

1 The genetic basis of phenotypic adaptation I: Fixation of  
2 beneficial mutations in the moving optimum model

3 Michael Kopp      Joachim Hermisson

4

5 Mathematics and Biosciences Group, Max F. Perutz Laboratories and Faculty of Mathemat-  
6 ics, University of Vienna, Dr. Bohr-Gasse 9, A-1030 Vienna, Austria

7 Author for correspondence: Michael Kopp, Dr. Bohrgasse 9, A-1030 Vienna, Austria. Phone:  
8 ++43-1-79044-4583, Fax: ++43-1-4277-9240, Email: michael.kopp@univie.ac.at

9 **Running Head**

10 Adaptation to a moving optimum

## 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. Based on 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 time scales acting in the system – the environmental, mutation, and fixation time scale – which define three parameter regimes leading to qualitative differences in the adaptive substitution pattern.

## Introduction

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- versus macro-mutationalism 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

1 models of so-called “adaptive walks” on rugged fitness landscapes (e.g., KAUFFMAN and  
2 LEVIN 1987; KAUFFMAN 1993). These approaches treat organismal adaptation as a process  
3 in a high-dimensional phenotype or genotype space characterized by pleiotropy and epistasis.  
4 Another common feature is that all these models assume a constant selection pressure. Yet,  
5 many organisms live in environments that are continually changing (e.g., HAIRSTON *et al.*  
6 2005; THOMPSON 2005; PARMESAN 2006; PERRON *et al.* 2008). This fact is acknowledged in  
7 the so-called moving optimum model, which assumes that the selectively favored phenotype  
8 changes over time and, hence, experiences a mixture of stabilizing and directional selection.

9 The moving optimum model was originally devised in the field of quantitative genetics, where  
10 its primary aim was to study adaptation at the phenotypic level, without an explicit focus  
11 on the underlying genetics. In this context, the model has been used, for example, to study  
12 the ability of populations to avoid extinction in the face of global change (e.g., LYNCH *et al.*  
13 1991; LYNCH and LANDE 1993; BÜRGER and LYNCH 1995; NUNNEY 2003), to study the  
14 role of temporal variation for the maintenance of genetic variation (BÜRGER 2000; BÜRGER  
15 and GIMELFARB 2002), and to understand the evolution of sex and recombination (BÜRGER  
16 1999; WAXMAN and PECK 1999). Only recently, several studies have started to investigate  
17 how a moving optimum affects the dynamics of adaptive substitutions (BELLO and WAXMAN  
18 2006; KOPP and HERMISSON 2007; COLLINS *et al.* 2007; SATO and WAXMAN 2008). A key  
19 result is that, under conditions of slow environmental change, mutations with small effect  
20 tend to become fixed earlier than those with large effect (BELLO and WAXMAN 2006; KOPP  
21 and HERMISSON 2007; COLLINS *et al.* 2007). This is in direct contrast to what is found under  
22 constant selection after a sudden change in the environment (e.g. ORR 1998, 2002, 2005a;  
23 KIM and ORR 2005).

24 In a previous note (KOPP and HERMISSON 2007), we have identified three parameter regimes  
25 in the moving optimum model that lead to qualitative differences in the dynamics of adaptive  
26 substitution. These regimes are defined by three time-scales that may dominate the adaptive  
27 process: the ecological time scale, which is given by the (inverse) speed of the optimum; the  
28 mutation time scale, which determines the waiting time for a new beneficial allele; and the  
29 time that it takes for this allele to rise to fixation. In particular, we found that small-effect  
30 mutations are fixed before large ones if the ecological time scale is dominating, that is, if

1 the speed of adaptation is limited primarily by the speed of environmental change (see also  
2 COLLINS *et al.* 2007 for similar findings). In KOPP and HERMISSON (2007), this result was  
3 obtained from an approximate calculation in a minimal model with two haploid biallelic loci.  
4 Here, we extend the previous analysis in two main ways. First, we significantly improve our  
5 approximation for the total fixation time of a given allele under moving-optimum selection.  
6 Second, we analyze the order of adaptive substitutions in a full haploid or diploid multilocus  
7 model. We confirm the preliminary results from the two-locus study, but show that regime  
8 boundaries depend on both the ecological parameters and the genetic architecture of the  
9 trait. The current paper focuses on the order of substitutions over relatively short time-  
10 scales. The complementary question about which mutations dominate long-term evolution  
11 will be treated in a separate study.

## 12 **The model**

### 13 **Assumptions on fitness**

14 Consider the evolution of a quantitative trait  $z$  that is under stabilizing selection with respect  
15 to a moving optimum  $z_{\text{opt}} = z_{\text{opt}}(t)$ . We assume that the optimum increases over time at a  
16 constant rate  $v$ , that is

$$z_{\text{opt}}(t) = vt. \quad (1)$$

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

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

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

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

1 As long as selection is weak ( $\sigma$  or  $|z - z_{\text{opt}}|$  small), equation (3) can be approximated by

$$s(t) \approx \lambda\left(t - \frac{\alpha}{2v}\right), \text{ where} \quad (4a)$$

$$\lambda = 2\sigma v\alpha. \quad (4b)$$

2 Thus, selection for  $\alpha$  increases approximately linearly over time.

### 3 **Assumptions on genetics**

4 The aim of this paper is to model the first steps of the adaptive process in the moving  
5 optimum model. We assume that the trait  $z$  is influenced by  $L$  additive, unlinked, haploid or  
6 diploid loci. Each locus has two alleles, which we will refer to as the wild-type and mutant  
7 allele, respectively. The contribution of the wild-type allele to the phenotype  $z$  is 0, and the  
8 contribution of the mutant allele is  $\alpha_i$  ( $i = 1 \dots L$ ). The  $\alpha_i$  will be referred to as (locus)  
9 mutational effects. In the one-locus case, we will suppress the index  $i$  and simply write  
10  $\alpha_1 = \alpha$ . Mutations from the wild-type to the mutant allele (and vice versa) occur recurrently  
11 at rate  $\mu$  per locus. In our calculations, we will usually use the population mutation parameter  
12  $\theta = 2N\mu$ , where  $N$  is the population size. Environmental variation is not modelled explicitly,  
13 but is subsumed in the selection parameter  $\sigma$  (e.g., BÜRGER 2000). Our analytical results  
14 will be compared to and extended by stochastic computer simulations, which are described  
15 in Appendix 1.

## 16 **Results**

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

## 1 **Expected time to fixation of a single mutant allele**

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

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

17 Denote the total time to fixation (i.e., to  $p = 1/2$ ) by  $\mathbb{T}$ . Following KOPP and HERMISSON  
18 (2007), this time can be subdivided into three periods:

$$\mathbb{T} = T_\ell + T_w + T_f, \tag{5}$$

19 where

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

20 is the *lag time* until the mutant allele becomes beneficial (i. e., has a positive selection  
21 coefficient, see eq. 2).  $T_w$  is the *waiting time* for a successful mutation (i.e., for the appearance  
22 of a mutant allele that is not subsequently lost due to drift). Finally,  $T_f$  is the (narrow-sense)  
23 *fixation time* of this successful mutation (i.e., the time needed for the frequency to increase  
24 from  $p = 1/(2N)$  to  $p = 1/2$ ). To calculate the waiting and fixation times, we need to

1 distinguish two different cases. In the first case, fixation occurs from a mutant allele that  
 2 appears after the end of the lag time, so that the waiting time is positive. We will refer  
 3 to this case as fixation from a “new” mutation. In the second case, fixation occurs from a  
 4 mutant allele that is already segregating in the population at the end of the lag time, and  
 5 the waiting time is zero.

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

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

(KOPP and HERMISSON 2007), which has mean

$$\bar{T}_w = \int_0^\infty T_w f(T_w) dT_w = \sqrt{\frac{\pi}{2\theta\lambda}} \quad (7b)$$

and variance

$$\text{Var}(T_w) = \frac{4 - \pi}{2\theta\lambda}. \quad (7c)$$

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

17 The fixation time  $T_f$  also is a random variable, but its variance is small relative to that of the  
 18 waiting time (as is well-known for constant selection, cf. ETHERIDGE *et al.* 2006). Therefore,  
 19 we will treat it as deterministic. However, for fixation from a new mutation,  $T_f =: T_f^n$  still  
 20 depends on the stochastic waiting time, because  $T_w$  determines the selection pressure during  
 21 the fixation phase. In Appendix 2, we derive

$$T_f^n(T_w) = \frac{\sqrt{(s^*)^2 + 2\lambda \ln\left(\frac{2Ns^*}{\theta+1} - 1\right)}}{\lambda} - T_w, \quad (8a)$$

where

$$s^* = \frac{\lambda}{2} \left( T_w + \sqrt{T_w^2 + \frac{4}{\lambda}} \right). \quad (8b)$$

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

$$E(\mathbb{T}^n) = T_\ell + \int_{T_w=0}^{\infty} [T_w + T_f^n(T_w)] f(T_w) dT_w. \quad (9)$$

2 Since  $T_f^n$  (as a function of  $T_w$ ) is approximately linear around  $T_w = \bar{T}_w$  (not shown), equa-  
 3 tion (9) can be approximated by

$$E(\mathbb{T}^n) \approx T_\ell + \bar{T}_w + T_f^n(\bar{T}_w). \quad (10)$$

4 Inserting  $\bar{T}_w$  from (7b) into (8b), the value of  $s^*$  becomes

$$s^*(T_w = \bar{T}_w) = \sqrt{\lambda} \left( \sqrt{\frac{\pi}{8\theta} + 1} + \sqrt{\frac{\pi}{8\theta}} \right), \quad (11)$$

