the dynamics of the mutant allele frequency p can be approximated deterministically by the differential equation

$$\dot{p} = s(\tilde{t})p(1-p) + \mu(1-2p) + \frac{1}{2N}.$$
 (B2)

Here, the first term on the right-hand side describes logistic growth due to selection, the second term describes the input from mutation (both forward and backward), and the third term describes the contribution of genetic drift to the frequency increase of a mutation that is conditioned to go to fixation (EwENS 2004). Numerical solutions of this equation are in excellent agreement with simulation results (not shown). Unfortunately, Equation B2 does not have a closed analytical solution. To derive a tractable approximation, we neglect selection as long as it is weaker than mutation and drift, and we neglect mutation and drift when, together, they are weaker than selection. As a consequence, the fixation time is subdivided into two components,

$$T_{\rm f} = T_{\rm f,1} + T_{\rm f,2},\tag{B3}$$

where $T_{f,1}$ is the time span dominated by mutation and drift, and $T_{f,2}$ is the time span dominated by selection. In the following, we need to distinguish between fixation from a new mutation and fixation from an already segregating allele.

Fixation from a new mutation: For $\tilde{t} < T_{f,1}$, we neglect selection (and back mutations) and assume that p initially increases due to mutation and drift alone, yielding $p(\tilde{t}) \approx ((\theta + 1)/2N)\tilde{t}$. Mutation and drift remain dominant until $s(\tilde{t})p(1-p) \approx s(\tilde{t})p > (\theta + 1)/2N$. Using (B1), this yields

$$T_{\rm f,1}^{\rm n} = \frac{1}{2} \left(\sqrt{T_{\rm w}^2 + 4/\lambda} - T_{\rm w} \right). \tag{B4}$$

At this time, the frequency of the mutant allele is

$$p(T_{\rm f,1}^{\rm n}) = p^* = \frac{\theta + 1}{2Ns^*}$$
(B5a)

with

$$s^* = \lambda (T_w + T_{f,1}^n). \tag{B5b}$$

This approximation is valid as long as p^* is small. At the very least, it must be <0.5, which is true if $s^* > 2\mu + 1/N$. For $\tilde{t} > T_{f,1}^n$, we neglect drift and mutation, such that Equation B2 becomes $\dot{p} = s(\tilde{t})p(1-p)$. With the initial

condition $p(T_{f,1}^n) = p^*$, this can be solved to yield

$$p(\tilde{t}) = \frac{p^* \exp[S(\tilde{t})]}{1 + p^* (\exp[S(\tilde{t})] - 1)},$$
(B6a)

where

$$S(\tilde{t}) = \int_{T_{t,1}}^{\tilde{t}} s(x) dx = s^* \Delta \tilde{t} + \frac{\lambda}{2} (\Delta \tilde{t})^2$$
(B6b)

and

$$\Delta \tilde{t} = \tilde{t} - T_{\rm f,1}^{\rm n},\tag{B6c}$$

and the condition $p(T_{f,2}^n) = \frac{1}{2}$ (our criterion for fixation) leads to

$$T_{\rm f,2}^{\rm n} = \frac{\sqrt{(s^*)^2 + 2\lambda \ln(2Ns^*/(\theta+1) - 1)} - s^*}{\lambda}.$$
 (B7)

Equation 8a follows from (B5b) and (B3).

Fixation from an already segregating allele: If fixation occurs from an already segregating allele, $T_w = 0$ and thus $\tilde{t} = t - T_\ell$ and $s = \lambda \tilde{t}$. The expected frequency of the mutant allele at the end of the lag time, p(0), can be calculated as follows: During the lag time (*i.e.*, for $\tilde{t} < 0$), p is small, and Equation B2 can be approximated by $\dot{p} = \mu + \lambda \tilde{t}p$. [Note that, here, the drift term 1/(2N) is omitted, because we are not conditioning on eventual fixation.] This can be solved to yield

$$p(\tilde{t}) = \exp(\frac{1}{2}\lambda \tilde{t}^2) \left(C + \frac{\mu}{\sqrt{\lambda}} \int_0^{\sqrt{\lambda/2}\tilde{t}} \exp(-x^2) dx \right).$$
(B8)

With the initial condition $p(-\infty) = 0$, we find

$$p(0) = \mathcal{C} = \frac{\mu}{2} \sqrt{\frac{\pi}{\lambda}}.$$
(B9)

For $\tilde{t} > 0$, we follow the same approach as for fixation from a new mutation. For $\tilde{t} < T_{f,1}^s$, p increases linearly according to $p(\tilde{t}) = p(0) + \mu \tilde{t}$ until $p = p^*$ (Equation B5a), leading to

$$T_{\rm f,1}^{\rm s} = \frac{1}{\sqrt{\lambda}} \left(\sqrt{\frac{\pi}{16} + 1} - \sqrt{\frac{\pi}{16}} \right) \approx \frac{0.65}{\sqrt{\lambda}}.$$
 (B10)

 $T_{f,2}^s$ is again given by (B7), but with $s^* = \lambda T_{f,1}^s$. Finally, the probability of fixation from an already segregating allele can be estimated as

$$\mathcal{P}^{s} = 1 - \exp\left(-p(0)N\sqrt{\frac{8\lambda}{\pi}}\right) \tag{B11}$$

(J. HERMISSON, unpublished data), where p(0) is the expected frequency of the mutant allele at the end of the lag time. Inserting (B9) into (B11) directly leads to (15).

