

The Distribution of Beneficial and Fixed Mutation Fitness Effects Close to an Optimum

Guillaume Martin¹ and Thomas Lenormand

Centre d'Ecologie Fonctionnelle et Evolutive, UMR CNRS 5175, 34295 Montpellier, France

Manuscript received January 18, 2008
Accepted for publication March 18, 2008

ABSTRACT

The distribution of the selection coefficients of beneficial mutations is pivotal to the study of the adaptive process, both at the organismal level (theories of adaptation) and at the gene level (molecular evolution). A now famous result of extreme value theory states that this distribution is an exponential, at least when considering a well-adapted wild type. However, this prediction could be inaccurate under selection for an optimum (because fitness effect distributions have a finite right tail in this case). In this article, we derive the distribution of beneficial mutation effects under a general model of stabilizing selection, with arbitrary selective and mutational covariance between a finite set of traits. We assume a well-adapted wild type, thus taking advantage of the robustness of tail behaviors, as in extreme value theory. We show that, under these general conditions, both beneficial mutation effects and fixed effects (mutations escaping drift loss) are beta distributed. In both cases, the parameters have explicit biological meaning and are empirically measurable; their variation through time can also be predicted. We retrieve the classic exponential distribution as a subcase of the beta when there are a moderate to large number of weakly correlated traits under selection. In this case too, we provide an explicit biological interpretation of the parameters of the distribution. We show by simulations that these conclusions are fairly robust to a lower adaptation of the wild type and discuss the relevance of our findings in the context of adaptation theories and experimental evolution.

UNDERSTANDING the distribution of fitness effects of beneficial mutation [hereafter $f_b(s_b)$] is necessary to predict the rate and genetic basis of adaptation (ORR 1998). It is also important to calibrate models of molecular evolution where positive selection is involved (EYRE-WALKER 2006) or to study processes involving the segregation of several beneficial mutations, like clonal interference (GERRISH and LENSKI 1998). So far, this distribution has been studied along two directions (for a historical review see ORR 2005a). The first is based on FISHER'S (1930) geometric model of adaptation, while the second uses GILLESPIE'S (1984) mutational landscape model. These two models differ in their basic assumptions, and each has its own limitations (discussed in ORR 2005b). Fisher's model (FM) considers stabilizing selection around an optimum in an n -dimensional phenotypic space and focuses on the fitness effect of random phenotypic changes. The mutational landscape model (MLM) directly focuses on the effect of single-nucleotide substitutions on fitness. The strength of the FM is to predict the full distribution of fitness effects of mutations [hereafter $f(s)$], including both deleterious and beneficial mutations and their respective proportions, which depends on the level of

adaptation of the wild type (distance to the optimum). The MLM is less general in that it considers only beneficial mutation, but its strength is to avoid explicit assumptions on the phenotype-to-fitness map inherent to the FM. This is made possible when beneficial mutations can be considered drawn from the extreme right tail of $f(s)$. In this case indeed, extreme value theory can be used to predict the (unique) limiting distribution of extreme draws, *i.e.*, $f_b(s_b)$. Importantly, this is robust to a wide range of $f(s)$ (GILLESPIE 1984; ORR 2002). A now famous and remarkably simple prediction of this theory is that $f_b(s_b)$ should be exponential (ORR 2003). This finding has, since then, been widely used (GERRISH and LENSKI 1998; WILKE 2004; PARK and KRUG 2007). Note that this is not the same as the distribution of effects fixed over a bout of adaptation, which is also predicted to be exponential (ORR 1998). In this article, we determine $f_b(s_b)$ under a general model of stabilizing selection, based on an extension of the FM (MARTIN and LENORMAND 2006b), but we study our model in the same biological conditions as assumed in the MLM, thus allowing the use of extreme value theory in this context. We show that under this general model, the exponential approximation for $f_b(s_b)$ can be substantially inaccurate unless there are a large number of weakly correlated traits under selection. We provide an alternative, in terms of a beta distribution, that includes the exponential as a limiting case. Before presenting

¹Corresponding author: Centre d'Ecologie Fonctionnelle et Evolutive, UMR CNRS 5175, 1919 Rte. de Mende, 34295 Montpellier, France.
E-mail: guillaume.martin@cefe.cnrs.fr

these results, we first discuss the limit of the MLM and classic FM approaches to predict $f_b(s_b)$.

The MLM approach has provided simple, robust, and testable conclusions, but, as any model, it has limits. First, the MLM is by construction valid only when the wild type is well adapted to its environment, so that mutations with fitness above the wild type's are indeed drawn from the rightmost tail of $f(s)$ (GILLESPIE 1984); this may not be the case in a new environment. Second, the MLM makes explicit assumptions on the genetic basis of adaptation (single-nucleotide substitutions with equal probability of occurrence). Although it is often seen as more realistic, these genetic assumptions may also limit the scope of the theory. Indeed, even in a simple situation involving only point mutations, corrections were required to compare empirical data to MLM theory because of differences in transition *vs.* transversion rates (ROKYTA *et al.* 2005). Third, even when considering well-adapted wild types, the MLM is not robust to *any* $f(s)$. Only the so-called Gumbel types of distributions, characterized by an exponential-like tail (ORR 2002; BEISEL *et al.* 2007), are consistent with the historical models by Gillespie and Orr. There are in fact three possible "domains of attraction" determining extreme values behavior: the Gumbel type discussed above, the Fréchet type, for heavily tailed distributions, and the Weibull type, for distributions that have a rightmost endpoint. There is no obvious reason to prefer one type over the others (discussed in BEISEL *et al.* 2007). Fourth, the MLM does not provide any prediction on how the distribution of mutant fitnesses should change over several generations, as the population adapts to its environment. In the classic formulation, the effect of adaptation is only to shift the wild type to higher and higher fitness ranks in an otherwise constant distribution of mutant fitnesses. By contrast, the FM is free of these limits but at the cost of explicit assumptions on the genotype-to-phenotype-to-fitness map, which could be unrealistic.

Overall, from an empirical point of view, it has proved difficult to validate predictions on beneficial mutation effects so far, because they are often rare. The exponential distribution appears to give a reasonable but still imperfect fit to empirical distributions of beneficial effects (ROKYTA *et al.* 2005; KASSEN and BATAILLON 2006;). From a more statistical point of view, no alternative theoretical $f_b(s_b)$ had been proposed until recently, when a statistical framework was proposed to test alternative predictions, all stemming from extreme value theory (BEISEL *et al.* 2007). Overall, empirical studies so far could neither clearly accept nor reject the predictions of the MLM, so that theoretical arguments may help settle the issue. In particular, it would be important that FM and MLM approaches yield consistent results. We now turn to this question.

