Evolution, 60(5), 2006, pp. 893–907

A GENERAL MULTIVARIATE EXTENSION OF FISHER'S GEOMETRICAL MODEL AND THE DISTRIBUTION OF MUTATION FITNESS EFFECTS ACROSS SPECIES

GUILLAUME MARTIN[1,2] AND THOMAS LENORMAND[1]
[1]*Centre d'Ecologie Fonctionnelle et Evolutive, Centre National de la Recherche Scientifique, UMR 5175, 1919 Route de Mende, 34 293 Montpellier, France*

Abstract.—The evolution of complex organisms is a puzzle for evolutionary theory because beneficial mutations should be less frequent in complex organisms, an effect termed ''cost of complexity.'' However, little is known about how the distribution of mutation fitness effects ($f(s)$) varies across genomes. The main theoretical framework to address this issue is Fisher's geometric model and related phenotypic landscape models. However, it suffers from several restrictive assumptions. In this paper, we intend to show how several of these limitations may be overcome. We then propose a model of $f(s)$ that extends Fisher's model to account for arbitrary mutational and selective interactions among n traits. We show that these interactions result in $f(s)$ that would be predicted by a much smaller number of independent traits. We test our predictions by comparing empirical $f(s)$ across species of various gene numbers as a surrogate to complexity. This survey reveals, as predicted, that mutations tend to be more deleterious, less variable, and less skewed in higher organisms. However, only limited difference in the shape of $f(s)$ is observed from *Escherichia coli* to nematodes or fruit flies, a pattern consistent with a model of random phenotypic interactions across many traits. Overall, these results suggest that there may be a cost to phenotypic complexity although much weaker than previously suggested by earlier theoretical works. More generally, the model seems to qualitatively capture and possibly explain the variation of $f(s)$ from lower to higher organisms, which opens a large array of potential applications in evolutionary genetics.

Key words.—Complexity, Fisher model, mutation fitness effect distribution, mutational and selective covariances, random matrices, survey.

Received July 22, 2005. Accepted February 18, 2006.

Mutations are the raw material for evolution. As a consequence, predicting and estimating the distribution of the effects of mutations on fitness (hereafter $f(s)$) is a central issue for many aspects of evolutionary theory (Lynch et al. 1999; reviewed in Charlesworth and Charlesworth 1998; Bataillon 2000). Considerable effort has been devoted, mostly empirically, to address this issue. Yet, we are still largely ignorant of how this distribution varies among organisms and how to explain this variation (Lynch et al. 1999; Bataillon 2003; Keightley and Lynch 2003; Shaw et al. 2003), although significant differences across species can be observed, even at small phylogenetic scales (Baer et al. 2005). From a theoretical standpoint, adaptive landscape modeling, such as Fisher's model (1930), is the most commonly used framework to predict $f(s)$. Fisher modeled an organism as a vector of n trait values (i.e., a position in the phenotypic space), whose fitness is determined by the distance of this phenotype to a given optimum. Mutation randomly displaces phenotypes in this n-dimensional space, which allows one to compute $f(s)$. Fisher's model is appealing because it predicts $f(s)$ based on selective and mutational assumptions on the underlying phenotypic traits (as in, e.g., Welch and Waxman 2003). Other approaches have been suggested to predict some features of $f(s)$. In particular, extreme value theory can be used to characterize the right-tail behavior of $f(s)$ (advantageous mutations) with a minimal set of assumptions regarding $f(s)$ itself (Orr 2003, 2005a). However, by avoiding describing $f(s)$, this approach also has the weakness of disconnecting properties of advantageous and deleterious mutations, which may jointly affect some evolutionary process (e.g., the rate of adaptation in asexuals; Gerrish and Lenski 1998). In contrast, Fisher's model predicts the full distribution with both deleterious and advantageous effects under a given set of assumptions. Its prediction may thus be compared with most empirical data on $f(s)$, which mainly describe the distribution of deleterious mutation effects (Lynch et al. 1999).

Fisher's model makes simplifying assumptions, whose realism may be questioned (Clarke and Arthur 2000; Orr 2000, 2001). As a consequence, this model is often considered to have a heuristic but not a quantitative value (Orr 2005a). For example, its predictions regarding $f(s)$ have never been confronted to empirical distributions although qualitative predictions on the adaptive process have received empirical support (Burch and Chao 1999; Imhof and Schlotterer 2001; Rozen et al. 2002; Rokyta et al. 2005). Similarly, its predictions regarding $f(s)$ are rarely used within theoretical models (but see Poon and Otto 2000) because its lack of realism may then compromise the models' conclusions. Therefore, many theoretical approaches are limited to those that do not depend too much on $f(s)$ (Orr 2005b; Otto 2004) which is a desirable property but potentially restrains the scope of theoretical investigation: when variation in $f(s)$ affects theoretical predictions, the problem is still ultimately to find what $f(s)$ really is. Paradoxically, the problem of $f(s)$ has been much more discussed for the statistical analysis of mutation accumulation experiments (Keightley 1994; Keightley and Lynch 2003; Shaw et al. 2003), without much guidance from a ''realistic'' theoretical expectation. More generally, although several papers sought to explain the variation in rates of mutation across taxa (Drake et al. 1998; Lynch et al. 1999; Keightley and Eyre-Walker 2000), variation in its fitness effect and its potential causes has received much less attention (but see Lynch et al. 1999; Bataillon 2000). The main goal of this paper is

[2] Present address: Department of Ecology and Evolution, Université de Lausanne, CH1015 Lausanne, Switzerland; E-mail: guillaume.martin@unil.ch.

to describe this variation with a survey of the empirical literature, and to analyze it in the light of a generalized version of Fisher's model.

Building and testing a model for $f(s)$ necessarily involves simplifying assumptions because an exhaustive description of the effect of all possible mutations on phenotype and their fitness consequences is obviously intractable. Fisher's model makes several specific simplifying assumptions that have been criticized. In this paper, we propose a model relaxing some of these assumptions, but before describing our approach, we discuss the main criticisms of Fisher's model that have been put forward. This brief presentation draws on previous discussions of this issue (Orr 1998, 2005a; Poon and Otto 2000; Welch and Waxman 2003).

Fisher's Model Realism

Optimum and fitness

The first key assumption of the model is that there is a single phenotypic optimum, so that it can only be used to model stabilizing selection (around a single optimum) or directional selection (away from the optimum). Quantitative genetic studies (reviewed in Kingsolver et al. 2001) and experimental evolution experiments (reviewed in Elena and Lenski 2003) provide evidence for both types of selection, but also for disruptive selection with alternative possible optima. However, using Fisher's model to predict $f(s)$ only requires that most of the range of possible mutations lies on the slope of a single local fitness peak. It may indeed be quite rare that a population stays long in a fitness valley. In any case, predicting local $f(s)$ is different from predicting long term evolutionary change, and Fisher's model may be less appropriate for the latter than for the former. Similarly, although the pattern of adaptation cannot be directly predicted with a moving optimum (Orr 2005a), $f(s)$ can still be predicted at any given time. An additional restriction is that a given fitness function must be chosen to map fitness with the phenotypic distance from the optimum. When close to the optimum, a quadratic or Gaussian fitness function is a straightforward local approximation for many arbitrary fitness functions (Lande 1980) and is therefore widely used. Overall, these assumptions of Fisher's model on the way phenotype determines fitness are not so unrealistic when considering a population close to a local optimum, but may be less accurate under strong environmental change.