M. Kopp and J. Hermisson

APPENDIX C: ANALYTICAL APPROXIMATION FOR α^*

Here, we derive an analytical approximation for α^* , the value of α that minimizes the total time to fixation in the one-locus model. Ignoring the dependence on α in the logarithmic term in Equation 14a (*e.g.*, by setting $\alpha = 1$ in this term only), the total time to fixation can be written as

$$\mathbb{T} \approx \frac{\alpha}{2v} + \frac{C(\sigma, \theta, v, N)}{\sqrt{\alpha}},\tag{C1a}$$

where the constant C is given by

$$C(\sigma, \theta, v, N) = (1 - \mathcal{P}^{S})C_{1}(\sigma, \theta, v, N) + \mathcal{P}^{S}C_{2}(\sigma, \theta, v, N)$$
(C1b)

with

$$C_{1}(\sigma, \theta, v, N) = \sqrt{\frac{\pi(\sqrt{1+8\theta/\pi}+1)^{2}}{16\theta\sigma v}} + \frac{\ln(2N\sqrt{2\sigma v}(\sqrt{\pi/8\theta+1}+\sqrt{\pi/8\theta})/(\theta+1)-1)}{\sigma v},$$
(C1c)

$$C_2(\sigma, \theta, v, N) = \sqrt{\frac{\pi(\sqrt{1+16/\pi}-1)^2}{32\sigma v} + \frac{\ln(2N\sqrt{2\sigma v}(\sqrt{\pi/16+1}-\sqrt{\pi/16})/(\theta+1)-1)}{\sigma v}}.$$
 (C1d)

In the following, we suppress the dependence on parameters in our notation. \mathbb{T} has a minimum at

$$\alpha^* = (vC)^{2/3}.$$
 (C2)

This approximation for α^* is very accurate for small θ , but shows small deviations from the true value if θ is large (see Figure 2). Furthermore, plugging (C2) into (C1a) shows that, for a mutation of size α^* , the total time to fixation, \mathbb{T}^* , is given by

$$\mathbb{T}^* = \frac{(vC)^{2/3}}{2v} + \frac{C}{(vC)^{1/3}} = \frac{1}{2} \left(\frac{C^2}{v}\right)^{1/3} + \left(\frac{C^2}{v}\right)^{1/3},\tag{C3}$$

where the first term corresponds to the lag time and the second term to the sum of waiting time and fixation time. Therefore, within the limits of the approximation, the lag time of the fastest mutation is exactly one-third of the total time to fixation.

Scaling relationships: We can use the above results to derive several useful scaling relationships. First, plugging (C1b) to (C1d) into (C1a) and ignoring the logarithmic term shows that α^* is proportional to $\sqrt[3]{v/\sigma}$. For small θ , $\mathcal{P}^{\rm S} \approx 0$, and Equation C2 reduces to $\alpha^* = (vC_1)^{2/3}$, which, if we again ignore the logarithmic term, is proportional to $\sqrt[3]{v/\sigma}$.

If we continue to focus on the case with small θ , we further see from (11) that $s^* \approx \sqrt{\lambda \pi/2\theta} = \lambda \overline{T}_w$. Thus, using (7b) and (8a), $T_w + T_f \approx \sqrt{T_w^2 + 2\ln(\ldots)/\lambda} \approx \overline{T}_w + \ln(\ldots)/(\lambda \overline{T}_w)$. For small θ , this sum is dominated by \overline{T}_w ; that is, T_f is small relative to T_w (and negatively correlated with it). Thus, the total time to fixation is determined by two components: the lag time T_ℓ , which is proportional to α/v , and the waiting time T_w (or rather, the sum of $T_w + T_f$), which is proportional to $1/\sqrt{\sigma v \theta \alpha}$. However, if we choose to measure mutational effects in units of $v/(\sigma \theta)$ —that is, we measure them relative to α^* —we find that both T_ℓ and $T_w + T_f$ can be measured in units of $1/\sqrt[3]{v^2\sigma\theta}$. At this scale, T_ℓ is proportional to α , and $T_w + T_f$ is proportional to $1/\sqrt{\alpha}$, but both are independent of the other parameters. In other words, α^* defines a characteristic mutational step size and a characteristic time for adaptation.