In a recent article, ORR (2006) sought to bridge the gap between the FM and the MLM and showed that they

share strong similarities regarding $f_b(s_b)$. Indeed, under the classic Fisher model, the distribution of fitness effects of single mutations is close to Gaussian, which pertains to the domain of attraction of the Gumbel-type extreme value distribution assumed in the MLM. This property ensures that the derivations of the MLM are at least approximately valid under the biological assumptions of the FM, which reinforces the view that $f_b(s_b)$ should indeed be exponential. However, this conclusion should be taken with caution. First, under the FM, $f(s)$ is necessarily bounded on the right: there is no better mutation than the one bringing the phenotype at the optimum. As we have seen, distributions that are bounded on the right pertain to the Weibull domain of attraction, not to the Gumbel, and this will be true of any model of selection for an optimum. ORR (2006) mentioned this problem and showed that the exponential approximation could nevertheless be accurate under the FM, provided that there are a large number of equivalent and independent traits affected by pleiotropic mutation and selection. However, (i) selective and mutational independence of the traits is often considered biologically unrealistic (ORR 2005b), and (ii) the number of traits affected by mutation may be limited, at least when considering single genes, as is usual in molecular evolution. Overall, the assumption of a large number of independent traits leads to an approximately Gaussian $f(s)$ (simply by the central limit theorem) whereas reviews of empirical $f(s)$ show that they are better approximated by a skewed gamma distribution when only deleterious mutations are observed (MARTIN and LENORMAND 2006b; EYRE-WALKER and KEIGHTLEY 2007). Because it is this Gaussian $f(s)$ that leads to an approximately exponential $f_b(s_b)$ in the FM (ORR 2006), the exponential result may not be robust if traits are fewer and correlated. Indeed, recent models have shown that nonequivalence between traits in the FM could substantially affect the predictions of the FM (WAXMAN and WELCH 2005; MARTIN and LENORMAND 2006b).

In this article, we apply extreme value theory to a general model of selection for an arbitrary optimum, where mutations have pleiotropic effects on an arbitrary number of potentially inequivalent and correlated traits. In this case, $f_b(s_b)$ belongs to the Weibull, rather than the Gumbel domain of attraction. Using a recent approximation for the tail behavior of quadratic forms in Gaussian vectors (JASCHKE *et al.* 2004), we show that beneficial effects are approximately beta distributed provided the wild type is relatively well adapted (as assumed in the MLM). This result is based on tail approximations (similar to the extreme value theory approach) and is robust to any continuous phenotype-to-fitness function close to an optimum (contrary to the classic FM). Our conclusions are checked using exact simulations, which show that the tail approximation yields surprisingly accurate results, even away from the tail. We discuss our results and compare them with results from the MLM.

MODEL

$f(s)$ under arbitrary stabilizing selection: We consider an extension of the FM that has been detailed previously (MARTIN and LENORMAND 2006b). The fitness $W(\mathbf{z})$ of a phenotype \mathbf{z} (of arbitrary dimension n , the number of phenotypic traits under selection) is a multivariate Gaussian function of \mathbf{z} , $W(\mathbf{z}) \equiv \exp(-\frac{1}{2}\mathbf{z}^T \cdot \mathbf{S} \cdot \mathbf{z})$, where superscript T denotes transposition, and \mathbf{S} is an arbitrary positive semidefinite matrix of selective interactions between phenotypic traits. This assumption is justified when close to the optimum, as many continuous fitness functions around a single optimum can be approximated by a Gaussian function close to that optimum (LANDE 1979). This does not preclude the existence of other optima, but it does assume that they are too remote from the mutant “cloud” around the wild type to influence $f(s)$. We consider an initial genotype (or wild type), with phenotype \mathbf{z}_o [and fitness $W(\mathbf{z}_o) = W_o$]. The distribution of mutant phenotypes ($d\mathbf{z}$) around \mathbf{z}_o is assumed to be multivariate Gaussian with mean $\mathbf{0}$ and arbitrary (positive semidefinite) covariance matrix \mathbf{M} . Again, this assumption is not as restrictive as it seems; what is required in fact is that there exists a set of trait definitions for which their mutational effect distribution is Gaussian (MARTIN and LENORMAND 2006b), which requires only that the distribution of mutant effects on the original traits be continuous, unimodal, and approximately centered on the wild type.

This model is a quite general description of stabilizing selection, when not too far from the optimum, and with universal pleiotropy of mutations (mutations affect all traits simultaneously). Contrary to the classic (isotropic) Fisher model, it allows for differences and correlations between traits for both mutation and selection. Following assumptions of the MLM, we assume that the wild type is well adapted (close to the optimum), so that beneficial mutation effects are small. Consequently, the selection coefficient of a beneficial mutant relative to the wild type ($s = W/W_o - 1$) is approximately equal to the log-relative fitness: $s \approx \log(1 + s) = \log(W/W_o)$. Therefore, s is approximately a quadratic function of mutational phenotypic effects $d\mathbf{z}$ (MARTIN and LENORMAND 2006b), *i.e.*, a quadratic form in Gaussian vectors (MATHAI and PROVOST 1992).

Beyond their mathematical convenience or robustness, these assumptions are supported by data: the model seems to correctly account for the variation of empirical distributions of mutational fitness effects across species (MARTIN and LENORMAND 2006b), across environments (MARTIN and LENORMAND 2006a), and among mutations (fitness epistasis; MARTIN *et al.* 2007). Under these assumptions, the probability density function (pdf) of s , $f(s)$, is entirely determined by the n eigenvalues of the matrix product $\mathbf{S}\mathbf{M}$ and the position \mathbf{z}_o of the initial phenotype relative to the optimum (APPENDIX A; Equation A2 in MARTIN and LENORMAND

2006b). There is no analytic expression for $f(s)$ in the general case, but it can be approximated by a displaced gamma distribution (JASCHKE *et al.* 2004; MARTIN and LENORMAND 2006b), as illustrated in APPENDIX A.

Importantly, the distribution of s is bounded on its rightmost endpoint by $s_o = \log(W_{\max}/W_o)$ that is the selection coefficient of the individual with optimal phenotype [with fitness $W(\mathbf{0}) = W_{\max}$] relative to the wild type [with fitness $W(\mathbf{z}_o) = W_o$]. As we saw above, this kind of right-bounded distribution is inherent to any model of selection for an optimum.

Distribution of beneficial fitness effects close to the optimum: A tail approximation for the distribution of quadratic forms in Gaussian vectors (such as s) has been derived recently (JASCHKE *et al.* 2004). When the wild type is well adapted (as $s_o \rightarrow 0$), a simple approximation can be deduced from this tail approximation, for the distribution $f_b(s_b)$ of beneficial mutations ($0 < s_b < s_o$), yielding

$$\frac{s_b}{s_o} \sim \text{beta}\left(1, \frac{m}{2}\right), \quad \text{when } s_o \ll 1 \quad (1)$$

(APPENDIX B), where s_o is the fitness distance to the optimum defined above, and $m = \text{rank}(\mathbf{S}\mathbf{M})$. The distribution of s_b scaled to its range $[0, s_o]$ is a beta distribution $\text{beta}[1, m/2]$. The two parameters of the distribution have an explicit biological interpretation and are *a priori* biologically independent of each other: s_o describes the level of adaptation of the wild type, and m describes the dimensionality of the phenotype-fitness landscape. Here “dimensionality” is understood as the number of *distinct* traits that are effectively affected by both mutation (matrix \mathbf{M}) and selection (matrix \mathbf{S}), *i.e.*, with correlation less than one. When there is no or little correlation between traits, both \mathbf{S} and \mathbf{M} are positive definite (with only strictly positive eigenvalues) so that $m = n$, the total number of traits that are pleiotropically affected by mutation. However, with many traits, as correlations between traits increase, either \mathbf{M} or \mathbf{S} or both quickly become semidefinite; *i.e.*, their rank $m < n$. This means that there are $n - m$ traits that can in fact be expressed as linear combinations of the m distinct traits of the system. Because $m = \text{rank}(\mathbf{S}\mathbf{M}) \leq \min(\text{rank}(\mathbf{M}), \text{rank}(\mathbf{S}))$, correlations among a large number of traits will result in a relatively small m , unless these correlations are weak.