Mutation and traits

In Fisher's model, fitness is determined by the phenotype, which is made of a set of n phenotypic traits describing adaptations (hereafter ''adaptation traits''). A more specific assumption is that the effect of mutation on these traits is continuous and symmetric. This assumption is also used in quantitative genetics and is useful for the mathematical analysis. It has been criticized (Clarke and Arthur 2000), but is consistent with some empirical evidence (Garcia-Dorado et al. 1999; Orr 2001); for example, based on the effect of single P-element inserts on bristle numbers (Lyman et al. 1996). There is obviously a qualitative difference between such adaptation traits for which we can reasonably assume symmetrical mutation effects and fitness traits (e.g., survival, fecundity, functional efficiency traits, etc.) which are known to decrease on average by mutations, since most mutations are deleterious (Lynch et al. 1999). For example, molecular properties of the active site of an enzyme (e.g., volume, charge, etc.) can be reasonably expected to change symmetrically by mutation. There are indeed no clear reasons why mutations should bias toward higher or lower charge or volume; and if such a bias is present, it is likely to be small compared to the mutational variance. In contrast, when considering a trait that measures a distance from a given value (e.g., departure from the charge or volume corresponding to maximal affinity), mutation effects cannot be symmetric since a distance is never negative: both an increase and a decrease in charge from the optimal value will decrease affinity. The same difference applies between adaptation traits (the axes in Fisher's model) and fitness traits: the former have an optimal value which is defined by the latter. The aim of the model is precisely to predict the distribution of the effect on fitness traits of mutations affecting adaptation traits.

The distribution of mutation phenotypic effects is also often assumed to be Gaussian (we will use this assumption too), again mainly for mathematical convenience. Empirically, this assumption is approximately valid for some traits, but clearly fails for others (Garcia-Dorado et al. 1999). We suggest here that the measure of each trait is somehow arbitrary, so that, with appropriate scaling, it is possible to reduce kurtosis as needed (in the same way as transforming a response variable in a linear model, to conform to the hypothesis of normal errors). Such a scaling would affect the fitness function accordingly, but this new fitness function could still be approximated by a quadratic function close to the optimum. In all that follows, it will be important to keep in mind the distinction between the effect of mutations on phenotype (i.e., on adaptation traits) and on fitness (i.e., $f(s)$): the former will be assumed Gaussian, whereas the latter will be predicted by the model.

A perhaps stronger assumption of Fisher's model is that a single mutation can potentially affect all the phenotypic traits of the organism (universal pleiotropy). There is empirical evidence showing that pleiotropy and compensatory mutations are widespread (reviewed in Poon and Otto 2000). However, studies of development and genetic regulatory networks strongly suggest that the genotype-phenotype map may be organized into modules. As a consequence, it is often argued that compensation can only occur within modules or that a change within a module leaves adaptation in other modules undisturbed (Wagner and Altenberg 1996). Although this view is certainly correct for phenotypic compensation, it may not hold when considering fitness compensation. For instance, if two traits in different modules are correlated in their effect on fitness, then the fitness effect of a mutation in a given module may be compensated by mutation in another module. Similarly, with selective correlation among traits in different modules, a mutation in a given module can cause maladaptation in another module. The modularity of mutation effects on phenotypes does not ensure the modularity of their fitness effects. In any case, Fisher's model only requires a description of the net phenotypic effect of all mutations averaged over modules and can thus accommodate

modularity or partial pleiotropy (as in Welch and Waxman 2003).

Finally, the distribution of the phenotypic effects of mutations in Fisher's model is always assumed to be independent of the genotype in which they arise (the genetic background). This assumption is not a problem when considering the distribution of mutations effects on a single genotype. However, considering that mutational effects vary with the background may alter long-term predictions in which the background changes. There is qualitative support for the idea that a given allele has similar phenotypic effects when introduced in different genetic backgrounds as demonstrated by the success of genetic engineering and improvement of domesticated species. It is probably quantitatively inexact, but remains a good working assumption, in the absence of alternative models.

Symmetry assumptions

Another strongly restrictive assumption in Fisher's model is that all traits are equivalent and independent of each other with respect to both mutation and selection (Welch and Waxman 2003). This spherical symmetry is indeed an oversimplification that can hardly apply to real organisms. Given appropriate scaling, it is possible to account for some correlations between traits, but such scaling can only be performed for either selection, or mutation effects, not for both (Orr 1998, 2005a; Poon and Otto 2000). There is empirical evidence showing that phenotypic correlations between traits are widespread for both selection (Kingsolver et al. 2001) and mutation (Garcia-Dorado et al. 1999; Keightley et al. 2000), so that relaxing the symmetry assumption would significantly improve the model's realism and the range of its potential applications. Recently, Waxman and Welch (2005) have proposed the first model to account for selective interactions between traits, but still neglecting mutational correlations. One of the aims of this paper is to relax the symmetry assumptions for both mutation and selection (using a different approach from that of Waxman and Welch 2005).

Scaling and empirical predictions

The last problem when trying to use Fisher's model is to find an appropriate scaling in terms of measurable quantities. For instance, we ignore the distribution of the size of mutation in the phenotypic space, the phenotypic distance to an optimum, or the number n of adaptation traits. As a consequence, even if it has a strong heuristic value, the whole model may appear arbitrary and of little use when it comes down to real and testable predictions (Orr 2005a). Fisher's model can make predictions that are scale independent and testable; for instance, regarding the distribution of factors fixed during a bout of adaptation. These predictions have received empirical support (Imhof and Schlotterer 2001; Rokyta et al. 2005). However, the agreement is qualitative and does not unambiguously support Fisher's model since several models make the same predictions (Orr 2005a).

Another prediction of the model with perhaps important evolutionary implication is that the number of traits, n, influences the fitness effect distribution of mutations in such a way that beneficial mutations are less likely and less favorable in more complex organisms (Orr 1998; Barton and Partridge 2000). An important consequence of this is that more complex organisms should adapt at slower rates, an effect dubbed ''cost of complexity'' (Orr 2000). This prediction has not been tested empirically. Similarly, it has not been tested whether the distribution of mutation fitness effects varies with complexity as predicted by this model because testing this prediction would require scaling mutation fitness effects to measurable quantities.

Possible Improvements and Tests

The first aim of this paper is to propose a model predicting the fitness effect distribution of mutations without assuming equivalence and independence between traits. This model attempts to predict in a simple analytic form how the moments of $f(s)$ should vary with phenotypic complexity (i.e., n, the number of adaptation traits under selection) and with the level of covariation between traits. We then make approximations that allow one to use the empirical distribution of mutation effects measured in a given environment (e.g., in the laboratory) to predict the new distribution in another environment. The second aim of this paper is to survey the available empirical data on $f(s)$ across taxa to test our model's predictions with the appropriate scaling. More precisely, we use gene number as a surrogate estimate of n and test for correlations between gene number and empirical moments of s across species. The third aim of this paper is to use our model and survey to quantify the cost of complexity by comparing predicted rates of adaptation across species, based on their empirical distribution of deleterious mutations. Overall, our survey and model indicate that complexity—as measured by gene number—and phenotypic correlations are critical factors shaping the fitness effect of mutations across taxa.

Model and Predictions

As with other models based on Fisher's geometric approach, we consider that fitness is determined by n adaptation traits. As explained in the introduction, we use a Gaussian distribution of mutation phenotypic effects on these traits. Furthermore, we assume that these traits are under Gaussian stabilizing selection around a fixed phenotypic optimum, which will work best when close to the optimum. These Gaussian assumptions allow the mathematical treatment of the model for any arbitrary selective and mutational covariance matrices for adaptation traits. This model of multivariate stabilizing selection and mutation is similar to that introduced by Zhang and Hill (2003) for the study of mutation-selection balance on a quantitative trait, but extended to account for beneficial mutations (i.e., phenotypes are not necessarily at their optimum). The presentation of the model takes several steps: we first derive the exact distribution of s under our assumptions, then we give a general exact expression for the moments of $f(s)$ at the optimum (i.e., when there are only deleterious mutations). Then we formulate testable predictions on the effect of n on these moments. Next, we derive an approximation for the probability density function of s, $f(s)$ at any distance to a new optimum defined by a new environment. We show how the parameters of this distribution can be estimated empirically and we use this approximation to compute the rate of adaptation in this new envi-

ronment. Next, because all our theoretical results depend on the strength of mutational and selective correlations between traits, we introduce a null model of phenotypic interactions to evaluate their influence on $f(s)$, based on random matrix theory. Next, we present simulations to validate our approximations. Finally, we confront our predictions with empirical data from the literature.