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

6 **Case 2: Fixation from an already segregating allele.** If  $\theta$  is large, there is a high  
 7 probability that fixation occurs from a mutant allele that has appeared already during the  
 8 lag time. In this case,  $T_w = 0$  and the fixation time  $T_f =: T_f^s$  is given by the same expression  
 9 as  $T_f^n$  (8a), but with

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

10 (see Appendix 2). The expected total time to fixation from an already segregating alleles,  
 11  $\mathbb{T}^s$ , is simply

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

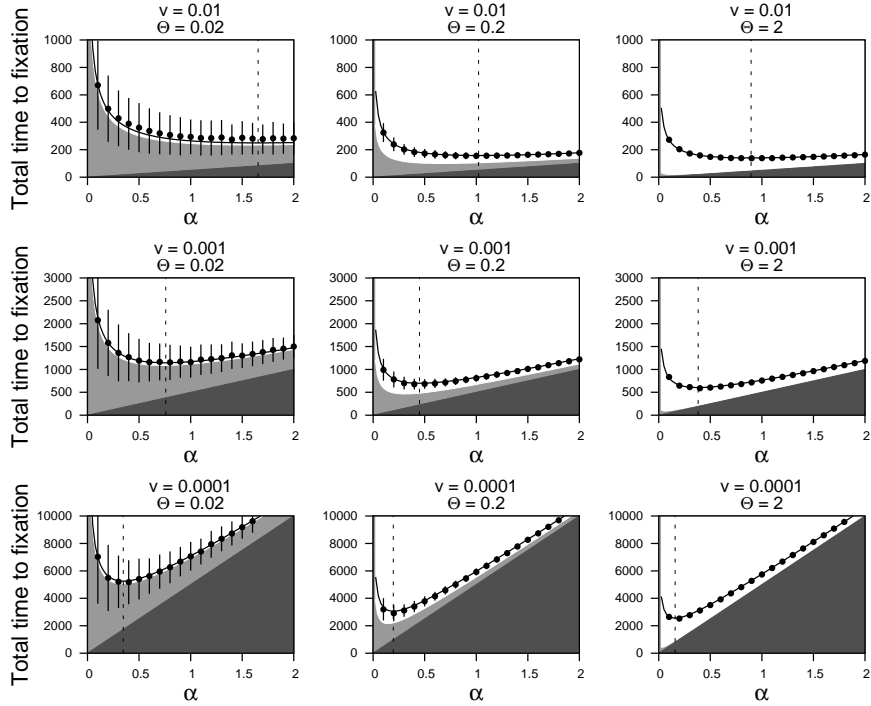


Figure 1: The total time to fixation (i.e., time until frequency  $p > 1/2$ ) of a mutant allele in the biallelic one-locus model as a function of the mutational effect  $\alpha$ . Each plot shows means and standard deviations from 1000 simulation runs, together with the prediction from the analytical approximation in equation (16) (solid line). The dotted line indicates  $\alpha^*$ , the value of  $\alpha$  with the shortest time to fixation (eq. A15). The dark grey area indicates the lag time  $T_\ell$  (equation 6) and the light grey area the mean waiting time (i.e.,  $\bar{T}_w$  from eq. 7b multiplied by the probability for fixation from a new mutation from eq. 15).  $\theta = 2N\mu$  with  $\mu = 10^{-5}$ .

- 1 In summary, combining equations (8a), (10), (13), (A7b), and (A9), the total time to fixation
- 2 can be estimated as

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

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

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

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

1 Thus, the expected total time to fixation from an either new or segregating alleles

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

2 Numerical analysis of equation (16) shows that the total time to fixation for a mutant with  
3 a given effect decreases with increasing  $v$ ,  $\sigma$  and  $\theta$ . This means that adaptation is fastest  
4 if the environment changes quickly, selection is strong, and the population-wide mutation  
5 rate is high. The predictions from equation (16) are in excellent agreement with results  
6 from stochastic simulations (Fig. 1). Visible deviations occur only for the combination of  
7 large  $v$  and small  $\theta$  ( $v = 0.01$ ,  $\theta = 0.02$  in Fig. 1), where the predicted time to fixation is  
8 slightly shorter than the observed one. The reason is that, for these parameters, the selection  
9 coefficient may become quite strong before a beneficial allele arrives. In this case, assuming a  
10 fixation probability of  $2s(t)$  (as is done in the derivation of eq. 7a) is an overestimation. The  
11 true fixation probabilities are lower and, hence, the true waiting times are longer. With the  
12 more accurate approximation for the fixation probability,  $1 - \exp(-2s(T_w + T_\ell))$ , the waiting  
13 time distribution becomes

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

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

16 For given,  $v$ ,  $\sigma$  and  $\theta$ , the total time to fixation is a U-shaped function of  $\alpha$  (Fig. 1). That  
17 is, mutations with intermediate effect become fixed faster than mutations with either small  
18 or large effects. The reason is that mutations with small effects are only weakly selected for  
19 and, thus, have long waiting times (due to a low fixation probability) and long fixation times.  
20 In contrast, mutations with large effects are deleterious for a long time and, therefore, have  
21 long lag times.

22 The value of  $\alpha$  that minimizes the total time to fixation will be denoted by  $\alpha^*$ . In Appendix 3,  
23 we derive an analytical approximation (eq. A15), which elucidates how  $\alpha^*$  depends on the

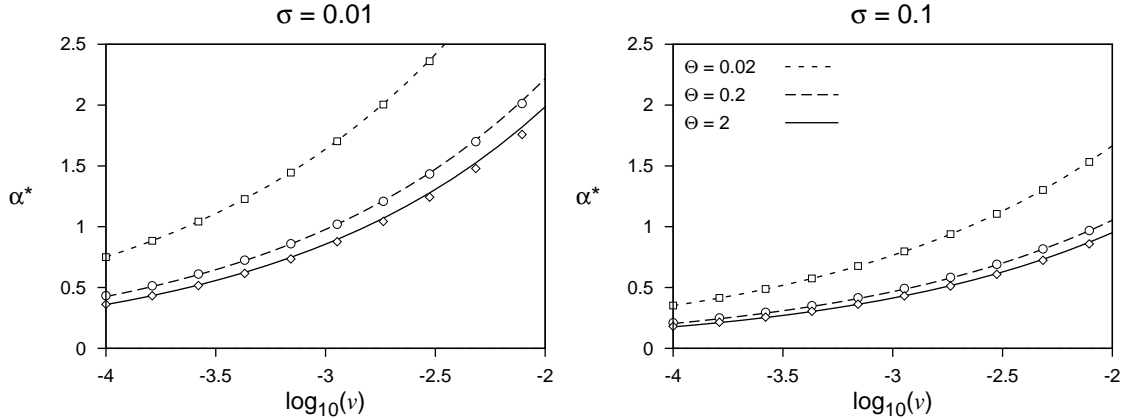


Figure 2:  $\alpha^*$ , the allelic effect of the mutant allele with the shortest time to fixation in the biallelic one-locus model, as function of the speed of the environmental change  $v$  for various values of the selection parameter  $\sigma$  and the mutation parameter  $\theta = 2N\mu$ . Lines are numerically calculated minima of equation (16) (see Fig. 1). Symbols show the approximation (A15).

1 model parameters (see Fig. 2). To a good approximation,

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

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

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

2 Thus, fast environmental change favors relatively large mutations, whereas a large population-  
3 wide mutation rate and strong selection both favor relatively small mutations. The reason  
4 is that fast environmental change leads to the build-up of a large lag between the wild-type  
5 phenotype and the optimal phenotype, which can then be bridged by a large mutation. In  
6 contrast, high mutation rates and strong selection lead to the fixation of small mutations  
7 before the lag becomes large. For a mutation of size  $\alpha^*$  we show in Appendix 3 that the  
8 lag time is approximately one third of the total time to fixation, suggesting that none of the  
9 three time scales is dominating the other two.