Domain of attraction: The beta distribution given in Equation 1 is an example of the so-called generalized Pareto distribution (GPD) that encompasses the three possible domains of attraction of extreme value theory (PICKANDS 1975). To use the classic formulation (used, *e.g.*, in BEISEL *et al.* 2007), this beta distribution in Equation 1 is a GPD with location $\mu = 0$, scale $\tau = 2/m$, and shape $\kappa = -2/m$. As long as m is not infinitely large, κ is negative so that $f(s)$ falls into the Weibull domain of attraction. However, with an infinitely large m , κ would

be zero, and $f(s)$ would fall into the Gumbel domain of attraction. This explains why the classic FM, which assumes a very large number of independent traits, is consistent with the MLM (ORR 2006) and close to a Gumbel-type distribution. Consistent with this, the cumulative distribution function of the beta distribution in Equation 1 converges to that of an exponential distribution when m is large:

$$\frac{s_b}{s_o} \sim \text{exponential} \frac{m}{2}, \quad \text{when } s_o \ll 1 \text{ and } m \gg 1 \quad (2)$$

(see APPENDIX B). Importantly, this result also provides an explicit formulation for the rate of the exponential in terms of biological parameters: with large m , beneficial effects (s_b) are exponentially distributed with rate $m/2s_o$.

Overall, under our general model of selection for an optimum, and with a well-adapted wild type, we obtain an approximately exponential distribution of beneficial mutations only when there are sufficiently many *independent* and *weakly correlated* traits under selection. When a limited number of traits are affected by the mutational target under consideration (*e.g.*, a single gene), or when there are many but strongly correlated traits, there is no reason to expect an exponential $f_b(s_b)$. In these cases, one should use the full model (beta approximation) given in Equation 1. Because the exponential approximation is a limiting case of the beta approximation, the two behaviors may be easily compared statistically. We now turn to the study of the fitness effect distribution of those beneficial mutations that reach fixation.

The distribution of fitness effects among beneficial mutations escaping drift loss: Not all beneficial mutations will fix in a population: even in an infinitely large population, most are lost soon after their appearance, due to the stochasticity of offspring number. From $f_b(s_b)$, it is possible to derive the distribution of selection coefficients among those beneficial mutations that escape drift loss when they are still rare (*i.e.*, those that reach fixation in a sexual population). From Equation 1, assuming a well-adapted wild type and a population not too small, we can use a weak selection approximation ($\pi(s) \propto 2s$) for the fixation probability of beneficial mutations (HALDANE 1927; WHITLOCK 2000) and find another beta approximation for the distribution of fixed mutation effects s_f :

$$\frac{s_f}{s_o} \sim \text{beta} \left(2, \frac{m}{2} \right), \quad \text{when } s_o \ll 1 \quad (3)$$

(APPENDIX C). The average fitness effect of *fixed* beneficial mutations is then (for small s_o)

$$E(s_f) \approx \frac{4s_o}{4+m}, \quad (4)$$

which, as expected, is larger than the average fitness effect of all beneficial mutations (from Equation 1),

$$E(s_b) \approx \frac{2s_o}{2+m}. \quad (5)$$

As expected, these two values increase with the wild-type maladaptation s_o and decrease with increased dimensionality m , which is a part of the “cost of complexity” defined by ORR (2000). The other part of this cost, not dealt with here, is the reduction of the fraction of beneficial mutations as m increases. When m is large, $E(s_f)$ and $E(s_b)$ are $\sim 4s_o/m$ (fixed effects) and $2s_o/m$ (beneficial effects), respectively, which converges to the results derived from the exponential approximation (see APPENDIX C).

Simulations: To check the accuracy of the above results, we simulated distributions of beneficial mutations in the Fisher model with correlated traits as in MARTIN and LENORMAND (2006b). We drew the mutational and selective covariance matrices (\mathbf{M} and \mathbf{S}) from $n \times n$ Wishart distributions $W_p(n, \mathbf{I})$ (where \mathbf{I} is the $n \times n$ identity matrix) so that the rank of $\mathbf{S}\mathbf{M}$ is $m = \min(p, n)$. \mathbf{M} and \mathbf{S} were then scaled to obtain an average deleterious effect of mutations $\bar{s} = \text{tr}(\mathbf{S}\mathbf{M})/2 = 0.05$, where $\text{tr}(\cdot)$ denotes matrix trace. Then the phenotype of the wild-type \mathbf{z}_o was drawn as a Gaussian vector and scaled so that $\log(W_{\max}/W_o) = \log(1/W_o) = s_o$, for a given fitness distance to the optimum. Finally, for each single mutant, we drew a mutation effect vector \mathbf{dz} from a multivariate Gaussian distribution $N(\mathbf{0}, \mathbf{M})$ and computed s as $s = \log(W(\mathbf{z}_o + \mathbf{dz})/W(\mathbf{z}_o))$. The resulting distributions of s are illustrated in supplemental Figure 1.

We chose p to get a distribution of fitness effects (among all mutations) with a large skewness, as is typically observed in empirical studies (*e.g.*, SANJUÁN *et al.* 2004). More precisely, a given distribution of deleterious s (when $s_o = 0$) corresponds to a given effective number of traits n_e (depending on the magnitude of correlations in \mathbf{M} and \mathbf{S}) that determines the shape of $f(s)$ (MARTIN and LENORMAND 2006b). With \mathbf{M} and \mathbf{S} drawn as Wishart deviates $\mathbf{S}, \mathbf{M} \sim W_p(n, \mathbf{I})$, $n_e \approx n/(1 + 2n/p)$ (for details, see Appendix 2 in MARTIN and LENORMAND 2006b), so we chose p as the integer part of $2n \cdot n_e/(n - n_e)$ to obtain a given n_e and n . Figures 1–3 show the same two examples corresponding to alternative levels of pleiotropy. In the low pleiotropy case, $n = 4$ and $n_e = 2.5$, so that $m = n = 4$ (\mathbf{S} and \mathbf{M} are positive definite). In the high pleiotropy case (still keeping a small n_e), $n = 40$ and $n_e = 4$, so that $m = p = 9$ (\mathbf{S} and \mathbf{M} are positive semidefinite). Therefore, these two cases correspond to a lower (resp. higher) number of traits jointly affected by mutation and selection and to a lower (resp. higher) dimensionality m . They are denoted low (resp. high) pleiotropy in the figures.