APPENDIX D: THE DISTRIBUTION OF THE FIRST FIXATION IN THE MULTILOCUS MODEL

Here, we derive the predicted distribution of the first fixation in the multilocus model. Assume that there are L loci with two alleles each. What is the probability that the mutant allele at locus i is the first to reach fixation? In the following, we present an approximation that is an extension of the two-locus approach by KOPP and HERMISSON (2007). The key simplification is that allele frequency changes at different loci are assumed to be independent (neglecting epistatic interactions for fitness).

We first need to combine the cases of fixation from a new mutation and from an already segregating allele. By a slight abuse of notation, we use $T_w = 0$ to indicate fixation from a segregating allele. We also introduce additional indexes for the focal locus. Using (4b), (7a), and (15), the waiting time distribution (7a) at locus *i* then becomes

$$g_i(T_{\mathbf{w}_i}) = \begin{cases} 0 & \text{for } T_{\mathbf{w},i} < 0\\ 1 - \exp\left(-\frac{\theta}{\sqrt{2}}\right) & \text{for } T_{\mathbf{w},i} = 0\\ \theta \lambda T_{\mathbf{w},i} \exp\left(-\frac{\theta}{\sqrt{2}} - \frac{\theta \lambda}{2} T_{\mathbf{w},i}^2\right) & \text{for } T_{\mathbf{w},i} > 0. \end{cases}$$
(D1)

Note that g_i has a discontinuity at $T_{w,i} = 0$. In particular, $g_i(0) \neq \lim_{T_{w,i} \to 0^+} g_i(T_{w,i})$. The corresponding cumulative distribution function is

$$G_{i}(T_{w,i}) = \int_{\tau=0}^{T_{w,i}} \tilde{f}_{i}(\tau) d\tau = \begin{cases} 0 & \text{for } T_{w,i} < 0\\ 1 - \exp\left(-\frac{\theta}{\sqrt{2}}\right) & \text{for } T_{w,i} = 0\\ 1 - \exp\left(-\frac{\theta}{\sqrt{2}} - \frac{\theta\lambda_{i}}{2}T_{w}^{2}\right) & \text{for } T_{w,i} > 0. \end{cases}$$
(D2)

For a given waiting time $T_{w,i}$, the total time to fixation is

$$\mathbb{T}_{i}(T_{\mathbf{w},i}) = \frac{\alpha_{i}}{2v} + \frac{1}{\lambda_{i}} \sqrt{s_{i}^{*}(T_{\mathbf{w},i})^{2} + 2\lambda_{i} \ln\left(\frac{2Ns_{i}^{*}(T_{\mathbf{w},i})}{\theta + 1} - 1\right)},\tag{D3}$$

where

$$s_{i}^{*}(T_{w,i}) = \begin{cases} \sqrt{\lambda_{i}} \left(\sqrt{\frac{\pi}{16} + 1} - \sqrt{\frac{\pi}{16}} \right) & \text{for } T_{w,i} = 0, \\ \frac{\lambda_{i}}{2} \left(T_{w} + \sqrt{T_{w}^{2} + \frac{4}{\lambda_{i}}} \right) & \text{for } T_{w,i} > 0. \end{cases}$$
(D4)

Assume the mutant allele at locus *i* has waiting time $T_{w,i}$. Then a mutant allele at another locus *j* will fix first only if its waiting time is less than $T_{w,j}^*$ ($T_{w,i}$), which can be obtained by numerically solving the equation $\mathbb{T}_i(T_{w,i}) = \mathbb{T}_j(T_{w,j})$ for $T_{w,j}$. [More precisely, we need to distinguish three cases: If $\mathbb{T}_i(T_{w,i}) < \mathbb{T}_j(0)$, then the allele at locus *j* can never fix before the one at locus *i*, even if it starts from an already segregating allele. Thus, $T_{w,j}^*$ can be set to a negative value (which has zero probability). If $\mathbb{T}_j(0) \leq \mathbb{T}_i(T_{w,i}) < \lim_{T_{w,j} \to 0^+} \mathbb{T}_j(T_{w,j})$, then the allele at locus *j* needs to start from a segregating allele and thus $T_{w,j}^* = 0$. Finally, if $\lim_{T_{w,j} \to 0^+} \mathbb{T}_j(T_{w,j}) \leq \mathbb{T}_i(T_{w,i})$, then the allele at locus *j* can start from a new mutation, and $T_{w,j}^*$ is the solution to the above equation.] Finally, the probability π_i that the mutant allele at locus *i* fixes before all other mutant alleles is given by

$$\pi_i = \int_0^\infty g_i(T_{\mathbf{w},i}) \prod_{j \neq i} (1 - G_j(T_{\mathbf{w},j}^*(T_{\mathbf{w},i})) dT_{\mathbf{w},i}.$$
(D5)