10 **The diploid case** So far, we have assumed that the locus under study is haploid. Here,  
11 we show how the previous analysis can be extended to the diploid case. We will assume that

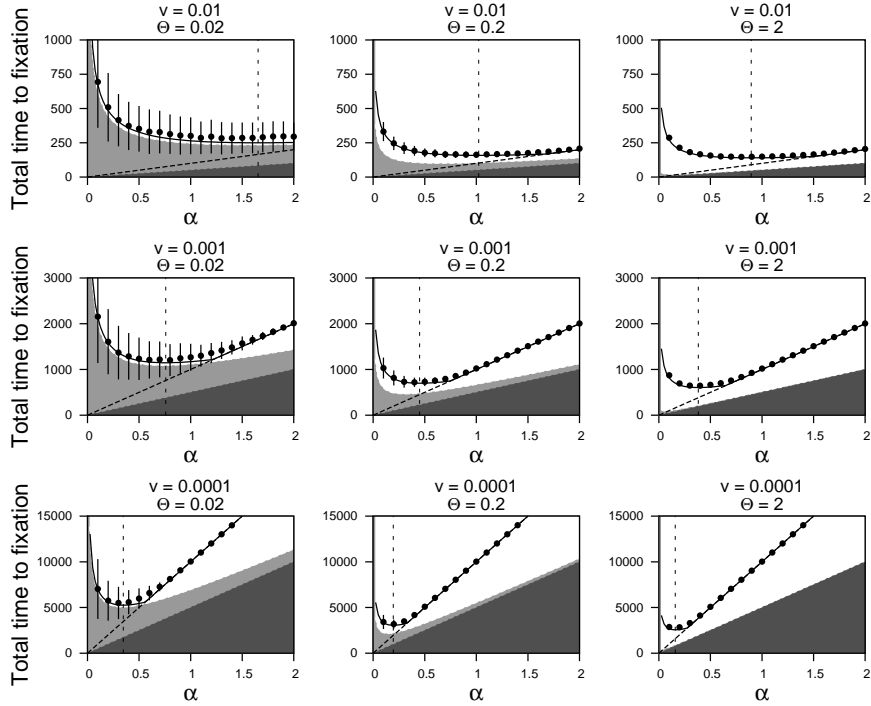


Figure 3: The total time to fixation of a single mutant allele in the diploid model. The Figure shows the additive case: Heterozygotes have phenotype  $\alpha$ , and mutant homozygotes have phenotype  $\beta = 2\alpha$ . The diagonal dashed line is  $\alpha/v$ , the time at which the mutant homozygote has the same fitness as the wild-type. The solid line is the maximum of the predicted time to fixation in the haploid model (eq. 16) and  $\alpha/v$ . The vertical dotted line is  $\alpha^*$ , the effect of the fastest mutation in the haploid model (eq. A15).  $\mu$  is the diploid mutation rate. For further details, see Fig. 1.

- 1 heterozygotes have phenotype  $\alpha > 0$ , and that mutant homozygotes have phenotype  $\beta \geq \alpha$ .
- 2 This allows for various degrees of (partial) dominance, but not for over- or underdominance
- 3 (nor for complete recessiveness). Furthermore, we define  $\mu$  as the diploid (twice the haploid)
- 4 mutation rate (such that the number of new mutations per generation remains  $N\mu$ ). Since,
- 5 initially, the mutant allele occurs only in heterozygotes, the lag time and the waiting time
- 6 are identical to those in the haploid case (i.e, eq. 6 and 7 remain valid). The same is not true
- 7 for the fixation time, however, because the further spread of the mutant allele depends on
- 8 both the heterozygous and homozygous fitnesses. Nevertheless, there is a simple argument
- 9 that allows us to derive a rough approximation based on the haploid result.
- 10 The mutant allele can invade the population once the heterozygotes have higher fitness than
- 11 the wild-type, which is the case for  $t > t_1 = \alpha/(2v)$ . However, the wild-type allele can be

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

### 13 **The order of fixations in the multilocus case**

14 In the single-locus model, we have seen that a moving optimum favors fast fixation of mu-  
15 tations with an intermediate phenotypic effect, and that the size of the fastest mutation  
16 depends on the mutation rate and on the ratio of the speed of environmental change to the  
17 strength of selection. What does this result predict (qualitatively and quantitatively) for  
18 the order of fixed mutations in a full multilocus model, in which alleles with different effects  
19 “compete” against each other? Figure 4 shows simulation results for a model with 40 bial-  
20 lelic loci and uniformly distributed mutational effects  $\alpha_i$ . Details of two example runs are  
21 shown in Figure 5. We see that the characteristic U-shape from Figure 1 is indeed preserved  
22 (Fig. 4A-C). Mutant alleles with intermediate effects are fixed earlier than those with either  
23 small or large effects. Not surprisingly, the full pattern is seen only if the range of mutational  
24 effects is sufficiently large (Figure 4D-F): If only small (large) alleles are available, then only  
25 the left (right) branch of the U appears, leading to a monotonic and almost linear decrease  
26 (increase) of fixed effects over time.

27 Despite the qualitative agreement with the single-locus results (Fig. 1), there are two quan-  
28 titative differences that require explanation. First, the size of the fastest alleles is shifted to  
29 the left, that is, to effects less than  $\alpha^*$ . And second, the pattern assumes a Y-shape, rather

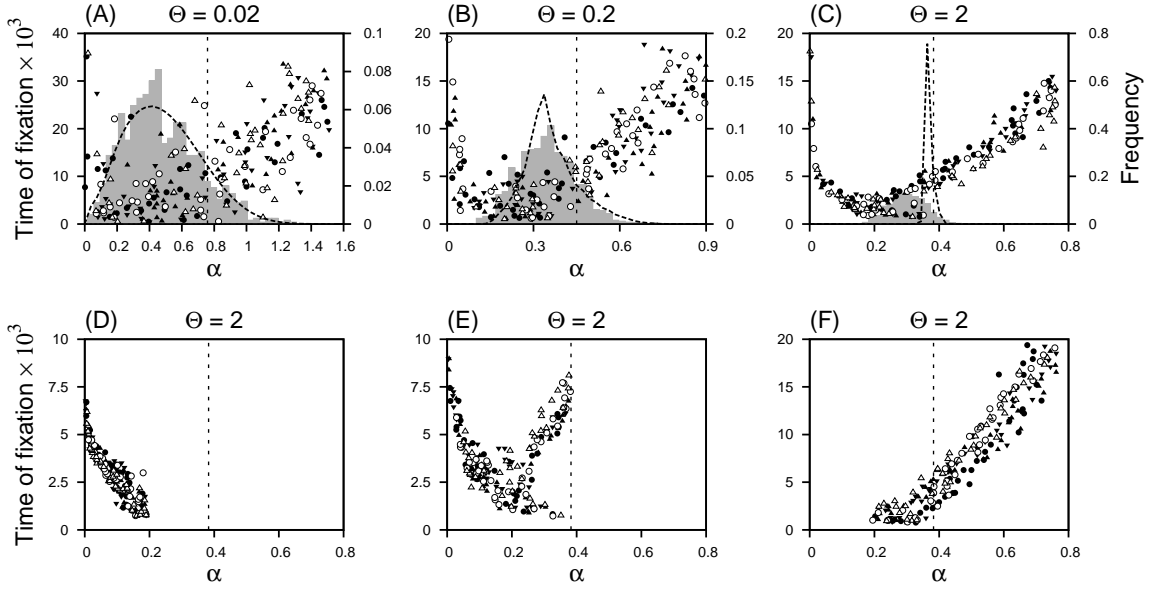


Figure 4: The order of fixations in the multilocus model. Symbols show the time to fixation of mutant alleles as a function of their effect  $\alpha$ . (Note that, as in Fig. 1, mutational effects are on the x-axis and time is on the y-axis.) Different symbols represent results from 5 replicated simulations with mutational effects drawn from a uniform distribution with a given range (between 0 and  $2\alpha^*$  in (A) to (C), between 0 and  $\alpha^*/2$  in (D), between 0 and  $\alpha^*$  in (E), and between  $\alpha^*/2$  and  $2\alpha^*$  in (F)). The vertical dotted line marks  $\alpha^*$ , the effect of the mutant allele with smallest expected time to fixation in the one-locus model (eq. A15). In (A) to (C), grey bars show the distribution of the effect of the first fixation as obtained from 1000 replicated simulations with fixed and evenly spaced  $\alpha$  values. The dashed line shows the predicted distribution of the first fixation according to equation (A21). Parameters:  $L = 40$ ,  $\mu = 10^{-5}$ ,  $\sigma = 0.1$ ,  $\theta = 2N\mu$ ,  $v = 0.001$ .

1 than a U-shape, if mutation rates are high (Fig. 4C). These differences can be explained by  
 2 two effects of multilocus genetics: A “sampling effect” for small  $\theta$  and an “interaction effect”  
 3 for large  $\theta$ .

4 **Sampling effect** The sampling effect arises from a combination of two factors: For small  
 5  $\theta$ , the waiting time has a large variance (eq. 7c), and small mutations have a shorter lag time  
 6 than large mutations. As a consequence, even though small mutations might have a high  
 7 expected total time to fixation, their minimal possible time is quite low. If the number of  
 8 loci is large, there is a high probability that one of the small mutations will fix before the lag  
 9 time of the larger mutations has passed. Thus, the sampling effect leads to an “advantage”  
 10 for small mutations, which is strongest if the waiting-time is the dominating time scale.