From a set of 400,000 simulated single mutants we kept only those with fitness higher than the wild type as beneficial mutants. To compute the fitness effect distribution among mutants that escape drift loss, we computed the exact fixation probability P_{fix} of each of the n_b

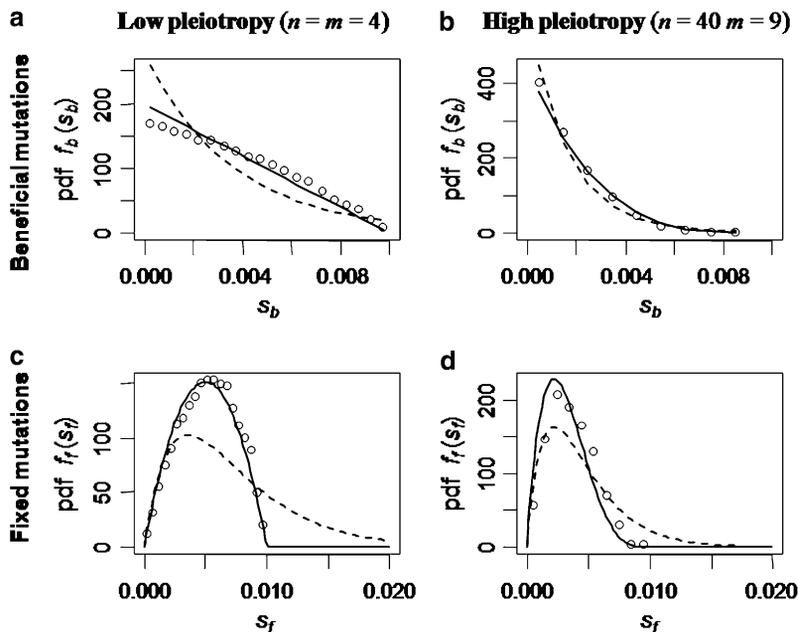


FIGURE 1.—Distribution of beneficial (s_b) and fixed (s_f) mutation fitness effects. The fitness effects distributions of beneficial (s_b) (a and b) and fixed (s_f) (c and d) mutations in exact simulations (from 400,000 random mutations, circles) are compared to alternative approximations: the beta (solid lines, Equations 1 and 3 for s_b and s_f , resp.) and the exponential (dashed lines, best-fitting exponential for s_b and Equation C4 for s_f). (a and c) A low pleiotropy case: $n = 4$ weakly correlated traits ($m = 4$). (b and d) A higher pleiotropy case: $n = 40$ traits that are more strongly correlated ($m = 9$). The wild type was well adapted [fitness distance to the optimum: $s_0 = 0.01$, with an average (deleterious) effect of all mutations: $-\bar{s} = -0.05$]. The beta approximation leads to a more accurate prediction than the exponential in both cases: the increase in accuracy is almost undetectable for high pleiotropy (b and d), but substantial for low pleiotropy (a and c).

beneficial mutants (with selection coefficient s) by numerically solving $P_{\text{fix}} = e^{-(1+s)P_{\text{fix}}}$, according to HALDANE (1927). Then we sampled n_b times the beneficial mutants according to their individual fixation probability P_{fix} .

RESULTS

Accuracy of the beta approximation for beneficial and fixed effects s_b : The beta approximation given in Equation 1 gives a very good fit to the simulations when s_0 is small relative to the average of all mutations ($s_0 \ll \bar{s}$), so that beneficial mutations are rare and at the rightmost tail of $f(s)$. Figure 1, a and b, illustrates $f_b(s_b)$ in this situation and supplemental Figure 1 shows the corresponding $f(s)$. However, the prediction is still fairly accurate for larger values of s_0 (and larger proportions of beneficial mutations), as illustrated in supplemental Figure 2, for $s_0 = \bar{s}$. As expected (Equation 2), when compared to the beta, even the best-fitting exponential distribution provides a less accurate description of $f_b(s_b)$ when m is small (Figure 1a), but a similarly satisfying one, even with a moderately large m ($m = 9$, Figure 1b, the two approximations are almost indistinguishable). Nevertheless, even in the latter case, a closer investigation (Figure 2) shows that the exponential approximation inaccurately captures the distribution on its rightmost part (for the largest beneficial effects), while the beta approximation (Equation 1) is accurate on the whole range of beneficial effects. This has little influence on the accuracy of the exponential model for beneficial effects (with large m), but is more problematic when deriving the distribution of *fixed* effects. Indeed, as for beneficial effects, the beta approximation for the distribution of fixed effects (Equation 3, Figure

1, c and d) gives a good fit to individual simulations (compare solid lines and open circles in Figure 1, c and d). As a comparison, the distribution of s_f under the (best-fitting) exponential approximation for s_b (Equation C4, APPENDIX C) gives a less good fit to simulations (dashed lines, Figure 1, c and d), worse when the dimensionality is low (Figure 1c, low pleiotropy case). The exponential approximation gives less accurate results for fixed effects (Figure 1, c and d) than for beneficial effects (Figure 1, a and b) because the exponential inaccurately describes the distribution of large beneficial s (Figure 2) that are overrepresented among fixed mutations.

Robustness of the results away from the optimum: As in the MLM, the results presented here are all weak selection approximations in that they assume that the wild type is close to the optimum (small s_0), so that beneficial mutations are all of small effect. We checked the robustness of these results when s_0 gets larger: supplemental Figure 2 shows that while the beta approximation for beneficial effects (s_b , Equation 1) is less (but still reasonably) accurate when $s_0 = \bar{s}$, the beta approximation for fixed effects (s_f in Equation 3) remains fairly accurate in this case. More surprisingly, the average value of fixed effects [$E(s_f)$, Figure 3] remains close to the tail approximation result (Equation 4), even for fairly large values of s_0 (up to 10 times the average effect of all mutations: $s_0 = 0.5 = 10\bar{s}$). As expected again, the prediction from the exponential approximation is less accurate, at least with a small m for beneficial effects, and in both cases for fixed effects. It becomes less and less accurate as s_0 increases (constant difference on log scale, Figure 3). The same pattern holds for $E(s_b)$ (not shown). Overall, while the shape of the distributions, away from the tail, is less accurately

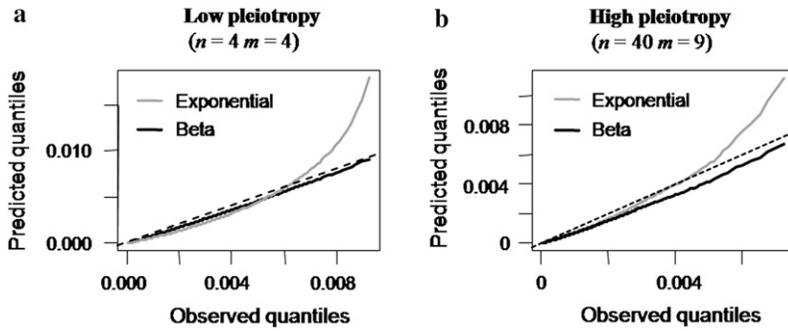


FIGURE 2.—Quantile–quantile plots for the beta *vs.* exponential approximations of $f_b(s_b)$. This is based on the same cases as in Figure 1, a and b. The quantiles of $f_b(s_b)$ from simulations (*x*-axis) are compared to predicted quantiles (*y*-axis) from either the beta approximation (Equation 1, solid line) or the best-fitting exponential (shaded line). When the points lie on a line with slope 1 (dashed line), the observed and expected distributions are the same. The beta approximation is accurate on the whole range of fitness effects while the exponential is inaccurate in the rightmost part of the distribution (largest *s* values).

described by Equations 1 and 3 (see supplemental Figure 2), their means are still fairly robustly predicted for large s_o .