Description of the Model

Phenotypes are modeled as a set of n continuous phenotypic traits represented by a column vector $\mathbf{z}$. The fitness $W(\mathbf{z})$ of phenotype $\mathbf{z}$ is a multivariate Gaussian function of the distance between $\mathbf{z}$ and a phenotypic optimum that is set to zero for all traits, without loss of generality: $W(\mathbf{z}) = \text{Exp}(-\frac{1}{2}\mathbf{z}^t\mathbf{S}\mathbf{z})$, where t denotes transposition. $\mathbf{S}$ is the $n \times n$ matrix of the selective effects of all traits. Diagonal elements in $\mathbf{S}$ measure the selection intensity on each trait while nondiagonal elements measure selective interactions between trait pairs. To describe the most general situation of stabilizing selection, we assume that $\mathbf{S}$ is positive semidefinite (not strictly definite as often assumed), which ensures that no phenotype has a higher fitness than phenotype $\mathbf{z} = 0$. Note that we will only consider traits under direct stabilizing selection (i.e., with a strictly positive diagonal term), although some linear combinations of these traits may be neutral because of selective interactions.

Consider now the effect of mutation on a single genotype, referred to as the initial genotype (with phenotype $\mathbf{z}_o$). We assume that the distribution of mutant phenotypes around $\mathbf{z}_o$ is multivariate Gaussian with mean zero and covariance matrix $\mathbf{M}$. As explained in introduction, this Gaussian assumption is valid as long as there is a transformation from the ''real'' traits to a set of traits that are distributed as a multivariate Gaussian. As for selection, the model allows both for differences in mutational variances across traits and for mutational correlations between traits (nondiagonal elements in $\mathbf{M}$). We assume that $\mathbf{M}$ is positive semidefinite (the most general structure for a covariance matrix). The selection coefficient s of a mutant phenotype $\mathbf{z}_o + \mathbf{dz}$ is defined relative to the initial phenotype $\mathbf{z}_o$ as $W(\mathbf{z}_o + \mathbf{dz})/W(\mathbf{z}_o) - 1$. We assume that s is small enough that $s \approx \log(1 + s)$ and we define $s_o \equiv -\log(W(\mathbf{z}_o)/W(\mathbf{0}))$, the selective disadvantage of the initial phenotype $\mathbf{z}_o$ relative to the optimal phenotype. Under these assumptions, the joint effects of all selective and mutational covariances (matrices $\mathbf{M}$ and $\mathbf{S}$) reduce to the n eigenvalues of the product $\mathbf{S}\cdot\mathbf{M}$. The exact distribution of s is a quadratic form in Gaussian vectors (Mathai and Provost 1992). The distribution is entirely determined by the distance to the optimum s_o, the direction to the optimum in the phenotypic space, and the n eigenvalues $\{\lambda_i\}_{i\in[1,n]}$ of $\mathbf{S}\cdot\mathbf{M}$, as shown in Appendix 1, available online only at http://dx.doi.org/10.1554/05-412.1.s1. Each λ_i corresponds to a phenotypic direction (a linear combination of biological traits z_i) on which mutation and selection act independently with a net effect λ_i on fitness. Thus, a large λ_i corresponds to a combination of traits that displays a large mutational variance and is under strong selection.

Exact Moments of f(s) at the Optimum

A distribution can be fully characterized by its moments. Since the central moments of quadratic forms of Gaussian vectors have analytic expression for any order (Mathai and Provost 1992; online Appendix 1), $f(s)$ is fully specified in our model. However, we focus on the first three central moments of s that are the most available empirically. When most mutations are deleterious, it can be assumed that the initial genotype is close to the optimum ($s_o \ll 1$). Then, defining the raw moments of the λ_i across traits i as $\overline{\lambda^r} \equiv 1/n \sum_{i=1}^{n} \lambda_i^r$, the three first central moments ($E(s)$, $V(s)$, and $\mu_3(s)$) of $f(s)$ are given by:

$$E(s) = -\frac{n}{2}\bar{\lambda}$$

$$V(s) = \frac{n}{2}\overline{\lambda^2}$$

$$\mu_3(s) = -n\overline{\lambda^3}. \qquad (1)$$

A more general expression for any distance to the optimum (i.e., with beneficial mutations) can be found in equation (A3) of Appendix 1 (available online). At the optimum, the moments of s depend only on the number of phenotypic traits n and on the distribution of the eigenvalues of $\mathbf{SM}$ (the λ_i).

Predicting the Effect of n on the Moments of f(s)

Equation (1) yields simple predictions on the three first moments of empirical distributions of mutation effects when there are few beneficial mutations ($s_o \ll 1$).

First, $E(s) = -n\bar{\lambda}/2$, so that the average deleterious effect of mutations should be larger in organisms with a presumably larger number of traits (e.g., in *Drosophila* vs. *Escherichia coli*). Larger $E(s)$ in *Drosophila* than in *E. coli* could be due either to a larger n (larger number of traits under selection in fruit flies than in bacteria) or a larger $\bar{\lambda}$ (same number of traits in both species, but a larger effect of mutation on each trait, in *Drosophila*) but the latter seems much less parsimonious. Note that, in addition, the expression for $E(s)$ in equation (1) is still valid when $s_o \neq 0$, that is, if beneficial mutations also occur (see eq. A3, Appendix 1 online). This outcome results directly from the Gaussian approximation (quadratic in log scale) of the fitness function: at the optimum, all mutations are weakly deleterious, whereas, when away from the optimum, deleterious mutations are more severe but compensated by some beneficial mutations. The net outcome depends only on the local curvature of the fitness function around the initial phenotype. This curvature is constant at any distance from the optimum with our quadratic fitness function, so that $E(s)$ is independent of s_o.

Second, rearranging equation (1) shows that both the coefficient of variation and the skewness of s should decrease with n. More precisely, if we note $\mu_2^* = V(s)/E(s)^2$ and $\mu_3^* = \mu_3(s)/E(s)^3$, the second and third moments of s scaled to the mean effect $E(s)$, we obtain from equation (1) two independent quantities that should increase linearly with n:

$$\frac{1}{\mu_2^*} = n\frac{\bar{\lambda}^2}{2\overline{\lambda^2}} \quad \text{and} \quad \frac{1}{\sqrt{\mu_3^*}} = n\sqrt{\frac{\bar{\lambda}^3}{8\overline{\lambda^3}}}. \qquad (2)$$

These predictions are based on scale invariant measures (scaled to $E(s)$), which should make them more robust for comparisons across species. Again, finding that $1/\mu_2^*$ and $1/\sqrt{\mu_3^*}$ are larger, for example in *Drosophila* than in *E. coli*, is more likely due to a larger number of traits in *Drosophila* than to variation in the distribution of the λ_i between these species.

We propose below (see A Model of Random Phenotypic Correlations) a null model for **S** and **M** to better understand the possible dependence between n and the distribution of the λ_i. Under this model, $E(s)$ increases linearly with n while $1/\mu_2^*$ and $1/\sqrt{\mu_3^*}$ increase but plateau with large n. If this model holds, we should indeed find larger $E(s)$ in more complex organisms (first prediction) but may not detect such an increase for $1/\mu_2^*$ and $1/\sqrt{\mu_3^*}$ (second prediction).

Distribution of Deleterious Effects and ''Effective Complexity'' n_e

Approximating the probability density of s is required for computing the rate of adaptation and may be useful in maximum likelihood analyses of empirical distributions of s. Equation (A2) in Appendix 1 (online), gives the exact distribution in the general case but its probability density is not known. However, an approximation can be obtained from the moments of s given above, using the moment matching method, which requires choosing an a priori distribution for the density. We chose the negative gamma distribution because it is the exact distribution of s corresponding to the simplest situation: when all $\lambda_i = \lambda$ are equal and $s_o = 0$, $f(s)$ is a negative gamma distribution with scale λ and shape $n/2$. Staying at the optimum ($s_o = 0$) but with the λ_i varying across traits, with coefficient of variation $CV(\lambda)$, $f(s)$ can be approximated by a negative gamma with scale $\lambda_e = \bar{\lambda}(1 + CV(\lambda)^2)$ and shape $n_e/2$ where n_e is the effective number of traits. n_e is defined as the number of traits that would generate the same mean and variance of $f(s)$ (i.e., the same parameters for the approximate gamma distribution), if all traits where independent and of equal effect (λ_e) as in the original Fisher model. From equation (1), it is given by

$$n_e = \frac{n}{1 + CV(\lambda)^2}. \tag{3}$$

This n_e is lower than n and decreases relative to n when the heterogeneity among traits (measured by $CV(\lambda)$) increases. This effect is strongest when only a few linear combinations of traits display both large mutational variance and are under strong selection (i.e., correspond to major λ_i). A small n_e means that the distribution of s is highly skewed, while a high n_e corresponds to more symmetrical distributions, closer to the Gaussian.