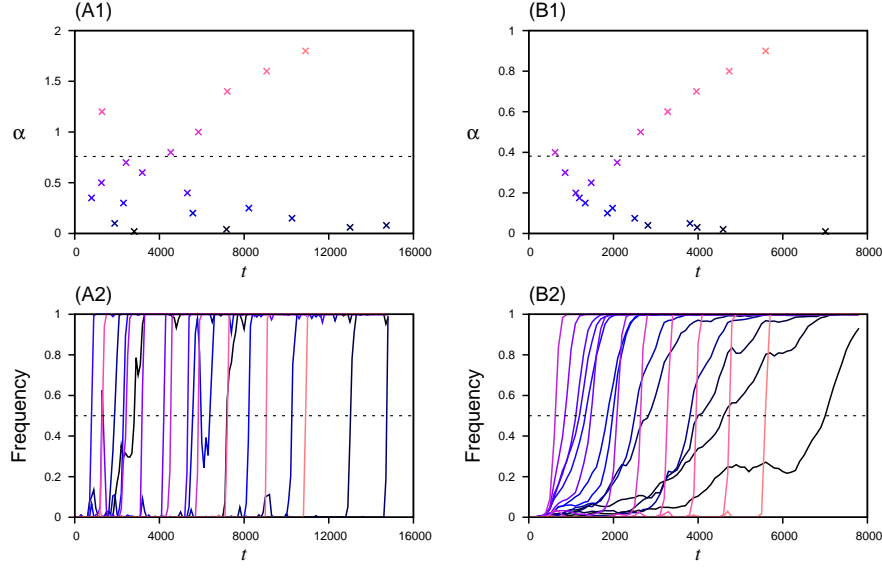


Figure 5: Adaptation in the multilocus model. (A1) is an example of a single simulation run with 20 loci ( $\alpha_i = 0.02, 0.04, 0.06, 0.08, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.5, 0.6, 0.7, 0.8, 1, 1.2, 1.4, 1.6, 1.8$ ) in a small population (parameters as in Fig. 4A). Each symbol shows the fixation time of the allele with the respective phenotypic effect. Note that, in contrast to Figure 4, time is on the x-axis and  $\alpha$  on the y-axis. (This is to allow for comparison with A2, see below). The dotted line marks  $\alpha^*$ , the effect of the mutant allele with the smallest expected time to fixation in the one-locus model (eq. A15). (A2) shows the corresponding allele frequency-dynamics. Colors indicate the allelic effects size  $\alpha_i$  (light red: small effect, dark blue: large effect). Note that each symbol in A1 corresponds to the line in A2 with the same color, and that 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  $\theta$  (parameters as in Fig. 4C) and  $\alpha_i$  values half the size of those in (A). Populations were sampled every 100 generations.

1 Since the sampling effect is a purely statistical phenomenon, it can still be estimated within  
 2 a single-locus framework. In Appendix 4, we use our results from the one-locus model to  
 3 derive an approximation for the distribution of the first fixation in the multi-locus model.  
 4 This approximation neglects epistatic interactions between mutant alleles at different loci  
 5 (see below). Nevertheless, it is in good agreement with simulation results, as long as  $\theta$  is  
 6 not too large (compare dashed lines and grey histograms in Figure 4A-C). The reason is  
 7 that, for small  $\theta$ , narrow sense fixation times are small, and different alleles rarely segregate  
 8 simultaneously (see Fig. 5A). Note also that the mode of the distribution of the first fixation  
 9 coincides well with the base of the U in Figure 4A-B.

1 **Interaction effect** If  $\theta$  is large, the probability of mutant alleles segregating simultane-  
2 ously is high (see Fig. 5B). These alleles interact with each other, because stabilizing selection  
3 entails epistasis for fitness: Each mutant allele that increases in frequency brings the popula-  
4 tion mean closer to the optimum and, thereby, decreases the selection coefficient of the other  
5 alleles. In particular, segregating small mutations can delay the fixation of large mutations  
6 by effectively prolonging their lag time.

7 The interaction effect has two main consequences: First, the size of the first fixation is smaller  
8 than predicted from the one-locus model, even if the sampling effect is negligible. This can  
9 be seen from Figure 4C, where the prediction for the first step according to equation (A21)  
10 (dashed line) is close to  $\alpha^*$ , but the observed distribution (grey bars) is considerably smaller.  
11 Second, the interaction effect causes a delay in the fixation of some intermediate and all  
12 large mutations, thus transforming the U-shaped into a Y-shaped pattern (Figure 4C, E).  
13 Figure 5B gives a more detailed view of this process. The initial decrease in effect sizes (the  
14 stem of the Y) arises because, at the beginning of the simulation, several alleles with small  
15 and intermediate effect start to increase in frequency more or less simultaneously. During  
16 their increase, these alleles suppress other mutant alleles, since their combined effect keeps  
17 the population mean close to the optimum. At about the “branching point”, the bulk of  
18 the initial alleles have become fixed, and only two classes of alleles remain: Those with very  
19 small effect, which have a very long fixation time and slowly increase in frequency “in the  
20 background”; and those with large effects and a long lag time, whose frequency up to this  
21 point was almost zero and who now, one by one, quickly go to fixation.

22 The interaction effect depends on the co-segregation of beneficial alleles at multiple loci, and is  
23 strongest if both small and large mutations are competing for fixation. This is demonstrated  
24 in Figure 4C-F. The Y-shaped pattern is seen only if the effects of available mutations cover  
25 a large range (4C, e). In contrast, if all mutations are large, they rarely co-segregate at  
26 higher frequencies. No delay in fixation is observed, and the size of fixed mutations increases  
27 almost linearly over time (Fig. 4F). Finally, if all mutations are small, epistatic interactions  
28 between them are even smaller. Co-segregating alleles hardly interfere, and the size of fixed  
29 mutations decreases linearly over time (Fig. 4D).

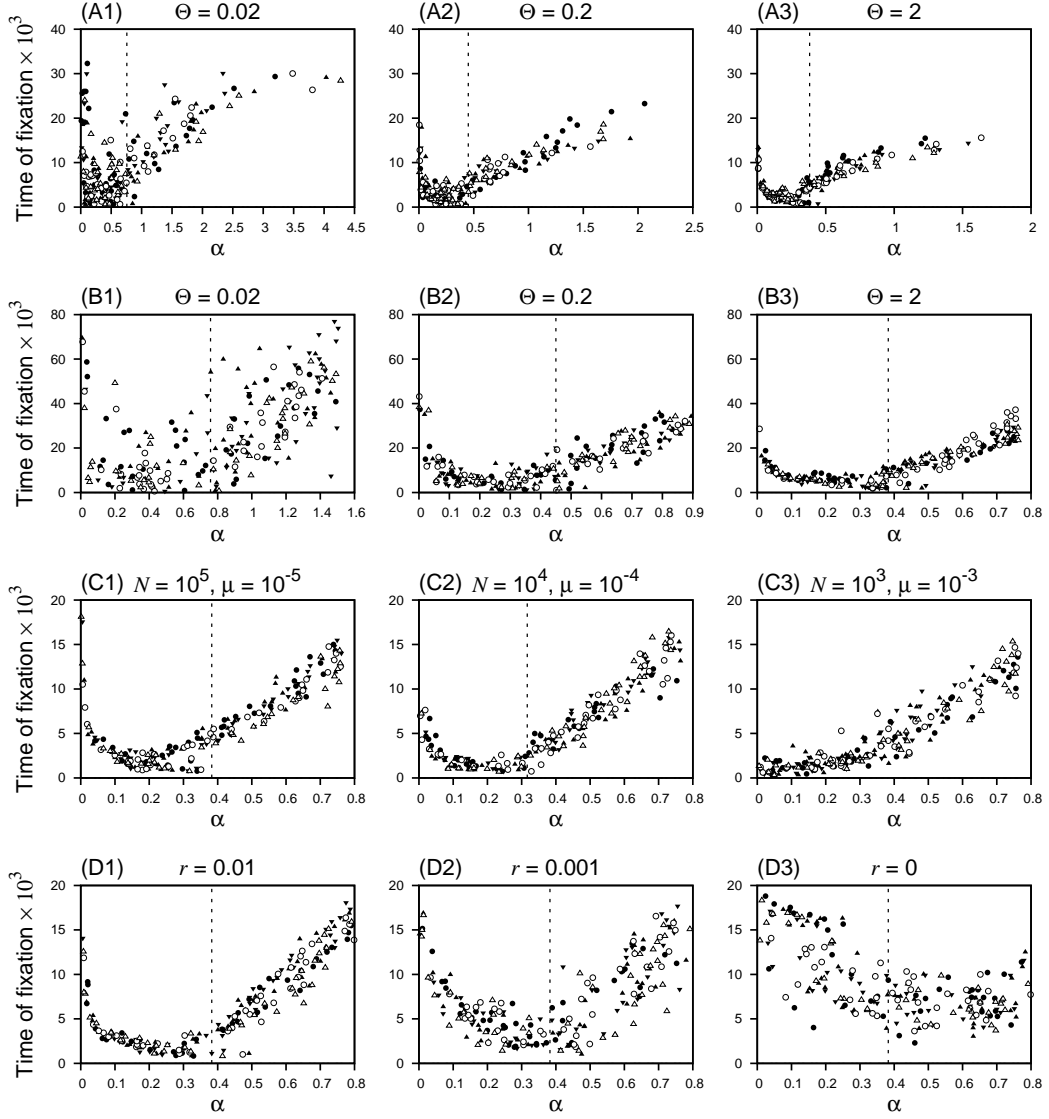


Figure 6: The order of fixations in variants of the multilocus model. (A) Mutational effects were drawn from an exponential distribution with mean  $\alpha^*$ . (B) Diploid case (no dominance). Here,  $\alpha$  is the effect of a single mutant allele. (C) Various combinations of  $N$  and  $\mu$  leading to the same value of  $\theta$ . (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  $\alpha$ . Different symbols represent results from 5 replicated simulations with randomly drawn mutational effects. The vertical dotted line marks  $\alpha^*$ , the effect of the mutant allele with smallest expected time to fixation in the one-locus model (eq. A15). In C3,  $\alpha^*$  could not be calculated because, for small  $\alpha$ , the approximation (14a) failed ( $\mu > s^*/2$ ). In B-D, mutational effects were drawn from a uniform distribution with mean  $\alpha^*$ . (Note that, as in Fig. 1 and 4, mutational effects are on the x-axis and time is on the y-axis.) See Fig. 4 for more details. Parameters:  $L = 40$ ,  $\sigma = 0.1$ ,  $v = 0.001$ ,  $\theta = 2N\mu = 2$  (C, B),  $\mu = 10^{-5}$  (A, B, D).