DISCUSSION

In this article, we derive the distribution of beneficial mutation fitness effects under a general model of selection for a phenotypic optimum with an arbitrary set of traits undergoing selection and pleiotropic mutation. The main restriction of the model is that the wild type in which mutations appear is assumed to be well adapted to the environment considered, an assumption shared with the mutational landscape approach.

Comparing the beta and exponential distributions for beneficial effects: Under these fairly general conditions, and although the exact system depends on many parameters, the distribution of beneficial effects, $f_b(s_b)$, is accurately approximated by a simple beta distribution (Equation 1). This distribution is an exam-

ple of the generalized Pareto distribution of the Weibull type, not of the Gumbel type as is classically assumed in most of the literature on adaptation theory. However, in the limit of a large number of weakly correlated traits (increased dimensionality m), our beta approximation converges to the exponential, consistent with previous results (ORR 2006). Overall, the distribution of beneficial effects (s_b , Equation 1) will substantially differ from an exponential when only a limited number of traits are considered (Figure 1a, supplemental Figure 2). Indeed, the convergence to an exponential is quick as dimensionality increases (Figure 1b, $m = 9$). However, this convergence to the exponential is slower for the distribution of fixed beneficial effects s_f (Figure 1, c and d). This occurs because, even for large m , the exponential distribution tends to particularly overestimate the proportion of largely beneficial effects [the really extreme right tail of $f(s)$] compared to the beta (Figure 2), and these effects are strongly overrepresented among fixed mutations.

Therefore, when assuming selection for an optimum, one should use the beta approximations proposed here (Equations 1–4), whenever possible, as they provide better accuracy in the general case, while retaining the simplicity that made the exponential approximation theoretically attractive. However, when considering beneficial-effect distributions (and with caution for fixed effects), and provided the dimensionality m is even moderately large (an issue we discuss more fully below), the exponential can provide an even simpler and similarly accurate approximation.

An important aspect of our result is that, beyond classic results from extreme value theory, our model provides a biological interpretation of the two parameters that emerge in the tail approximation: s_o is the selection coefficient of the optimal genotype relative to the wild type; m measures the level of pleiotropy (*i.e.*, the number of not fully dependent dimensions of the phenotypic space under selection). Note that when $n > m$, there are $n - m$ traits that are completely determined by linear combinations of the first m traits: m is therefore akin to a “degree of freedom,” the number of traits that suffice to fully describe the fitness landscape. Although included in the model for the sake of generality, the

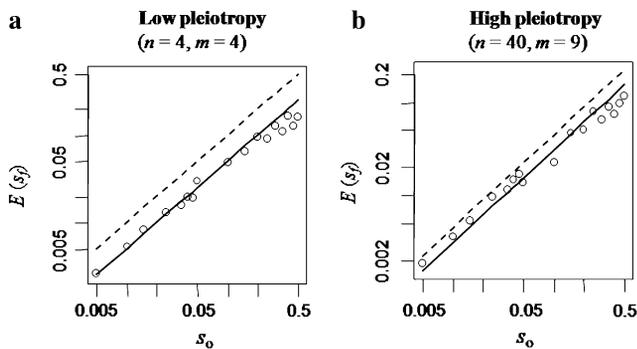


FIGURE 3.—Variation of $E(s_f)$ with the distance to the optimum s_o . The average selection coefficient of fixed mutations, $E(s_f)$, is given for the same parameter values as in Figures 1 and 2, except that the distance to the optimum (*x*-axis) varies from $s_o = 0.005 = 0.1\bar{s}$, to $s_o = 0.5 = 10\bar{s}$. Open circles give $E(s_f)$ from simulations, while lines give the prediction from either the beta approximation (Equation 4, solid lines), or the best-fitting exponential (Equation C5, dashed lines). a and b show the same two levels of pleiotropy as in Figures 1 and 2. Note the log–log scale. Equation 4 is accurate even for surprisingly large values of s_o . In contrast, results from the exponential approximation are moderately accurate in the high pleiotropy case (b) and inaccurate in the low pleiotropy case (a).

extra $n - m$ traits are somehow meaningless in terms of pleiotropy. These two parameters (m and s_0) are, *a priori*, biologically independent. For instance, because s_0 measures adaptation of the wild type, we may predict how this parameter changes through time as individuals adapt, while m could be expected to remain constant, at least over short evolutionary timescales. Beyond characterizing the distribution of beneficial effects, this model therefore provides a means to predict how this distribution changes through time, as in the FM, while preserving the robustness provided by the use of tail behaviors (extreme value theory) as in the MLM. This is true, including when the exponential approximation is valid (large m , Figure 1b): our model and simulations then show that beneficial effects s_b are exponentially distributed, with rate $m/2s_0$.

Robustness of the results: There are few other assumptions in the model, apart from the existence of an optimum, to which the wild type is well adapted. Indeed, the model is approximately valid near a local maximum of any continuous fitness function and for any selective or mutational covariance between traits (by construction). Another relevant issue is modularity: our model assumes total pleiotropy of all mutations on all the traits considered. If distinct mutational targets (*e.g.*, genes) affect at least partly distinct sets of traits, then there is modularity in the effect of mutation (WELCH and WAXMAN 2003). We suspect that the total $f(s)$ would then be a sum of each module's $f(s)$, weighted by the probability of mutation in each module. The effect of such modularity would have to be studied in more detail, but even then there would still be a maximum value of s , so that $f(s)$ would likely pertain to the Weibull domain of attraction, leading to a distribution of beneficial effects of the type of Equation 1. However, the biological interpretation of the two parameters is probably less straightforward in this case.

A surprising property of our model is the robustness of the predictions when away from the tail. The approximate distribution of both beneficial and fixed effects is still reasonably accurate when they are of the same order as the mean effect of all mutations ($s_0 = \bar{s}$, supplemental Figure 2), for which the proportion of beneficial mutations is $>10\%$. Even more surprisingly, the mean of these distributions (both beneficial and fixed effects) is accurately predicted, even farther away from the tail (up to $s_0 = 10\bar{s}$, Figure 3). Overall, the average log-fitness gain per adaptive fixation [$E(s_f)$, Equation 4] is $\sim 4s_0/(4 + m)$, for fairly arbitrary levels of adaptation of the wild type. However, the prediction would probably fail, away from the tail, if the phenotype-to-fitness function $W(\mathbf{z})$ was not close enough to a Gaussian, which is possible when away from the optimum. Whether or not fitness functions are Gaussian remains an open question, although a review of mutation effects in stressful environments (wild-type ill-adapted) did suggest that it may be a reasonable approximation even away from the

optimum (MARTIN and LENORMAND 2006a). Overall, while the simple results derived here may apply also in new and stressful environments (away from the optimum), they are *a priori* more likely to be valid in benign ones (close to it).

How large is m ? An important issue is whether m is large or not, as it determines the accuracy of the exponential approximation. When m is at least moderately large ($m \geq 10$), then the classic exponential (Equation 2) could be a sufficient approximation, when describing beneficial effects (although it would be less accurate for fixed effects, as we already mentioned). Although a large m seems intuitively likely because many traits are under selection, it needs not be so: as we have seen above, (i) with even weak mutational and selective correlations, a large set of traits is necessarily mutually dependent, which reduces m , and (ii) traits may be organized in modules. That empirical $f(s)$ are not Gaussian suggests that these effects are important. In fact, whether m is large enough that the exponential is a sufficient description of $f_b(s_b)$ is mainly an open empirical question; we now turn to this issue.