Based on a distribution of deleterious mutation effects (i.e., at the optimum $s_o = 0$), both n_e and λ_e can be directly estimated from the mean and variance of s as

$$n_e = 2\frac{E(s)^2}{V(s)} \quad \text{and} \tag{4a}$$

$$\lambda_e = \frac{V(s)}{-E(s)}. \tag{4b}$$

This will be used later to estimate the effective number of traits n_e across species from empirical distributions of deleterious mutation effects.

Predicting f(s) Away from the Optimum

When the initial genotype is at the optimum, empirical distributions of deleterious mutation effects can be used to estimate n_e and λ_e. It is thus tempting to ask whether we can then use this information to predict $f(s)$ in any new environment, in which the initial genotype is not at the optimum but at a distance s_o from the optimum. When the initial genotype is away from the optimum ($s_o > 0$), we use a similar approximation as above, and $f(s)$ becomes a ''displaced gamma'' (Shaw et al. 2002)—the sum of a negative gamma and the constant s_o: $s = s_o - \gamma$, where γ is approximately gamma distributed with scale α and shape β (see online Appendix 1). We will denote $f_\Gamma(s)$ this approximation for the probability density $f(s)$ of s, because it rests on the approximation that the random part γ of the distribution of s is a gamma deviate:

$$f_\Gamma(s) = \frac{e^{-(s_o - s)/\alpha}(s_o - s)^{\beta-1}\alpha^{-\beta}}{\Gamma(\beta)}. \tag{5}$$

Such a distribution can account for both advantageous and deleterious mutations in a continuous manner and with only three parameters and can be implemented in maximum likelihood analysis of empirical $f(s)$ (Shaw et al. 2002). As for the case $s_o = 0$, α and β must be set to match the mean and variance of s away from the optimum $E(s)$' and $V(s)$'. The resulting approximation is accurate but depends on both the distance to the optimum s_o and the particular direction $\mathbf{z}_o$ of the initial genotype (see eq. A3 in online Appendix 1). The distance to the optimum s_o could be estimated in principle (e.g., using long-term experimental evolution), whereas the direction ($\mathbf{z}_o$) may not be measurable. As a consequence, this approximation may be of little interest. Fortunately, $f(s)$ does not vary too much with the direction of $\mathbf{z}_o$, so that we can find a less accurate approximation for $f(s)$ that depends only on the distance to the optimum s_o and on the moments of $f(s)$ at the optimum (eq. A4 in online Appendix 1). The resulting α and β of the displaced gamma approximation in equation (5) can then be predicted based on estimable quantities (λ_e, n_e, and s_o) yielding

$$\beta = \frac{n_e}{2}\frac{(1 + \varepsilon)^2}{1 + 2\varepsilon} \quad \text{and} \tag{6a}$$

$$\alpha = \lambda_e\frac{1 + 2\varepsilon}{1 + \varepsilon}, \tag{6b}$$

where n_e and λ_e can be estimated from a distribution of deleterious effects (equation 4) , and $\varepsilon = s_o/|E(s)| = 2s_o/(n_e\lambda_e)$ is the distance to the optimum relative to the average fitness effect of a mutation $E(s)$. With this approximation, it is possible to predict $f(s)$ in a new environment (with a given $s_o > 0$) from the mean and variance of mutation effects measured close to the optimum (i.e., on deleterious mutations). Importantly, this means that we can predict the approximate $f(s)$ for any species in which deleterious mutation effects have been measured and in an environment for which s_o is known.

Rates of Adaptation and Cost of Complexity

The approximation $f_\Gamma(s)$ above (eq. 5) can now be used to compute the rate of adaptation to a new environment (corresponding to a given s_o), defined as the per generation increase in mean fitness, $d\bar{W}/dt$. In a population of very large size N, only beneficial mutations ($s > 0$) reach fixation (with probability $2s$, we ignore possible complications due to linkage) so that using the displaced gamma approximation yields:

$$\frac{d\bar{W}}{dt} \approx NU \int_{s>0} 2s^2 f_\Gamma(s)\, ds = N\mu E(s)^2 F(n_e, \varepsilon), \qquad (7)$$

where U is the per generation per genome mutation rate. $E(s)$, $\varepsilon = s_o/|E(s)|$, and n_e have been defined above. The function $F(\cdot)$ has a complicated expression but can be computed simply from equations (5) and (7) Interpretation of equation (7) is consistent with previous studies based on the Fisher-Orr model (Orr 2000; Welch and Waxman 2003). First, F is an increasing function of ε, meaning that the rate of adaptation increases with the maladaptation of the initial genotype. Second, F decreases with increasing complexity n: this is Orr's (2000) cost of complexity. The effect of n on the rate of adaptation may also be influenced by any covariation of other parameters (N, U, s_o, $E(s)$) with n. In particular, $|E(s)|$ may increase with n, as suggested by equation (1) (and confirmed by our survey below). This has antagonistic effects on the rate of adaptation by increasing $E(s)^2$ but decreasing ε. Most importantly, our model shows that it is n_e, not n, that determines the rate of adaptation. Therefore, including heterogeneity between traits greatly reduces the cost of complexity by reducing the effective number of traits n_e. Overall, variation in $f(s)$ affects the rate of adaptation by changing $E(s)^2 F(n_e, s_o/E(s))$, in which both $E(s)$ and n_e can be estimated from empirical distributions of deleterious mutation effects (see eq. 4).

A Model of Random Phenotypic Correlations

We considered so far arbitrary mutational and selective matrices **S** and **M**. To understand the relationship between the strength of phenotypic correlations and $f(s)$, we propose here a model of random interactions between many traits, using results from random matrix theory, a mathematical tool widely used in physics and finance to model complex interactions (Forrester et al. 2003). The available evidence suggests that both positive and negative phenotypic correlations are widespread (Lynch and Walsh 1998), so we considered a case where correlations of both signs are equally probable. Mutational and selective covariance matrices **S** and **M** can, for example, be assumed to be drawn randomly into independent Wishart distributions (a classic model for random covariance matrices, see Appendix 2, available online only at http://dx.doi.org/10.1554/05-412.1.s2). These matrices are built by drawing the elements of a first matrix into the standard Gaussian distribution, and multiplying it by its transpose to obtain a symmetric positive semidefinite matrix. Standard Wishart matrices contain a random set of both positive and negative correlations with a zero average. If the number of traits is sufficiently large (e.g., $n > 15$), the distributions of phenotypic correlations in **S** and **M** converge to a simple distribution with known probability density (see eq. A5 in online Appendix 2). Let ρ_S and ρ_M be the standard deviations of these asymptotic distributions around zero. ρ_S and ρ_M measure the strength of selective and mutational correlations averaged over all traits: a large ρ_S (respectively, ρ_M) means that there are many large correlations (of any sign) within matrix **S** (respectively, **M**). Similarly with large n, the distribution of the eigenvalues of $\mathbf{S}\cdot\mathbf{M}$, for any random draw of **S** and **M**, converges to an asymptotic distribution with known moments such that $CV(\lambda)^2 = n(\rho_S^2 + \rho_M^2)$ (online Appendix 2). We can then directly obtain an expression of n_e (from eq. 3) in terms of phenotypic correlations:

$$n_e = \frac{n}{1 + n(\rho_S^2 + \rho_M^2)} \xrightarrow{n\to\infty} \frac{1}{\rho_S^2 + \rho_M^2}. \qquad (8)$$

Equation (8) shows that with random selective and/or mutational correlations between traits, trait heterogeneity $CV(\lambda)$ increases with the number of traits n. This effect drastically reduces n_e and hence the cost of complexity: as n increases indefinitely, n_e reaches a plateau $1/\rho_S^2 + \rho_M^2$ that depends only on the strength of phenotypic correlations and can be very small. This behaviour is consistent with the observation that empirical distributions of s are typically more asymmetric than the Gaussian. This suggests that n_e is typically small although n is expected to be very large in most species.