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

3 (i) *Distribution of locus effects*: Both theoretical and empirical evidence suggests that small  
4 beneficial mutations are more common than large ones (e.g., BÜRGER 2000; BARTON and  
5 KEIGHTLY 2002; ORR 2005a; EYRE-WALKER and KEIGHTLEY 2007). As seen in Figure 4,  
6 a limited range of mutational effects can have a large influence on the order of fixations.  
7 However, if the range is sufficiently large, the exact shape of the distribution from which  
8 mutational effects are drawn seems to matter little. For example, results qualitatively similar  
9 to those in Fig. 4 are obtained if the distribution of mutational effects is exponential instead  
10 of uniform (Fig. 6A).

11 (ii) *Diploidy*: As in the single-locus case, haploid and diploid genetics lead to qualitatively  
12 identical results (Fig. 6B).

13 (iii) *Variable mutation rate*: So far, we have assumed that the per-locus mutation rate  $\mu =$   
14  $10^{-5}$ , and we have varied the population size  $N$  to obtain various values of  $\theta = 2N\mu$ . Choosing  
15 a different value for  $\mu$  has little effect in most cases (Fig. 6C), showing that the crucial  
16 parameter is  $\theta$ , not  $N$  or  $\mu$  individually. Substantial differences occur only if  $\mu$  is extremely  
17 high ( $10^{-3}$  or higher in Fig. 6C). In this case, recurrent mutation may become more important  
18 than selection in determining the dynamics of allele frequencies ( $\mu > s^*/2$ , see remarks to  
19 eq. 14a and A7b). This leads to a “boost” for small mutations, which are under the weakest  
20 selection and have the shortest lag times. As a consequence, the U-shape disappears, and  
21 the size of fixed mutations increases monotonically over time (Fig. 6C3).

22 (iv) *Linkage*: The multilocus simulations presented above assume linkage equilibrium (see  
23 Appendix 1). However, additional simulations using individual-based modeling (Fig. 6D;  
24 Kopp and Hermisson, in prep.) show that the results are robust in the presence of moderate  
25 linkage (recombination rates  $\geq 0.01$ , Fig. 6D1). Strong linkage in combination with large  
26  $\theta$ , in contrast, leads to an increase in stochasticity and an increase in the size of the first  
27 fixation (Fig. 6D2, 3). Such a shift is also known from other models of adaptation and  
28 has been described in the context of clonal interference (e.g., GERRISH and LENSKI 1998;  
29 PARK and KRUG 2007) and the Hill-Robertson effect (HILL and ROBERTSON 1966; BARTON  
30 1995). Individual-based simulations also show that our results are not biased by the quadratic  
31 approximation (A2).

1 (v) *Standing genetic variation*: The simulations above were started with frequency  $p = 0$  for  
2 all mutant alleles at time  $t = 0$ , but genetic variation was allowed to accumulate during the lag  
3 time (see fixation from “already segregating alleles” above). In another set of simulations, we  
4 let the population equilibrate to the initial conditions before starting to move the optimum.  
5 This allows for the build-up of standing genetic variation (HERMISSON and PENNING 2005;  
6 BARRETT and SCHLUTER 2008). It turns out that the order of fixations is only slightly  
7 affected, with small loci benefiting more from standing variation than large ones (which are  
8 more deleterious initially). The reason for the small effect is that, for large  $\Theta$ , most fixations  
9 occur from mutations that arise during the lag time. However, this is already accounted for  
10 in our model (eq. 15), and adding genetic variation that is present even before the lag time  
11 does not change the qualitative picture.

12 (vi) *Speed of environmental change*: As in the one-locus case, we performed simulations with  
13  $v = 0.01$  and  $v = 0.0001$ , and obtained qualitatively identical results (not shown). For  $v >$   
14  $0.01$ , small populations sometimes are unable to follow the moving optimum and, instead, go  
15 to extinction (BÜRGER and LYNCH 1995). While the persistence or extinction of populations  
16 in the face of environmental change is an important topic with obvious implications for  
17 conservation, it is not the focus of the present study (but see LYNCH *et al.* 1991; LYNCH and  
18 LANDE 1993; BÜRGER and LYNCH 1995; NUNNEY 2003).

## 19 Discussion

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

26 In this paper, we have focused on a specific issue: Assume that mutations with various effects  
27 on the trait compete to be recruited in the adaptive process: What is the expected order of  
28 adaptive fixations? Which mutations – small, intermediate, or large – are fixed first under the

1 given environmental and genetic conditions? We have taken a two-step approach to address  
2 this question. In the first step, we consider a single mutant allele. We have derived a highly  
3 accurate approximation for how the expected total time to fixation of this allele depends on  
4 the internal (genetic) and external (environmental) factors. In the second step, we examine  
5 how these results are altered by multilocus effects in the context of a polygenic trait.

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

## 11 **Fixation of a single mutation**

12 Our analysis for the total time to fixation of a single allele is based on the insight that  
13 fixation in the moving optimum model is controlled by three different time scales (KOPP  
14 and HERMISSON 2007). The first time scale is ecological. It is given by the (inverse) speed  
15 of the trait optimum and governs the strength of selection. The other two time scales are  
16 genetic and measure the waiting time for a successful beneficial mutation, and the time for the  
17 mutant allele to rise to fixation. Accordingly, the total time to fixation can be subdivided  
18 into three components that correspond to these time scales: The *lag time* until the allele  
19 becomes beneficial as a result of environmental change, the *waiting time* for a successful  
20 mutation (which may be zero if fixation occurs from an already segregating allele), and the  
21 (narrow-sense) *fixation time*.

22 The three time scales partition the genetic and environmental parameter space into three  
23 evolutionary regimes (KOPP and HERMISSON 2007): In the *environmentally limited regime*,  
24 the total time to fixation of an allele is dominated by the lag time. In the *mutation-limited*  
25 *regime*, it is dominated by the waiting time, and in the *fixation time-limited regime* by  
26 the fixation time. The latter two regimes may be summarized as *genetically limited*. The  
27 principal distinction between them is the degree to which adaptation is rather stochastic  
28 (mutation-limited regime, population-wide mutation rate  $\theta \lesssim 1$ ) or rather deterministic

1 (fixation time-limited regime,  $\theta \gtrsim 1$ ). For our present issue, however, – that is, the order  
2 of adaptive substitutions – what matters most is the distinction between the two genetically  
3 limited regimes and the environmentally limited regime.

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

16 In consequence,  $\alpha^*$  provides a natural scale for measuring allelic effects (see Appendix 3).  
17 Its numerical value decides which effect sizes are “small” or “large” in the context of the  
18 adaptive process. Of course, this classification is a purely relative one. Whether a mutation  
19 of a given absolute effect will be counted as small or large depends on the value of  $\alpha^*$  and,  
20 hence, on the ecological and genetic parameters.

21 Indeed, the relationship between  $\alpha^*$  and the model parameters is very simple (eq. 17). To  
22 a good approximation,  $\alpha^*$  depends on the speed of environmental change relative to the  
23 strength of selection,  $v/\sigma$ , and on the population-wide mutation rate  $\theta$ . If  $\theta$  is small,  $\alpha^*$  only  
24 depends on a single composite parameter

$$\gamma = \frac{v}{\sigma\theta} \quad (18)$$

25 (see eq. 17b), which summarizes the genetic and ecological factors and should be amenable to  
26 empirical study. In nature,  $\gamma$  is expected to vary over several orders of magnitude. The speed  
27 of environmental change  $v$  experienced by a population depends on the ecological scenario  
28 and on the generation time, and it may vary from virtually imperceptible to extremely rapid.

1 Similarly, the effective population size (and hence  $\theta$ ) differs greatly among species. Thus, for  
2 a given ecological change,  $\alpha^*$  will usually be larger for microbes (with large populations and  
3 short generation times) than, for example, mammals. On the other hand, the dependence  
4 of  $\alpha^*$  on  $\gamma$  as a third root is rather weak (eq. 17). For example, a 10-fold increase in the  
5 speed of environmental change only leads to a 2.14-fold increase in  $\alpha^*$ . For the parameter  
6 values covered in Figure 2,  $\alpha^*$  is on the order of 0.1 or 1. If all phenotype related variables  
7 ( $\alpha$ ,  $\sigma$ ,  $v$ ) are measured in units of the environmental standard deviation, this is well within  
8 the range of mutational effects estimated from empirical data, which vary between  $10^{-2}$  and  
9 1 (BÜRGER 2000, p. 264). In addition, the distribution of mutational effects is thought to  
10 be highly leptocurtic, that is, mutations with large effects are overrepresented relative to a  
11 Gaussian distribution (BÜRGER 2000, p. 264). Therefore, it is likely that the majority of  
12 mutations will be smaller than  $\alpha^*$ , but a non-negligible part is bound to be larger. This  
13 means, in particular, that the fastest mutations are likely to be available in the population.