Empirical estimation of m : One may estimate m empirically, with a similar approach as proposed to estimate n_e (MARTIN and LENORMAND 2006b): from empirical distributions of single-mutant fitnesses [empirical $f(s)$]. Indeed, at any distance from the optimum, m may be estimated by fitting a generalized Pareto distribution to the extreme right tail of mutation fitness effects. The shape of the GPD, $\kappa = -2/m$, will then provide an estimate for m . Such a fit can be performed using the method of BEISEL *et al.* (2007) or routines proposed in the "POT" R package (<http://r-forge.r-project.org/projects/pot/>), for example. Alternatively, m may be measured from evolution experiments data (TENAILLON *et al.* 2007). However, some simulations would be needed to check whether this latter method does estimate the required quantity (*i.e.*, m not n) when traits are correlated.

Predicting the proportion of beneficial mutations? The tail approximation used in this article (JASCHKE *et al.* 2004) can also be used to derive the proportion p_b of beneficial mutations when s_0 is small (Equation B3 of APPENDIX B). Unfortunately, this prediction is much less robust than that on $f_b(s_b)$, giving strongly inaccurate results unless p_b is of the order of $\leq 10^{-3}$ (simulations not shown). Consistent with this, JASCHKE *et al.* (2004) showed that their tail approximation gave a poor fit to quadratic forms distributions, unless considering values very close to the rightmost endpoint. Because our prediction for $f_b(s_b)$ depends on the same approximation, modulo the scaling constant d (see APPENDIX B), we believe the poor robustness of the approximation for $f(s)$ and p_b comes from the expression for d being valid only very close to the rightmost endpoint of the distribution. Anyhow, this lack of fit means that our results do not provide a satisfactory expression for p_b , and that the displaced gamma approximation should be preferred

for this purpose (MARTIN and LENORMAND 2006b), although it yields a slightly more complicated expression.

Extreme value theory and clonal interference: Clonal interference is the mechanism by which beneficial mutations occurring in different individuals compete for ultimate fixation in asexuals. GERRISH and LENSKI (1998; GERRISH 2001) showed that, in the limit of fairly low mutation rates, this process implies that the mutations that fix are the ones with the largest selection coefficient among all those that appear during a selective sweep. Such a sieving process therefore consists in drawing the maximum value among a set of draws from a distribution, which is exactly what is described by extreme value theory or tail approximations. Therefore, the MLM and the model discussed in this article may prove useful for describing the distribution of fixed mutations in asexuals, in a much more general context than for sexuals, *i.e.*, at any distance from the optimum. As extreme s values (largely beneficial) are strongly over-represented among mutations that escape clonal interference, it will probably be safer to use the general model (beta approximations) than the exponential limit in this context, as the latter is less accurate when it comes to describing the very right tail of $f(s)$ (Figure 2).

Conclusion: Overall, our study clarifies the conditions under which the different ways to model the distribution of fitness effects of beneficial mutations give similar or different results and why. In particular, we stress that beneficial mutations may not be exponentially distributed. Under selection for an optimum, the fitness effects of both beneficial and fixed mutations are Beta distributed, which is close to exponential only when there are a large number of weakly correlated traits subject to selection and pleiotropic mutation.

We thank Allen Orr, Paul Joyce, and two anonymous reviewers for their helpful comments on a previous version of this manuscript. G.M. was funded by a Centre National de la Recherche Scientifique postdoctoral grant to Sylvain Gandon, and T.L. was supported by a starting grant from the European Research Council.

LITERATURE CITED

- BEISEL, C. J., D. R. ROKYTA, H. A. WICHMAN and P. JOYCE, 2007 Testing the extreme value domain of attraction for distributions of beneficial fitness effects. *Genetics* **176**: 2441–2449.
- EYRE-WALKER, A., 2006 The genomic rate of adaptive evolution. *Trends Ecol. Evol.* **21**: 569–575.
- EYRE-WALKER, A., and P. D. KEIGHTLEY, 2007 The distribution of fitness effects of new mutations. *Nat. Rev. Genet.* **8**: 610–618.
- FISHER, R. A., 1930 *The Genetical Theory of Natural Selection*. Oxford University Press, Oxford.
- GERRISH, P., 2001 The rhythm of microbial adaptation. *Nature* **413**: 299–302.
- GERRISH, P. J., and R. E. LENSKI, 1998 The fate of competing beneficial mutations in an asexual population. *Genetica* **103**: 127–144.
- GILLESPIE, J. H., 1984 Molecular evolution over the mutational landscape. *Evolution* **38**: 1116–1129.
- HALDANE, J. B. S., 1927 A mathematical theory of natural and artificial selection V. Selection and mutation. *Proc. Camb. Philos. Soc.* **26**: 220–230.
- JASCHKE, S., C. KLUPPELBERG and A. LINDNER, 2004 Asymptotic behavior of tails and quantiles of quadratic forms of Gaussian vectors. *J. Multivariate Anal.* **88**: 252–273.
- KASSEN, R., and T. BATAILLON, 2006 The distribution of fitness effects among beneficial mutations prior to selection in experimental populations of bacteria. *Nat. Genet.* **38**: 484–488.
- LANDE, R., 1979 Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. *Evolution* **33**: 402–416.
- MARTIN, G., and T. LENORMAND, 2006a The fitness effect of mutations in stressful environments: a survey in the light of fitness landscape models. *Evolution* **12**: 2413–2427.
- MARTIN, G., and T. LENORMAND, 2006b A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. *Evolution* **60**: 893–907.
- MARTIN, G., S. F. ELENA and T. LENORMAND, 2007 Distributions of epistasis in microbes fit predictions from a fitness landscape model. *Nat. Genet.* **39**: 555–560.
- MATHAI, A. M., and S. B. PROVOST, 1992 *Quadratic Forms in Random Variables*. Marcel Dekker, New York.
- ORR, H. A., 1998 The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution* **52**: 935–949.
- ORR, H. A., 2000 Adaptation and the cost of complexity. *Evolution* **54**: 13–20.
- ORR, H. A., 2002 The population genetics of adaptation: the adaptation of DNA sequences. *Evolution* **56**: 1317–1330.
- ORR, H. A., 2003 The distribution of fitness effects among beneficial mutations. *Genetics* **163**: 1519–1526.
- 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.
- ORR, H. A., 2006 The distribution of fitness effects among beneficial mutations in Fisher's geometric model of adaptation. *J. Theor. Biol.* **238**: 279–285.
- PARK, S., and J. KRUG, 2007 Clonal interference in large populations. *Proc. Natl. Acad. Sci. USA* **104**: 18135–18140.
- PICKANDS, I. J., 1975 Statistical inference using extreme order statistics. *Ann. Stat.* **3**: 119–131.
- ROKYTA, D. R., P. JOYCE, S. B. CAUDLE and H. A. WICHMAN, 2005 An empirical test of the mutational landscape model of adaptation using a single-stranded DNA virus. *Nat. Genet.* **37**: 441–444.
- SANJUÁN, R., A. MOYA and S. F. ELENA, 2004 The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. *Proc. Natl. Acad. Sci. USA* **101**: 8396–8401.
- SHAW, F. H., C. J. GEYER and R. G. SHAW, 2002 A comprehensive model of mutations affecting fitness and inferences for *Arabidopsis thaliana*. *Evolution* **56**: 453–463.
- TENAILLON, O., O. K. SILANDER, J. UZAN and L. CHAO, 2007 Quantifying organismal complexity using a population genetic approach. *Plos One* **2**: e217.
- WAXMAN, D., and J. J. WELCH, 2005 Fisher's microscope and Haldane's ellipse. *Am. Nat.* **166**: 447–457.
- WELCH, J. J., and D. WAXMAN, 2003 Modularity and the cost of complexity. *Evolution* **57**: 1723–1734.
- WHITLOCK, M. C., 2000 Fixation of new alleles and the extinction of small populations: drift load, beneficial alleles, and sexual selection. *Evolution* **54**: 1855–1861.
- WILKE, C. O., 2004 The speed of adaptation in large asexual populations. *Genetics* **167**: 2045–2053.