Simulations and Scaling Issues

Simulations were run using the software R (Ihaka and Robert 1996) to jointly check the displaced gamma approximation for $f(s)$ in equation (5), and the approximation for n_e based on random matrix theory in equation (8). Mutational and selective covariance matrices (**S** and **M**) were randomly drawn as Wishart matrices with fixed correlation strength ρ_S and ρ_M, and scaled to obtain a given value of $E(s) = -\frac{1}{2}\mathrm{Tr}(\mathbf{S}\cdot\mathbf{M}) = -n\bar{\lambda}/2$, where Tr(.) denotes matrix trace. The phenotypic distance to the optimum $\mathbf{z}_o$ was then drawn as a vector of n independent Gaussian deviates $n(0,0.1)$, and scaled to obtain a given value of $s_o = \frac{1}{2}\mathbf{z}_o^t\mathbf{S}\mathbf{z}_o$. Each mutant phenotype was then drawn from a multivariate Gaussian with mean zero and covariance **M** and the corresponding s was computed following the exact distribution in equation (A1) of Appendix 1 (available online).

Figure 1 shows how the approximate distribution matches simulations of the exact distribution of s when mutational and selective covariances (**M** and **S**) are drawn randomly. The approximation fits the exact distribution very well when only deleterious effects are considered ($s_o = 0$, Fig. 1a for $n = 40$ traits). The shape and scale of the gamma approximation on Figure 1 were computed using the asymptotic result from random matrix theory: $CV(\lambda)^2 = n(\rho_S^2 + \rho_M^2)$. This approximation is almost as accurate as if using the exact $CV(\lambda)$ computed from the eigenvalues of simulated matrices. The left top panels in Figure 1a also show the distribution of phenotypic correlations within the simulated Wishart matrices together with the predicted asymptotic distribution from equation (A5). The prediction remains accurate even with a limited number of traits (e.g., $n = 15$, not shown). Finally, Figure 1 shows that stronger correlations can considerably reduce n_e, resulting in more skewed $f(s)$ (a gamma

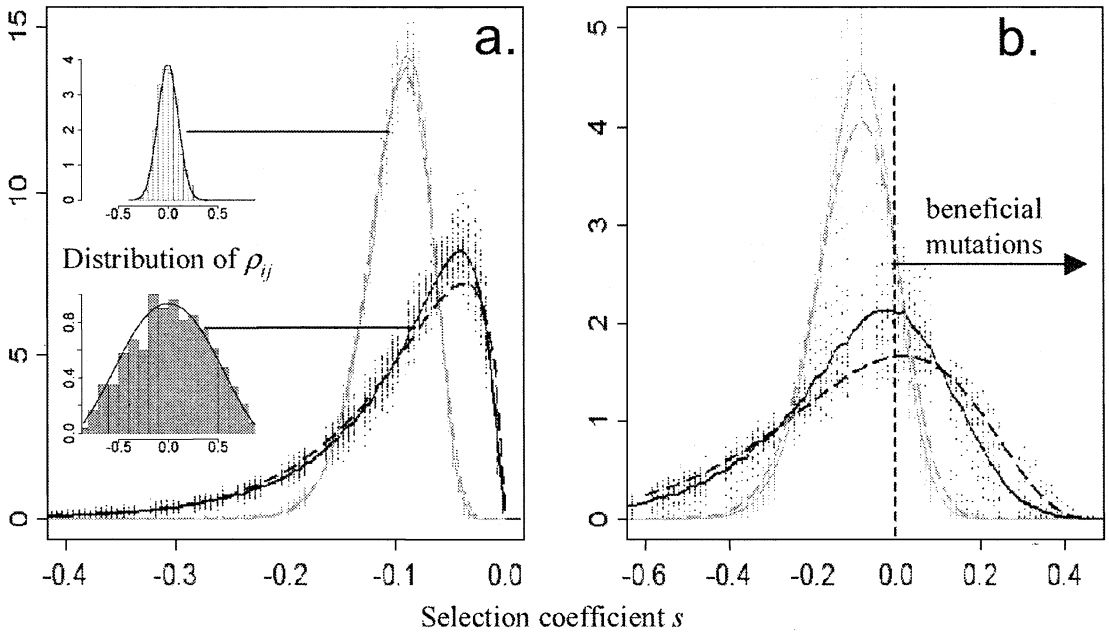

FIG. 1. Distribution of s with random phenotypic covariances and displaced-gamma approximation. Distribution of mutation fitness effect with $n = 40$ traits and the initial genotype perfectly adapted (a; $s_o = 0$) or away from the optimum (b; $s_o = 0.5$; i.e., the fitness of the initial genotype is 50% that of the best genotype). Dots show the density of s from 50 simulations (each with 20,000 mutants and independent draws of **S** and **M** as 40×40 Wishart matrices, see A Model of Random Phenotypic Correlations). Plain lines show the mean over these 50 simulated densities. In each panel, $f(s)$ is illustrated with strong (black) or weak (gray) phenotypic correlations, whose distributions are illustrated with histograms. Strong and weak phenotypic correlations are $\rho_S = \rho_M = 0.38$ (corresponding to $n_e = 3.2$) or 0.1 (corresponding to $n_e = 22$), respectively. Dashed lines show the displaced gamma approximation in equation (5) with n_e values computed as in equation (8), $n_e = n/(1 + n(\rho_S^2 + \rho_M^2))$.

with smaller shape), at the optimum ($s_o = 0$, Fig. 1a) or away from it ($s_o = 0.5$, Fig. 1b).

When $s_o > 0$ (e.g., $s_o = 0.5$ in Fig. 1b.) the displaced gamma approximation for $f(s)$ is reasonably accurate, although it tends to overestimate the proportion of advantageous mutations. The discrepancy is almost entirely due to our approximation for the variance of s at a given fitness distance to the optimum s_o (eq. A4 in online Appendix 1). Using the exact expression for this variance (eq. A3 in online Appendix 1) results in a very accurate displaced gamma approximation. However, this more accurate approximation involves some nonestimable quantities and was thus discarded here.

TESTING THE MODEL WITH EMPIRICAL F(S)

Survey of the Moments of Empirical f(s) across Species

Principle of the approach

To test our predictions on the effect of n on the moments of s (eq. 2), we compared the moments of empirical distributions of mutational fitness effects across species ranging from viruses to higher plants using the number of protein genes per haploid genome as a surrogate measure of complexity. We chose this measure mainly because it is the only one available for all the species considered in our survey, and to avoid any ranking of species according to an a priori qualitative ''complexity.'' However, we do not equate genes and traits in this approach. By definition, a single mutation affects only one gene sequence but may have many pleiotropic effects on other gene products, so that a simple proportional relationship between gene number and phenotypic complexity is likely to be rather unrealistic. Instead, we assume that, at the phylogenetic scale of our comparisons, gene number is positively correlated with the number of adaptation traits, because gene number is expected to determine the number of gene products and their interactions, and because the evolution of new functions is currently sought to occur by the addition of new genes (Vassilieva et al. 2000). This quantitative measure also provides an explicit way to describe the (dis-)similarity between species more precisely and objectively than some a priori attributed rank of complexity between species. We therefore test our model only by assuming some ordered relationship: the larger the gene number, the larger the number of traits. Such a relationship may not hold at a finer scale of comparison (e.g., the plant *Arabidopsis thaliana* has more genes than the fruit fly *D. melanogaster*, but which has the largest number of traits under selection?). However, it does not seem too unrealistic when comparing viruses, microbes and higher organisms.