## 14 **The order of fixations at multiple loci**

15 In the single-locus model, we have argued that the fastest mutations are those at the bound-  
16 ary of the environmentally limited regime. A similar analysis can also be applied to the  
17 order of fixations in the multilocus model. However, to do so, we first need to define the  
18 notion of evolutionary regimes in a multilocus context. We will do so on a per-locus basis.  
19 Thus, in a multilocus model with unequal locus effects, one part of the loci may belong to  
20 one regime (e.g., environmentally limited), whereas others fall into a different regime (e.g.,  
21 mutation limited). Note that this definition deviates slightly from the one used in KOPP and  
22 HERMISSON (2007), where regimes were defined for the whole model, not for single loci.

23 Using only the time-scale arguments from the single-locus case, the results shown in Figure 4  
24 can be interpreted in the following way. (i) If all loci are in one of the genetically limited  
25 regimes, large effect loci adapt before small effect loci. This corresponds to the classical  
26 pattern known for adaptation under constant selection (e.g., ORR 1998, 2002; KIM and  
27 ORR 2005). (ii) If all loci are in the environmentally limited regime, loci with small effect  
28 are favored over ones with larger effect. This confirms the pattern described by BELLO and  
29 WAXMAN (2006), KOPP and HERMISSON (2007) and COLLINS *et al.* (2007). (iii) In a general

1 model with loci in both regimes, we see a branching pattern in the order of fixations, where  
2 loci with an intermediate effect fix first and mutations with either small or large effects follow  
3 up later. Thus, the linear patterns described under (i) and (ii) can be seen as special cases,  
4 which only arise if the trait has a genetic architecture with a restricted range.

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

9 (1) For small  $\theta$ , a *sampling effect* causes the average size of the first fixations to be smaller  
10 than  $\alpha^*$  (the fastest mutation in the one-locus model). The reason is that the waiting time for  
11 the first successful mutation in a “race” at multiple loci is shorter than the average waiting  
12 time at a single locus. This advantage is larger for loci with small effect, for which the waiting  
13 time is the dominating time scale. The sampling effect can still be captured analytically by  
14 using results from the one-locus model and ignoring interactions between loci (Appendix 4).

15 (2) For large mutation rates ( $\theta \gtrsim 1$ ), the fixation pattern is affected by interactions between  
16 co-segregating alleles. Figure 4 (B,C,D) shows that the split into two branches is again  
17 shifted to smaller locus effects and sets in only after an initial phase (Y-shape). The reason  
18 are epistatic interactions: For large  $\theta$ , mutations with small effect are dominated by the  
19 (narrow sense) fixation time. At any given time, therefore, there will be multiple segregating  
20 mutations that have not yet reached fixation. These mutations contribute to a narrowing of  
21 the gap between the mean phenotype and the optimum phenotype and reduce the selection  
22 pressure on mutations at other loci. In particular, they increase the lag time for mutations  
23 with larger effect, which leads to a shift in the pattern of early fixations towards smaller loci.

## 24 **Discussion of the modelling approach**

25 Previous modellers have used different approaches for studying the genetics of adaptation  
26 (ORR 2005a). Given the complex nature of the adaptive process, it is not surprising that dif-  
27 ferent models have focused on different aspects. The adaptive walk and mutational landscape  
28 models view evolution as a search process on a high-dimensional fitness surface (GILLESPIE  
29 1983; KAUFFMAN and LEVIN 1987; GILLESPIE 1984; KAUFFMAN 1993; ORR 2002). Here,

1 the key factor is epistasis, which makes the landscape rugged. Models of clonal interference  
2 focus on the effects of linkage and recombination on the fixation process (BARTON 1995;  
3 GERRISH and LENSKI 1998; KIM and ORR 2005; PARK and KRUG 2007). None of the above  
4 models contains an explicit phenotype. In contrast, Fisher’s geometric model explores the  
5 consequences of pleiotropy in a high-dimensional genotype-phenotype map (FISHER 1930;  
6 ORR 1998; WELCH and WAXMAN 2005; MARTIN and LENORMAND 2006). In our moving  
7 optimum model, finally, the key point is the inclusion of a changing environment, that is, of  
8 ecology.

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

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

23 Another way of comparing the various models is with respect to the included time-scales.  
24 In the models of Kauffman, Gillespie and Orr, only the mutation time-scale is relevant.  
25 KIM and ORR (2005) allow for an increased mutation rate, which introduces the fixation  
26 time as a second important scale. The interaction of the mutation and fixation time-scales  
27 also determines the outcome of clonal interference models. Our approach adds as a third  
28 component the ecological time-scale. As we have shown, the inclusion of this scale can  
29 change the predicted pattern of fixations substantially.

30 Models of constant selection predict that large mutations fix before small mutations (re-

1 regardless of whether mutational effects are measured in terms of phenotype or fitness; ORR  
2 2005a,b). Our model contains this result as a special case if all alleles are in the mutation-  
3 limited regime (i.e., if their effects are sufficiently small). However, in the general case, it  
4 is mutations of intermediate size that fix first. If all mutations are in the environmentally  
5 limited regime, we can even observe a complete reversal of the pattern, with small mutations  
6 fixing before large ones.

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

19 In this paper, we have focused on the order of adaptive substitutions. Another important  
20 question is the *distribution* of fixed mutations over a longer bout of adaptation (ORR 1998,  
21 2002, 2005a,b). However, while the present biallelic model can successfully predict the size  
22 of the first fixation (eq. A21), it is less suited for studying the distribution of subsequent  
23 adaptive steps. The reason is that we assume only a limited number of mutant alleles, all of  
24 which eventually go to fixation. As a consequence, these mutations are no longer available  
25 for future steps. For long-term adaptation, a more realistic approach is to use a continuum-  
26 of-alleles model (e.g., BÜRGER 1999), in which the effects of new mutations at each step are  
27 picked from a (fixed) distribution. The main question in such a model is how the (long-term)  
28 distribution of fixed substitutions is related to the distribution of new mutations. This will  
29 be the subject of a separate study (Kopp and Hermisson, manuscript in preparation).

30 Key concepts and findings of the present study include the role of time-scales and evolu-

1 tionary regimes, and the dependence of regime boundaries on the ecological parameters and  
2 the genetic architecture. They are not the results of specific model assumptions, but were  
3 developed and refined in a step-wise process from a single haploid locus to the full diploid  
4 multilocus case. We therefore expect that these findings also remain valid in a general context  
5 and can serve as guide for future model extensions.

6 **Acknowledgements** We thank N. Barton, R. Bürger, C. Rueffler, A. de Visser, and two  
7 anonymous reviewers for helpful comments on the manuscript. MK and JH are members of  
8 the Mathematics and BioSciences Group at the University of Vienna, which is funded by a  
9 grant from the Vienna Science and Technology Fund (WWTF) to JH.

## 10 **Appendices**

### 11 **Appendix 1: Simulation methods**

12 Here, we describe the computer simulations. We assume that the population has a constant  
13 size  $N$ , individuals are hermaphroditic and generations are non-overlapping. Simulations  
14 start at time  $t = 0$ , where the population is assumed to be monomorphic for the wild-type  
15 alleles at each locus. Each generation is modeled in two steps (see KOPP and HERMISSON  
16 2007). First, we use deterministic equations to calculate the *expected* genotype frequencies  
17 after selection, (free) recombination and mutation. Then, we use stochastic multinomial  
18 sampling to obtain the *actual* genotype frequencies in a finite population subject to genetic  
19 drift (e. g., GILLESPIE 1993). This means that the next generation is obtained by drawing  
20  $N$  individuals (with replacement) from the frequency distribution calculated in step one. (In  
21 the diploid case,  $2N$  haplotypes are drawn, assuming Hardy-Weinberg proportions before  
22 selection.)

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

1 quadratic function

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

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

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

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

## 8 **Appendix 2: Derivation of the fixation time $T_f$**

9 Here, we derive the narrow-sense fixation time  $T_f$ . For ease of notation, we define an al-  
10 ternative time scale  $\tilde{t}$ , which measures time since the start of the fixation process (so that  
11  $\tilde{t} = t - T_\ell - T_w$ ). The selection coefficient on this time scale is given by

$$s(\tilde{t}) = \lambda(T_w + \tilde{t}). \quad (\text{A3})$$

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

$$\dot{p} = s(\tilde{t})p(1 - p) + \mu(1 - 2p) + \frac{1}{2N}. \quad (\text{A4})$$

15 Here, the first term on the right-hand side describes logistic growth due to selection, the  
16 second term describes the input from mutation (both forward and backward), and the third  
17 term describes the contribution of genetic drift to the frequency increase of a mutation that  
18 is conditioned to go to fixation (EWENS 2004). Numerical solutions of this equation are

1 in excellent agreement with simulation results (not shown). Unfortunately, equation (A4)  
 2 does not have a closed analytical solution. To derive a tractable approximation, we neglect  
 3 selection as long as it is weaker than mutation and drift, and we neglect mutation and  
 4 drift when, together, they are weaker than selection. As a consequence, the fixation time is  
 5 subdivided into two components,

$$T_f = T_{f,1} + T_{f,2}, \quad (\text{A5})$$

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

9 **Fixation from a new mutation** For  $\tilde{t} < T_{f,1}$ , we neglect selection (and back-mutations),  
 10 and assume that  $p$  initially increases due to mutation and drift alone, yielding  $p(\tilde{t}) \approx \frac{\theta+1}{2N}\tilde{t}$ .  
 11 Mutation and drift remain dominant until  $s(\tilde{t})p(1-p) \approx s(\tilde{t})p > \frac{\theta+1}{2N}$ . Using (A3), this yields

$$T_{f,1}^n = \frac{1}{2} \left( \sqrt{T_w^2 + 4/\lambda} - T_w \right). \quad (\text{A6})$$

12 At this time, the frequency of the mutant allele is

$$p(T_{f,1}^n) = p^* = \frac{\theta + 1}{2Ns^*} \text{ with} \quad (\text{A7a})$$

$$s^* = \lambda(T_w + T_{f,1}^n). \quad (\text{A7b})$$

13 This approximation is valid as long as  $p^*$  is small. At the very least, it must be smaller than  
 14 0.5, which is true if  $s^* > 2\mu + 1/N$ .