Communicating editor: M. W. FELDMAN

APPENDIX A: EXACT DISTRIBUTION OF MUTATION
FITNESS EFFECTS AND DISPLACED
GAMMA APPROXIMATION

The distribution of the log-relative fitness among random mutants in the general version of the FM can be expressed as

$$s = \log\left(\frac{W(\mathbf{z}_o + \mathbf{dz})}{W(\mathbf{z}_o)}\right) = -\mathbf{z}_o^T \cdot \mathbf{S} \cdot \mathbf{dz} - \frac{1}{2} \mathbf{dz}^T \cdot \mathbf{S} \cdot \mathbf{dz} \quad (\text{A1})$$

(MARTIN and LENORMAND 2006b), where \mathbf{dz} (the mutational effects on phenotypic traits) is distributed as a multivariate normal $\mathbf{dz} \sim N(\mathbf{0}, \mathbf{M})$. This is a quadratic form in Gaussian vectors; it can always, without loss of generality, be expressed in diagonal form (JASCHKE *et al.* 2004); *i.e.*, in a new basis where the new phenotypic vectors (\mathbf{x}) are linear combinations of the original phenotypic vectors (\mathbf{z}),

$$s = \log\left(\frac{W(\mathbf{x}_o + \mathbf{dx})}{W(\mathbf{x}_o)}\right) = \boldsymbol{\delta}^T \cdot \mathbf{dx} + \frac{1}{2} \mathbf{dx}^T \cdot \boldsymbol{\Lambda} \cdot \mathbf{dx}, \quad (\text{A2})$$

where $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n)$ is an $n \times n$ diagonal matrix where the $\lambda_i \leq 0$ are the n eigenvalues of $-\mathbf{S} \cdot \mathbf{M}$, \mathbf{dx} is distributed as a standard multivariate Gaussian, $\mathbf{dx} \sim N(\mathbf{0}, \mathbf{I}_n)$, and $\boldsymbol{\delta}^T = \mathbf{x}_o^T \cdot \boldsymbol{\Lambda} = \{\delta_1, \dots, \delta_n\}$, $\mathbf{x}_o = \{x_1, \dots, x_n\}$ is simply \mathbf{z}_o expressed in the new basis. It is important to note that the expression in (A2) is a particular type of quadratic form as $\boldsymbol{\delta} = \boldsymbol{\Lambda} \cdot \mathbf{x}_o$; it has a zero element wherever there is a zero eigenvalue in matrix $\boldsymbol{\Lambda}$. This implies that the dimension of the whole system is not n , but the number $m \leq n$ of nonzero eigenvalues λ_i (*i.e.*, the rank of $\boldsymbol{\Lambda}$). As a consequence, we can always express (A2) in a positive-definite form by focusing only on these m dimensions by setting $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_m)$, where all $\lambda_i < 0$ and $\boldsymbol{\delta}^T = \mathbf{x}_o^T \cdot \boldsymbol{\Lambda} = \{\delta_1, \dots, \delta_m\}$, where, by identification, $\delta_i = \lambda_i x_i$. This argument guarantees that we can apply JASCHKE *et al.*'s (2004) proposition 3.3, which is valid for positive definite $\boldsymbol{\Lambda}$, even when \mathbf{S} and \mathbf{M} are not positive definite but only semidefinite. The distribution of s defined in (A2) is bounded on its rightmost end by $s_o = -\log(W(\mathbf{0})/W(\mathbf{x}_o)) = -\frac{1}{2} \mathbf{x}_o \cdot \boldsymbol{\Lambda} \cdot \mathbf{x}_o \geq 0$, which is the selection coefficient of the optimum phenotype ($\mathbf{x} = \mathbf{0}$) relative to the wild type ($\mathbf{x} = \mathbf{x}_o$). For consistency with JASCHKE *et al.*'s (2004) notation, note that s_o can also be expressed as $s_o = -\frac{1}{2} \mathbf{x}_o \cdot \boldsymbol{\Lambda} \cdot \mathbf{x}_o = -\sum_{i=1}^m \lambda_i x_i^2 / 2$ or equivalently as $s_o = -\sum_{i=1}^m \delta_i^2 / (2\lambda_i)$.

The distribution of s on its whole range $[-\infty, s_o]$ can be approximated by a displaced gamma distribution that has been introduced by SHAW *et al.* (2002) for the analysis of mutation fitness effects distributions with beneficial mutations. The resulting approximate pdf of s is given by

$$f_I(s) = \frac{e^{-((s-s_o)/\alpha)} (s-s_o)^{\beta-1} \alpha^{-\beta}}{\Gamma(\beta)}, \quad (\text{A3})$$

where $\Gamma(\cdot)$ is the gamma function, the shape β and scale α are chosen to fit the mean and variance of $f(s)$, and the displacement parameter is s_o , the maximum s (MARTIN and LENORMAND 2006b). When the wild type is at any fitness distance s_o from the optimum, these parameters are approximately

$$\beta \approx \beta_o \frac{(1+\varepsilon)^2}{(1+2\varepsilon)} \quad \text{and} \quad \alpha \approx \alpha_o \frac{(1+2\varepsilon)}{(1+\varepsilon)} \quad (\text{A4})$$

(from Equation 6 of MARTIN and LENORMAND 2006b), where β_o and α_o are the shape and scale (respectively) of the gamma distribution that fits $f(s)$ at the optimum (*i.e.*, when there are only deleterious mutations), and where $\varepsilon = s_o/\bar{s}$ is the distance to the optimum of the wild type (s_o), measured in terms of fitness and scaled by the average fitness effect of mutation $\bar{s} \equiv -E(s)$. The variable ε describes the degree of adaptation of the wild type (s_o) relative to the average fitness effect of single mutations (\bar{s}). Note that the expressions for α and β in (A4) are only approximate, based on an approximation for the moments of $f(s)$ as a function of s_o , Equation A4 of MARTIN and LENORMAND (2006b), but the displaced gamma remains a good approximation of $f(s)$, even when the best-fitting parameters are not exactly those given in (A4) (see supplemental Figure 1).