Estimates of $E(s)$ were drawn from the literature on mutation accumulation (MA), discarding cases with a large proportion of beneficial mutations, suggesting that the initial genotype was far from the optimum, in which case our Gaussian fitness function may not apply. Our survey updates the review by Bataillon (2000) including recent studies on *Drosophila, Caenorhabditis,* vesicular stomatitis virus (VSV), *Saccharomyces cerevisiae* and the fungus *Cryptococcus neoformans,* to a total of 33 $E(s)$ estimates in eight taxa. Esti-

TABLE 1. Gene count and average effect of mutations from mutation accumulation (MA) experiments. VSV, vesicular stomatitis virus; spont, accumulation of spontaneous mutation; mmr, MA on mismatch repair deficient strains; ems, MA using EMS mutagenesis. II or III refer to mutations on the second or third chromosomes in *Drosophila*. The fitness traits measured were: relative viability in competition (viab), lifetime reproductive success (LRS), intrinsic growth rate (r) or relative growth rate in competition (RGR). GC is haploid gene count. $E(s)$ is the average haploid or homozygous effects of mutations.

Species	Method	Trait	GC	$E(s)$	Reference
VSV	spont	RGR	5	−0.0022	Elena and Moya 1999
VSV	spont	RGR	5	−0.0024	Elena and Moya 1999
VSV	spont	RGR	5	−0.0022	Elena and Moya 1999
Escherichia coli	spont	*r*	4970	−0.034	Loewe et al. 2003
E. coli	spont	*r*	4970	−0.012	Bataillon 2000
Saccharomyces cerevisiae[1]	mmr	*r*	5855	−0.075	Zeyl and DeVisser 2001
Cryptococcus neoformans[2]	spont	*r*	6475	−0.045	Xu 2004
Drosophila melanogaster	spont	viab	16,130	−0.16	Bataillon 2000
D. melanogaster	spont (II)	viab	16,130	−0.11	Bataillon 2000
D. melanogaster[3]	spont (II)	viab	16,130	−0.03	Bataillon 2000
D. melanogaster[3]	spont (II)	viab	16,130	−0.03	Bataillon 2000
D. melanogaster	spont (II)	RGR	16,130	−0.1	Avila and Garcia-Dorado 2002
D. melanogaster	spont (II)	viab	16,130	−0.08	Chavarrias et al. 2001
D. melanogaster[4]	spont (III)	viab	16,130	−0.1	Charlesworth et al. 2004
D. melanogaster[5]	ems	viab	16,130	−0.11	Keightley and Ohnishi 1998
Caenorhabditis elegans	spont	*r*	21,357	−0.1	Bataillon 2000
C. elegans	spont	*r*	21,357	−0.2	Bataillon 2000
C. elegans	mmr	*r*	21,357	−0.413	Estes et al. 2004
C. elegans	ems	*r*	21,357	−0.15	Keightley et al. 2000
C. elegans	spont	*r*	21,357	−0.364	Baer et al. 2005
C. elegans	spont	*r*	21,357	−0.25	Baer et al. 2005
C. briggsae	spont	*r*	<21,357	−0.1	Baer et al. 2005
C. briggsae	spont	*r*	<21,357	−0.198	Baer et al. 2005
Arabidopsis thaliana	spont	LRS	28,159	−0.23	Bataillon 2000
Triticum durum	spont	LRS	40,000	−0.2	Bataillon 2000

[1] M grande lines in Zeyl and DeVisser (2001), that is, keeping only ''petite'' mutations in mmr strain. The normal ''F'' strain produced only a single ''grande'' mutant and is not reported here. Only heterozygous effect $E(hs)$ are reported in Zeyl and DeVisser (2001). Homozygous effects in the table are corrected using $h = 0.2$ based on dominance estimates for point mutations in *S. cerevisiae* (Korona 2004).
[2] Average effect in optimal environment (YEPD/37°) from all MA lines in table 2 of Xu (2004).
[3] Mukai-Ohnishi studies.
[4] Average $E(s)$ for all nonlethal mutations in table 4 of Charlesworth et al. (2004).
[5] Maximum likelihood estimate in table 4 of Keightley and Ohnishi (1998).

mation of higher moments of s is very difficult using MA data (Lynch et al. 1999; Keightley 2004), so only few estimates of $V(s)$ and $\mu_3(s)$ are available. To obtain them, we surveyed empirical studies that directly measured the distribution of fitness among lines carrying a single mutation, so that the observed distribution of mutant fitnesses directly gives estimates of the first moments of s. We found nine estimates from five taxa.

Gene number

For most species, we used the number of open reading frames (available on the Kyoto Encyclopedia of Genes and Genomes [KEGG] website http://www.genome.jp). The value used for wheat (*Triticum durum*) was a recent estimate from the rice genome (40,000; Bennetzen et al. 2004) based on strong similarities between cereal genomes (Ware and Stein 2003). The number of protein genes in *C. briggsae* was considered equal to that of *C. elegans* for the statistical analysis (Stein et al. 2003). The number of protein genes in VSV is five (Sanjuan et al. 2004a).

Empirical moments of s

Estimates of $E(s)$ in Table 1 were obtained by surveying mutation accumulation experiments, using either Bateman-Mukai (BM) estimates or maximum likelihood estimates when the latter were significantly better (i.e., in Keightley and Ohnishi 1998; Vassilieva et al. 2000). We discarded two recent studies on yeast (Joseph and Hall 2004) and *Arabidopsis* (Shaw et al. 2002) in which a very large proportion of mutations were beneficial, suggesting that the initial genotype was far from the optimum. $E(s)$ values in Table 2 are direct measures from single mutation effects, except in two cases. For *C. elegans* we used the corrected BM estimate given in Vassilieva et al. (2000). For transposable element (TE) single inserts on chromosomes II and III of *Drosophila* (Lyman et al. 1996), $E(s)$ is biased by a direct TE effect and not given in this study. Therefore, we used the per insert viability effect of third chromosome TE insertions for chromosome III given by Mackay et al. (1992). For chromosome II, we used the average $E(s)$ of all second chromosome viability effects in Table 1. This average estimate was not, of course, included in the statistical analyses of $E(s)$. We did not include the Mukai et al. (1972) and Ohnishi (1977) results for the computation of this average $E(s)$, because their validity has been questioned (Garcia-Dorado et al. 1999), but we did include them in the statistical analysis of $E(s)$ estimates.

We surveyed estimates of higher moments of s from single mutations and from one study reporting a precise estimate (i.e., with limited confidence interval) of CV(s) obtained by maximum-likelihood analysis of MA data in *C. elegans* (Vas-

TABLE 2. Empirical distributions of s and empirical estimates of first moments of $f(s)$. In all these studies, all moments of s (haploid or homozygous effects) are estimated directly using single mutation lines except in *Caenorhabditis elegans* (see Empirical Moments of s). Abbreviations as in Table 1 except subst, single point substitutions; TE, single transposable element insertion. μ_2^*, scaled variance (i.e., squared coefficient of variation) $CV(s)^2 = V(s)/E(s)^2$; μ_3^*, scaled third moment $\mu_3(s)/E(s)^3$.