15 For  $\tilde{t} > T_{f,1}^n$ , we neglect drift and mutation, such that equation (A4) becomes  $\dot{p} = s(\tilde{t})p(1-p)$ .  
 16 With the initial condition  $p(T_{f,1}^n) = p^*$ , this can be solved to yield

$$p(\tilde{t}) = \frac{p^* \exp[S(\tilde{t})]}{1 + p^* (\exp[S(\tilde{t})] - 1)}, \text{ where} \quad (\text{A8a})$$

$$S(\tilde{t}) = \int_{T_{f,1}}^{\tilde{t}} s(x) dx = s^* \Delta \tilde{t} + \frac{\lambda}{2} (\Delta \tilde{t})^2 \text{ and} \quad (\text{A8b})$$

$$\Delta \tilde{t} = \tilde{t} - T_{f,1}^n, \quad (\text{A8c})$$

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

$$T_{f,2}^n = \frac{\sqrt{(s^*)^2 + 2\lambda \ln\left(\frac{2Ns^*}{\theta+1} - 1\right)} - s^*}{\lambda}. \quad (\text{A9})$$

2 Equation 8a follows from (A7b) and (A5).

3 **Fixation from an already segregating allele** If fixation occurs from an already segre-  
 4 gating allele,  $T_w = 0$  and thus  $\tilde{t} = t - T_\ell$  and  $s = \lambda \tilde{t}$ . The expected frequency of the mutant  
 5 allele at the end of the lag time,  $p(0)$ , can be calculated as follows: During the lag time (i. e.,  
 6 for  $\tilde{t} < 0$ ),  $p$  is small, and equation (A4) can be approximated by  $\dot{p} = \mu + \lambda \tilde{t} p$ . (Note that,  
 7 here, the drift term  $1/(2N)$  is omitted, because we are not conditioning on eventual fixation.)  
 8 This can be solved to yield

$$p(\tilde{t}) = \exp\left(\frac{1}{2}\lambda \tilde{t}^2\right) \left( C + \frac{\mu}{\sqrt{\lambda}} \int_0^{\sqrt{\lambda/2}\tilde{t}} \exp(-x^2) dx \right). \quad (\text{A10})$$

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

$$p(0) = C = \frac{\mu}{2} \sqrt{\frac{\pi}{\lambda}}. \quad (\text{A11})$$

10 For  $\tilde{t} > 0$ , we follow the same approach as for fixation from a new mutation. For  $\tilde{t} < T_{f,1}^s$ ,  $p$   
 11 increases linearly according to  $p(\tilde{t}) = p(0) + \mu \tilde{t}$  until  $p = p^*$  (eq. A7a), leading to

$$T_{f,1}^s = \frac{1}{\sqrt{\lambda}} \left( \sqrt{\frac{\pi}{16} + 1} - \sqrt{\frac{\pi}{16}} \right) \approx \frac{0.65}{\sqrt{\lambda}}. \quad (\text{A12})$$

12  $T_{f,2}^s$  is again given by (A9), but with  $s^* = \lambda T_{f,1}^s$ . Finally, the probability of fixation from an  
 13 already segregating allele can be estimated as

$$\mathcal{P}^s = 1 - \exp\left(-p(0)N\sqrt{\frac{8\lambda}{\pi}}\right) \quad (\text{A13})$$

1 (JH, unpublished), where  $p(0)$  is the expected frequency of the mutant allele at the end of  
 2 the lag time. Inserting (A11) into (A13) directly leads to (15).

### 3 **Appendix 3: Analytical approximation for $\alpha^*$**

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

$$\mathbb{T} \approx \frac{\alpha}{2v} + \frac{C(\sigma, \theta, v, N)}{\sqrt{\alpha}} \quad (\text{A14a})$$

where the constant  $C$  is given by

$$C(\sigma, \theta, v, N) = (1 - \mathcal{P}^s)C_1(\sigma, \theta, v, N) + \mathcal{P}^s C_2(\sigma, \theta, v, N) \quad (\text{A14b})$$

with

$$C_1(\sigma, \theta, v, N) = \sqrt{\frac{\pi \left(\sqrt{1 + \frac{8\theta}{\pi}} + 1\right)^2}{16\theta\sigma v} + \frac{\ln\left(\frac{2N\sqrt{2\sigma v}\left(\sqrt{\frac{\pi}{8\theta} + 1} + \sqrt{\frac{\pi}{8\theta}}\right) - 1}{\theta + 1}\right)}{\sigma v}}, \quad (\text{A14c})$$

$$C_2(\sigma, \theta, v, N) = \sqrt{\frac{\pi \left(\sqrt{1 + \frac{16}{\pi}} - 1\right)^2}{32\sigma v} + \frac{\ln\left(\frac{2N\sqrt{2\sigma v}\left(\sqrt{\frac{\pi}{16} + 1} - \sqrt{\frac{\pi}{16}}\right) - 1}{\theta + 1}\right)}{\sigma v}}. \quad (\text{A14d})$$

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

$$\alpha^* = (vC)^{2/3} \quad (\text{A15})$$

10 This approximation for  $\alpha^*$  is very accurate for small  $\theta$ , but shows small deviations from the  
 11 true value if  $\theta$  is large (see Fig. 2). Furthermore, plugging (A15) into (A14a) shows that, for  
 12 a mutation of size  $\alpha^*$ , the total time to fixation,  $\mathbb{T}^*$ , is given by

$$\mathbb{T}^* = \frac{(vC)^{2/3}}{2v} + \frac{C}{(vC)^{1/3}} = \frac{1}{2} \left( \frac{C^2}{v} \right)^{1/3} + \left( \frac{C^2}{v} \right)^{1/3}, \quad (\text{A16})$$

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

4 **Scaling relationships** We can use the above results to derive several useful scaling rela-  
 5 tionships. First, plugging (A14b) to (A14d) into (A14a) and ignoring the logarithmic term,  
 6 shows that  $\alpha^*$  is proportional to  $\sqrt[3]{v/\sigma}$ . For small  $\theta$ ,  $\mathcal{P}^S \approx 0$ , and eq. (A15) reduces to  
 7  $\alpha^* = (vC_1)^{2/3}$ , which, if we again ignore the logarithmic term, is proportional to  $\sqrt[3]{v/(\sigma\theta)}$ .

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

## 19 **Appendix 4: The distribution of the first fixation in the multilocus** 20 **model**

21 Here, we derive the predicted distribution of the first fixation in the multilocus model. As-  
 22 sume that there are  $L$  loci with two alleles each. What is the probability that the mutant  
 23 allele at locus  $i$  is the first to reach fixation? In the following, we present an approximation  
 24 that is an extension of the two-locus approach by KOPP and HERMISSON (2007). The key

1 simplification is that allele frequency changes at different loci are assumed to be independent  
 2 (neglecting epistatic interactions for fitness).

3 We first need to combine the cases of fixation from a new mutation and from an already  
 4 segregating allele. By a slight abuse of notation, we will use  $T_w = 0$  to indicate fixation from  
 5 a segregating allele. We also introduce additional indices for the focal locus. Using (4b),  
 6 (7a), and (15), the waiting time distribution (7a) at locus  $i$  then becomes

$$g_i(T_{w,i}) = \begin{cases} 0 & \text{for } T_{w,i} < 0 \\ 1 - \exp\left(-\frac{\theta}{\sqrt{2}}\right) & \text{for } T_{w,i} = 0 \\ \theta\lambda T_{w,i} \exp\left(-\frac{\theta}{\sqrt{2}} - \frac{\theta\lambda}{2}T_{w,i}^2\right) & \text{for } T_{w,i} > 0 \end{cases} \quad (\text{A17})$$

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

$$G_i(T_{w,i}) = \int_{\tau=0}^{T_{w,i}} \tilde{f}_i(\tau) d\tau = \begin{cases} 0 & \text{for } T_{w,i} < 0 \\ 1 - \exp\left(-\frac{\theta}{\sqrt{2}}\right) & \text{for } T_{w,i} = 0 \\ 1 - \exp\left(-\frac{\theta}{\sqrt{2}} - \frac{\theta\lambda_i}{2}T_w^2\right) & \text{for } T_{w,i} > 0 \end{cases} \quad (\text{A18})$$

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

$$\mathbb{T}_i(T_{w,i}) = \frac{\alpha_i}{2\nu} + \frac{1}{\lambda_i} \sqrt{s_i^*(T_{w,i})^2 + 2\lambda_i \ln\left(\frac{2Ns_i^*(T_{w,i})}{\theta + 1} - 1\right)} \quad \text{where} \quad (\text{A19})$$

$$s_i^*(T_{w,i}) = \begin{cases} \sqrt{\lambda_i} \left(\sqrt{\frac{\pi}{16} + 1} - \sqrt{\frac{\pi}{16}}\right) & \text{for } T_{w,i} = 0, \\ \frac{\lambda_i}{2} \left(T_w + \sqrt{T_w^2 + \frac{4}{\lambda_i}}\right) & \text{for } T_{w,i} > 0. \end{cases} \quad (\text{A20})$$