APPENDIX B: TAIL BEHAVIOR OF QUADRATIC
FORMS IN GAUSSIAN VECTORS

There is no analytic expression for $f(s)$ as defined in (A2), but it has a simple tail behavior as s approaches its rightmost endpoint s_o . We have seen above that, even when \mathbf{S} or \mathbf{M} are positive semidefinite (the most general case), the system can always be reduced to positive definite of dimension $m = \text{rank}(\mathbf{S} \cdot \mathbf{M})$. Therefore we can always apply JASCHKE *et al.*'s (2004) asymptotic approximation (Equations 3.23 and 3.24, p. 261 of their article with all $\lambda_i < 0$) in our context. In what follows we denote results derived from this tail approximation by an asterisk (*). Applying the tail approximations shows that, to the leading order in $(s - s_o)$, $f(s)$ approaches

$$f(s) \underset{s \rightarrow s_o}{\approx} f^*(s) = d(s_o - s)^{(m/2)-1}, \quad (\text{B1})$$

where

$$d = \frac{e^{-(1/2) \sum_{i=1}^m (\delta_i/\lambda_i)^2}}{\Gamma(m/2) \prod_{i=1}^m \sqrt{|\lambda_i|}} \quad (\text{B2})$$

is a constant depending on λ_i and δ_i . Here we have assumed that all nonzero eigenvalues have multiplicity of 1 (*i.e.*, they are distinct for each of m traits). With arbitrary multiplicity, the only change is in the expression of d in (B2) (for details, see JASCHKE *et al.* 2004). From the above tail approximation one easily retrieves the proportion of beneficial mutations

$$p_b \underset{s_b \rightarrow 0}{\approx} p_b^* = \int_{s=0}^{s_b} f^*(s) ds = 2d \frac{s_b^{m/2}}{m} \quad (\text{B3})$$

and their pdf $f_b(s_b)$, provided that s_o is close to 0, so that all $s_b > 0$ are also close to 0:

$$f_b(s_b) \underset{s_b \rightarrow 0}{\approx} f_b^*(s_b) = \frac{f^*(s_b)}{p_b^*} = \frac{m}{2s_o} \left(1 - \frac{s_b}{s_o}\right)^{(m/2)-1}. \quad (\text{B4})$$

This is equivalent to stating that the approximate distribution of s_b/s_o , when s_o is small, is a beta with shape parameters 1 and $m/2$:

$$\frac{s_b}{s_o} \sim \text{beta}\left(1, \frac{m}{2}\right). \quad (\text{B5})$$

The cumulative distribution function (cdf) of the beta distribution above ($F_\beta(x)$) is approximately equal to that of the exponential distribution (with rate $m/2$) as m gets large,

$$F_\beta\left(x = \frac{s_b}{s_o}\right) = 1 - (1-x)^{m/2} \xrightarrow{m \rightarrow \infty} 1 - e^{-mx/2}, \quad (\text{B6})$$

which is why the classic FM (many independent traits, *i.e.*, m large) yields an exponential distribution of beneficial effects, consistent with the MLM.

Finally, note that the same kind of tail behavior as in Equation B1 is obtained with the displaced gamma approximation defined in Equation A3,

$$f_\Gamma(s) \underset{s \rightarrow s_o}{\approx} d'(s_o - s)^{\beta-1}, \quad (\text{B7})$$

where $d' = \alpha^{-\beta}/\Gamma(\beta)$ is a constant. However, while the exact $f(s)$ and the gamma approximation $f_\Gamma(s)$ have the same tail behavior *qualitatively* (compare Equations B1 and B7), they differ *quantitatively*, as $d \neq d'$ and $m/2 \neq \beta$. Therefore, while the displaced gamma approximation is fairly accurate for the whole distribution of s (supplemental Figure 1), it is less so for the subset of beneficial mutations $f_b(s_b)$, when s_o is small, in which case the beta approximation in Equations B4 and B5 is the most accurate. As s_o gets large, the displaced gamma approximation will become the most accurate for both $f(s)$ and $f_b(s_b)$. Notably, as s_o gets close to 0, $\beta \rightarrow \beta_o = n_e/2$ (see Equation A2 and MARTIN and LENORMAND 2006b), so that the discrepancy between the two tail behaviors [(B1) *vs.* (B7)] depends on the difference between the “effective number of traits” n_e and the “dimensionality” m .

APPENDIX C: APPROXIMATE DISTRIBUTION OF FIXED EFFECTS

The distribution of fixed effects s_f is that of beneficial effects, conditional on escaping drift loss, with probability $\pi(s_f)$. The pdf of fixed mutation effects, $f_{\text{fix}}(s_f)$, is therefore

$$f_{\text{fix}}(s_f) = \frac{\pi(s_f) f_b(s_f)}{\int_{x=0}^{s_o} \pi(x) f_b(x) dx}, \quad (\text{C1})$$

ignoring the fixation of deleterious mutations. In the general case, this expression cannot be explicitly derived. However, when the pdf $f_b(s_b)$ of beneficial mutations follows the Jäschke tail approximation of Equation B1, a simple approximation for $f_{\text{fix}}(s_f)$ can be derived. To do so, one simply replaces $f_b(\cdot)$ in Equation C1 by the tail approximation $f_b^*(\cdot)$ (in Equation B4) and uses the weak selection, large population approximation for the fixation probability: $\pi(s) \approx 2s$ (HALDANE 1927). As all beneficial mutations are $< s_o$, which is assumed to be small, this weak selection approximation should always be fairly accurate whenever the Jäschke tail approximation is valid (wild type well adapted). The pdf of the distribution of fixed effects is approximately

$$\begin{aligned} f_{\text{fix}}(s_f) \underset{s_o \rightarrow 0}{\approx} f_{\text{fix}}^*(s_f) &= \frac{2s_f f_b^*(s_f)}{\int_{x=0}^{s_o} 2x f_b^*(x) dx} \\ &= \frac{m(2+m)}{4s_o^2} \left(1 - \frac{s_f}{s_o}\right)^{(m/2)-1} s_f, \end{aligned} \quad (\text{C2})$$

which means that the distribution of s_f/s_o is a beta with shape parameters 2 and $m/2$:

$$\frac{s_f}{s_o} \sim \text{beta}\left(2, \frac{m}{2}\right). \quad (\text{C3})$$

Note that for populations of finite (though not too small) size N and effective size N_e (and still with weak selection as assumed here), the probability of fixation of a beneficial allele becomes $\sim 2s^* N_e/N$ (WHITLOCK 2000). This scaling factor does not affect Equation C2 so that the distribution of fixed mutation effects is not affected by finite population sizes (*i.e.*, only the probability of a beneficial mutation fixing is affected, not its effect distribution). Note, however, that for smaller N , the above approximation is inaccurate *a priori*; in particular, deleterious mutations can fix, which was neglected here.

As a comparison, we compute, in the same way, the distribution of fixed effects when the distribution of beneficial mutation effects is exponential with rate λ (note that the integral in the numerator of C2 is over the range $[0, \infty]$ for the exponential). The distribution of fixed mutation effects is then

$$f_{\text{exp}}(s_f) = \lambda^2 s_f e^{-\lambda s_f}, \quad (\text{C4})$$

and the average fitness effect of fixed mutations when beneficial effects are exponentially distributed is

$$E_{\text{exp}}(s_f) = \frac{2}{\lambda}, \quad (\text{C5})$$

where in the limiting case of a large m , $\lambda = m/2s_o$ is the rate of the exponential distribution of beneficial effects s_b (as $x = s_b/s_o$ is exponential with rate $m/2$, see Equation B6).