Species	Method	Trait	GC	$E(s)^2$	$CV(s)^2$	μ_3^*	$\hat{n}_e = 2/\mu_2^*$	Reference
VSV[1]	subst	RGR	5	−0.139	1.87*	4.58	1.07	Sanjuan et al. 2004a
Escherichia coli	TE	*r*	4970	−0.0275	9.55	259	0.21	Elena et al. 1998
Saccharomyces cereviae[1]	spont	*r*	5855	−0.109	2.07	8.59	0.96	Wloch et al. 2001
S. cereviae[1]	TE	*r*	5855	−0.05	2.25	4.59	0.89	Thatcher et al. 1998
S. cereviae[1]	ems	*r*	5855	−0.171	1.80	3.69	1.11	Wloch et al. 2001
S. cereviae[1]	mmr	*r*	5855	−0.183	1.12	1.64	1.78	Wloch et al. 2001
Drosophila melanogaster	TE II	viab	16,130	−0.098	0.79	0.54	2.23	Lyman et al. 1996
D. melanogaster	TE III	viab	16,130	−0.122	0.90	1.03	2.52	Lyman et al. 1996
Caenorhabditis elegans[2]	spont	*r*	21,537	−0.12	0.77	?	2.58	Vassilieva et al. 2000

[1] Moments of s directly computed from fitness effect values in supporting information of Sanjuan et al. (2004a), table 1 of Thatcher et al. (1998), and provided by Wloch et al. (2001).
[2] Maximum likelihood estimate of μ_2^* and correspondingly corrected BM estimate of $E(s)$ given in Vassilieva et al. (2000).

silieva et al. 2000). Studies on *E. coli* (Elena et al. 1998), *S. cerevisiae* (Thatcher et al. 1998) and *D. melanogaster* (Lyman et al. 1996) used TE insertions to generate single mutations. Note that results on *Drosophila* are based on the viability effects of either second or third chromosome TE inserts (as indicated in Table 2), instead of the whole genome. In another study on *S. cerevisiae*, Wloch et al. (2001) used tetrad analysis to isolate single spontaneous or induced mutation events and measure their fitness effect in the haploid stage. In the VSV (Sanjuan et al. 2004a), single nucleotide substitutions are produced by site-directed mutagenesis. No or very few beneficial mutations were detected in these studies, except in VSV with 4% of advantageous mutations (Sanjuan et al. 2004a). The initial genotype was therefore considered well adapted to the laboratory environment ($s_o \ll E(s)$) so that equation (2) applies.

Statistical analyses

We tested our predictions on $E(s)$ using a linear model accounting for gene number and four other potentially confounding variables that could covary with gene number and produce false positive correlations. First, the measure of s might be biased between microbes and higher organisms because a more integrative measure of fitness is used in the former (e.g., growth rate over many generations) instead of fitness components (e.g., viability) over one generation in the latter. The level of integration of the fitness measure was included as an ordered variable with values: 4, growth rate in competition; 3, intrinsic growth rate; 2, lifetime reproductive success; and 1, viability. Second, MA experiments in microbes may underestimate $E(s)$ because selection within sublines is more likely when several generations occur between population bottlenecks (i.e., severely deleterious mutations may not be detected in MA on microbes; Kibota and Lynch 1996). Therefore, we included a factor discriminating microbes versus nonmicrobes to avoid detecting a correlation between $E(s)$ and gene number that would in fact reflect a bias in estimates between lower and higher organisms. Third, the type of mutation was included as a factor: spontaneous versus TE versus point mutations, the latter referring to experiments based on single nucleotide substitutions, or mutagenesis by EMS and mismatch repair deficiency, which are known to cause mainly point mutations (Wloch et al. 2001). Note that spontaneous mutations, as accumulated in standard MA experiments, are a mix of different types of mutations including TE and point mutations. Fourth, the type of estimate and method of mutation accumulation, that is, MA (Table 1) versus direct (Table 2) estimates of $E(s)$, was included as a factor. This factor discriminates between measures based on an unknown number of mutational events (MA) versus a single mutation per line (direct). The full model including pairwise interactions was simplified backward to isolate the significant factors. Regarding our predictions on μ_2^* and μ_3^*, we only tested for a correlation to gene number due to the limited number of estimates available. Finally, we did not correct for phylogenetic independence given the phylogenetic scale of our comparisons. Similarly, we did not pool different estimates for the same species, which would artificially mask the quite large within-species variation of the estimates. However, pooling estimates per species does not qualitatively alter our conclusions below.

The Effect of n on Empirical Moments of Deleterious Mutation Effects

As predicted from equation (1) $E(s)$ (33 estimates from Tables 1 and 2) increases with gene number, our surrogate measure of n ($R^2 = 0.33$, $F_{1,31} = 16.1$, $P = 0.0004$, Fig. 2), and the trend remains significant when discarding viruses ($R^2 = 0.27$, $F_{1,27} = 10$, $P = 0.0035$) or among only eukaryotes ($R^2 = 0.18$, $F_{1,24} = 5.2$, $P = 0.031$). It has been suggested that, for *D. melanogaster*, the results of Mukai et al. (1972) and Ohnishi (1977) may be less reliable than more recent ones (Garcia-Dorado et al. 1999). When removing these estimates, the correlation is stronger ($R^2 = 0.38$, $F_{1,29} = 18$, $P = 0.0002$). Finally, in the VSV, direct and BM estimates are very different, and the most reliable measure is probably the direct estimate (S. F. Elena, pers. comm.). Including only this measure for the VSV in the analysis does not remove the global trend of an increase in $E(s)$ with gene number ($R^2 = 0.24$, $F_{1,28} = 8.8$, $P = 0.006$).

The effect of gene number also remains significant ($F_{1,30} = 24.7$, $P < 0.0001$) when including potentially confounding factors (see Statistical Analyses). Among these, only the method to obtain mutants has a significant effect: point mu-

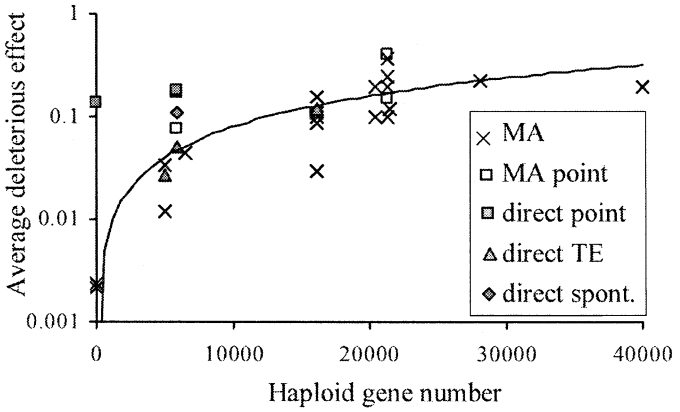

FIG. 2. Variation of the average deleterious effect of mutations with gene number. Estimates of $E(s)$ (y-axis, log-scale) are either indirect estimates from mutation accumulation (MA) experiments (MA, Table 1) or direct estimates from mutagenesis experiments (direct, Table 2). The type of mutation is indicated on the graph as: spont (spontaneous), TE (transposons), point (point mutations; i.e., ems, mmr, and subst in Tables 1 and 2). The line indicates the regression with intercept set to zero on all 33 estimates. $E(s)$ increases by about two orders of magnitude from viruses to higher organisms.

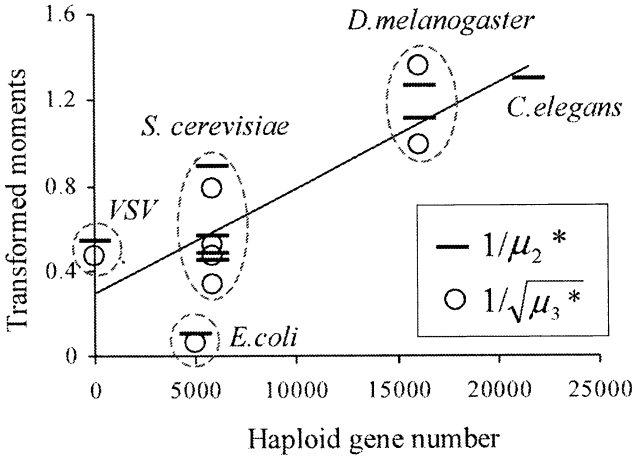

FIG. 3. Variation of second and third moments of $f(s)$ with gene number and of the inverse of the scaled second moment, $1/\mu_2^*$ (dashes), and the square root of the scaled third moment $1/\sqrt{\mu_2^*}$ (circles) with gene count. These measures are expected to correlate with n in equation (2). μ_2^*, μ_3^*, and gene count are from Table 2. The line shows the regression of $1/\mu_2^*$ on gene number. The second and third moments of $f(s)$ ($\mu_2^*$and μ_3^*) tend to decrease with gene number.