10 Assume the mutant allele at locus  $i$  has waiting time  $T_{w,i}$ . Then a mutant allele at another  
 11 locus  $j$  will fix first only if its waiting time is less than  $T_{w,j}^*(T_{w,i})$ , which can be obtained by  
 12 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  
 13 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  
 14 one at locus  $i$ , even if it starts from an already segregating allele. Thus,  $T_{w,j}^*$  can be set to a  
 15 negative value (which has zero probability). If  $\mathbb{T}_j(0) \leq \mathbb{T}_i(T_{w,i}) < \lim_{T_{w,j} \rightarrow 0^+} \mathbb{T}_j(T_{w,j})$ , then

1 the allele at locus  $j$  needs to start from a segregating allele and thus,  $T_{w,j}^* = 0$ . Finally, if  
 2  $\lim_{T_{w,j} \rightarrow 0^+} \mathbb{T}_j(T_{w,j}) \leq \mathbb{T}_i(T_{w,i})$  then the allele at locus  $j$  can start from a new mutation, and  
 3  $T_{w,j}^*$  is the solution to the above equation.) Finally, the probability  $\pi_i$  that the mutant allele  
 4 at locus  $i$  fixes before all other mutant alleles is given by

$$\pi_i = \int_0^\infty g_i(T_{w,i}) \prod_{j \neq i} (1 - G_j(T_{w,j}^*(T_{w,i}))) dT_{w,i}. \quad (\text{A21})$$

## 5 References

- 6 BARRETT, R. D. H., and D. SCHLUTER, 2008 Adaptation from standing genetic variation.  
 7 Trends Ecol. Evol. **23**: 38–44.
- 8 BARTON, N. H., 1995 Linkage and the limits to natural selection. Genetics **140**: 821–841.
- 9 BARTON, N. H., and D. KEIGHTLY, 2002 Understanding quantitative genetic variation.  
 10 Nat. Rev. Gen. **3**: 11–21.
- 11 BELLO, Y., and D. WAXMAN, 2006 Near-periodic substitutions and the genetic variance  
 12 induced by environmental change. J. Theor. Biol. **239**: 152–160.
- 13 BÜRGER, R., 1999 Evolution of genetic variability and the advantage of sex and recombina-  
 14 tion in a changing environment. Genetics **153**: 1055–1069.
- 15 BÜRGER, R., 2000 *The mathematical theory of selection, recombination, and mutation*. Wi-  
 16 ley, Chichester.
- 17 BÜRGER, R., 2005 A multilocus analysis of intraspecific competition and stabilizing selection  
 18 on a quantitative trait. J. Math. Biol. **50**: 355–396.
- 19 BÜRGER, R., and A. GIMELFARB, 2002 Fluctuating environments and the role of mutation  
 20 in maintaining quantitative genetic variation. Genet. Res. **80**: 31–46.
- 21 BÜRGER, R., and M. LYNCH, 1995 Evolution and extinction in a changing environment: a  
 22 quantitative genetic analysis. Evolution **49**: 151–163.

- 1 COLLINS, S., J. DE MEAUX, and C. ACQUISTI, 2007 Adaptive walks toward a moving  
2 optimum. *Genetics* **176**: 1089–1099.
- 3 ELENA, S. F., and R. E. LENSKI, 2003 Evolution experiments with microorganisms: the  
4 dynamics and genetic bases of adaptation. *Nature Reviews Genetics* **4**: 457–469.
- 5 ETHERIDGE, A., P. PFAFFELHUBER, and A. WAKOLBINGER, 2006 An approximate sam-  
6 pling formula under genetic hitchhiking. *Ann. Appl. Probab.* **16**: 685–729.
- 7 EWENS, W. J., 2004 *Mathematical population genetics*. Springer-Verlag, Berlin, 2nd edition.
- 8 EYRE-WALKER, A., and P. D. KEIGHTLEY, 2007 The distribution of fitness effects of new  
9 mutations. *Nature Reviews Genetics* **8**: 610–618.
- 10 FISHER, R., 1930 *The genetical theory of natural selection*. Clarendon Press, Oxford.
- 11 GERRISH, P. J., and R. E. LENSKI, 1998 The fate of competing beneficial mutations in an  
12 asexual population. *Genetica* **102/103**: 127–144.
- 13 GILLESPIE, J. H., 1983 A simple stochastic gene substitution model. *Theor. Pop. Biol.* **23**:  
14 202–215.
- 15 GILLESPIE, J. H., 1984 Molecular evolution over the mutational landscape. *Evolution* **38**:  
16 1116–1129.
- 17 GILLESPIE, J. H., 1993 Substitution processes in molecular evolution. I. Uniform and clus-  
18 tered substitutions in a haploid model. *Genetics* **134**: 971–981.
- 19 HAIRSTON, N. G., S. P. ELLNER, M. A. GEBER, T. YOSHIDA, and J. A. FOX, 2005  
20 Rapid evolution and the convergence of ecological and evolutionary time. *Ecology Letters*  
21 **8**: 1114–1127.
- 22 HERMISSON, J., and P. S. PENNING, 2005 Soft sweeps: molecular population genetics of  
23 adaptation from standing genetic variation. *Genetics* **169**: 2335–2352.
- 24 HILL, W. G., and A. ROBERTSON, 1966 The effect of linkage on limits to artificial selection.  
25 *Gen. Res.* **8**: 269–294.

- 1 KAUFFMAN, S. A. A., 1993 *Origins of order: self-organization and selection in evolution*.  
2 Oxford University Press, Oxford.
- 3 KAUFFMAN, S. A. A., and S. LEVIN, 1987 Towards a general theory of adaptive walks on  
4 rugged landscapes. *J. Theor. Biol.* **128**: 11–45.
- 5 KIM, Y., and H. A. ORR, 2005 Adaptation in sexuals vs. asexuals: clonal interference and  
6 the Fisher-Muller model. *Genetics* **171**: 1377–1386.
- 7 KIMURA, M., 1983 *The neutral theory of molecular evolution*. Cambridge University Press,  
8 Cambridge.
- 9 KOPP, M., and J. HERMISSON, 2007 Adaptation of a quantitative trait to a moving optimum.  
10 *Genetics* **176**: 715–718.
- 11 LYNCH, M., W. GABRIEL, and A. M. WOOD, 1991 Adaptive and demographic responses of  
12 plankton populations to environmental change. *Limnology and Oceanography* **36**: 1301–  
13 1312.
- 14 LYNCH, M., and R. LANDE, 1993 Evolution and extinction in response to environmental  
15 change. In P. Kareiva, J. G. Kingsolver and R. B. Juey, editors, *Biotic interactions and*  
16 *global change*. Sinauer, Sunderland, 43–53.
- 17 MARTIN, G., and T. LENORMAND, 2006 A general multivariate extension of Fisher’s geo-  
18 metrical model and the distribution of mutation fitness effects across species. *Evolution*  
19 **60**: 893–907.
- 20 NUNNEY, L., 2003 The cost of natural selection revisited. *Ann. Zool. Fennici* **40**: 185–194.
- 21 ORR, H. A., 1998 The population genetics of adaptation: the distribution of factors fixed  
22 during adaptive evolution. *Evolution* **52**: 935–949.
- 23 ORR, H. A., 2002 The population genetics of adaptation: the adaptation of DNA sequences.  
24 *Evolution* **56**: 1317–1330.
- 25 ORR, H. A., 2005a The genetic theory of adaptation: a brief history. *Nat. Rev. Gen.* **6**:  
26 119–127.

- 1 ORR, H. A., 2005b Theories of adaptation: what they do and don't say. *Genetica* **123**:  
2 3–13.
- 3 PARK, S.-C., and J. KRUG, 2007 Clonal interference in large populations. *Proc. Natl. Acad.*  
4 *Sci. USA* **104**: 18135–18140.
- 5 PARMESAN, C., 2006 Ecological and evolutionary responses to recent climate change. *Annu.*  
6 *Rev. Ecol. Evol. Syst.* **37**: 637–669.
- 7 PERRON, G. G., A. GONZALEZ, and A. BUCKLING, 2008 The rate of environmental change  
8 drives adaptation to an antibiotic sink. *J. Evol. Biol.* **21**: 1724–1731.
- 9 PROVINE, W. B., 2001 *The origins of theoretical population genetics*. Chicago University  
10 Press, Chicago.
- 11 SATO, M., and D. WAXMAN, 2008 Adaptation to slow environmental change, with apparent  
12 anticipation of selection. *J. Theor. Biol.* **252**: 166–172.
- 13 THOMPSON, J. N., 2005 *The geographic mosaic of coevolution*. The University of Chicago  
14 Press, Chicago.
- 15 WAXMAN, D., and J. R. PECK, 1999 Sex and adaptation in a changing environment. *Ge-*  
16 *netics* **153**: 1041–1053.
- 17 WELCH, J. J., and D. WAXMAN, 2005 Fisher's microscope and Haldane's ellipse. *Am. Nat.*  
18 **166**: 447–457.