tations tend to have more deleterious effects than spontaneous mutations and transposable element insertions ($F_{1,30} = 9.5$, $P = 0.004$). However, point mutations were often obtained by methods that artificially increase the mutation rate (using EMS mutagenesis or mismatch repair deficient strains) so that the effect detected might be due to increased number of mutation per line (through, e.g., negative epistatic interactions), rather than to the molecular nature of point mutations. Our model also predicts that $E(s)$ should not depend on the level of adaptation of the initial genotype (whenever the Gaussian fitness approximation is still valid; i.e., not too far from the optimum). This is consistent with the estimates of $E(s)$ in the MA experiment on VSV (Table 1), which are very similar for three initial genotypes that differed in fitness (0.8, 1, 2.5) (Elena and Moya 1999). The second prediction (eq. (2) was that both $1/\mu_2^*$ and $1/\sqrt{\mu_3^*}$ should increase linearly with n. Although limited, data from Table 2 indicate that both these quantities positively correlate with gene number: Pearson's $\rho = 0.84$, $n = 9$, $P = 0.0043$ and $\rho = 0.80$, $n = 8$, $P = 0.018$, respectively (see Fig. 3). The VSV is clearly an outlier in the dataset, as it has much larger $1/\mu_2^*$ and $1/\sqrt{\mu_3^*}$ values than would be expected from its gene number. This might be due to the fact that $E(s)$, which scales both μ_2^* and μ_3^* may have been overestimated in the VSV direct measure, as could be suggested by the fact that the $E(s)$ estimate is much larger in the direct measure than in the MA experiments (see Fig. 2).

Shape of Empirical Distributions and Estimates of n_e

We now turn to testing whether our approximation for $f(s)$ in terms of a gamma is consistent with the available empirical data. If $f(s)$ is gamma distributed, we expect a quadratic relationship between the scaled second and third moments such that $\mu_3^* = 2\mu_2^{*2}$. The values of μ_2^* and μ_3^* (Table 2) exhibit such a relationship across species ($\mu_3^* \approx 3.07\ \mu_2^{*2}$; 95% bootstrap slope CI 1.01–3.09; $R^2 = 0.99$; $P < 0.0001$, Fig. 4). This strong relationship indicates that empirical $f(s)$ belong to a gamma-like distribution family. Under such a gamma approximation, the shape of the empirical $f(s)$ for deleterious effects (i.e., when $s_o = 0$), is simply $\frac{1}{2}n_e$ (see above), where n_e can be estimated as $n_e = 2E(s)^2/V(s) = 2/\mu_2^*$ (eq. 4). The resulting n_e estimates (reported in Table 2) increase with increasing gene number (same P-values as $1/\mu_2^*$, above), as predicted among species of increasing complexity. Overall, with the exception of the VSV, increasing gene number (complexity) results qualitatively in gamma $f(s)$ of increasing

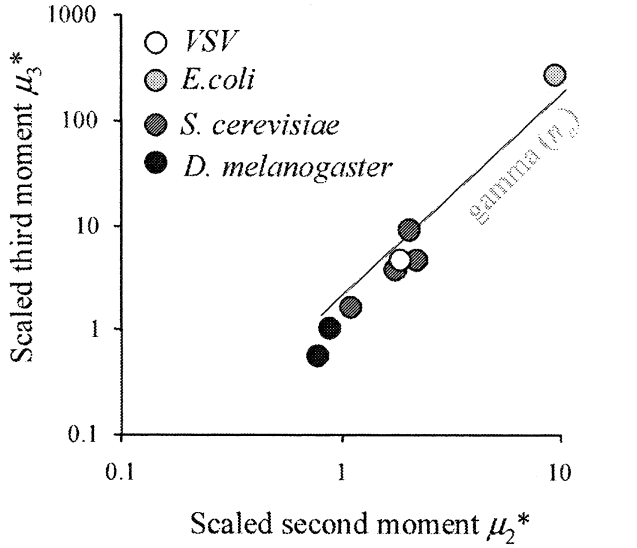

FIG. 4. Variation of the second and third scaled moments of $f(s)$ (μ_2^* and μ_3^* respectively, both in log-scale, from Table 2) across species. The line indicates the predicted relationship if s follows a gamma distribution with shape $n_e/2$ and $s_o = 0$, where n_e is $2/\mu_2^*$. Second and third moments show a strong quadratic relationship consistent with a gamma-like distribution.

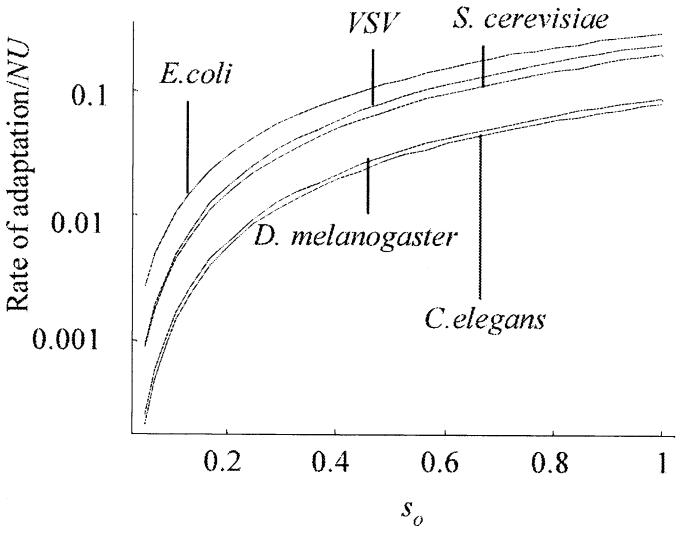

FIG. 5. Predicted rates of adaptation across species. Rates of adaptation (scaled to *NU*) for different species for different distance of the initial genotype relative to the optimum (s_o, x-axis). They are computed from equation (7) using $E(s)$ and n_e estimates in Table 2. For yeast and *Drosophila*, the mean over 4 (2, respectively) estimates is indicated. Predicted rates of adaptation decrease by an order of magnitude from bacteria to fruit flies or nematodes.

shape parameter ($\frac{1}{2}n_e$, see Fig. 4). In the following section we use this finding to quantify the cost of complexity by estimating how the effects of phenotypic complexity that we observe in our survey may influence the rate of adaptation across species.

Rates of Adaptation and Cost of Complexity across Species

Equation (7) can be used to evaluate approximately how variation in $f(s)$ across species may influence their rates of adaptation for a given fitness distance to the optimum s_o and a given flow of mutations $N\mu$. Rates of adaptation (scaled to $N\mu$), based on $E(s)$ and n_e estimates in Table 2 for the different species, are shown in Figure 5 for s_o varying from 0.05 to 1. First, Figure 5 shows that the increase in n_e values from *E. coli* to higher organisms indeed translates into a reduction in rates of adaptation, confirming the existence of a cost of complexity. Second, Figure 5 also shows that only a modest difference in the rate of adaptation is predicted among species (as expected due to the small variation in n_e), with an increase of only an order of magnitude from *E. coli* to *C. elegans*. In the next section we discuss how phenotypic correlations may explain such a reduced cost of complexity.

Why Such a Low Cost? The Effect of Phenotypic Correlations

A main conclusion of our survey is that the effective number of dimensions n_e increases with complexity as predicted by our model, but varies only by an order of magnitude from bacteria to fruit flies, and remains very small (0.2–2.5). This finding is surprising because our model predicts that $n_e = n/[1 + \mathrm{CV}(\lambda)^2]$ (eq. 3) should be proportional to the actual number of traits under selection, which, intuitively, should vary by more than an order of magnitude across the species considered. One possible explanation may be that trait heterogeneity $\mathrm{CV}(\lambda)$ increases with the number of traits n, leading to a diminishing return of n_e on n. We showed that such a phenomenon occurs in the simplest case where mutational and selective covariance matrices, **S** and **M**, are drawn randomly (see eq. 8). For example, n_e values as small as 2.5 (as observed for *C. elegans* or *D. melanogaster*) can be obtained for any large number of traits with correlation strength $\rho_S^2 = \rho_M^2 \approx 0.2$. Note, however, that the very small n_e value estimated for *E. coli* ($n_e = 0.21$) cannot be explained by this model, even assuming very strong correlations (in eq. 8, n_e cannot be less than 0.5 for large n). Nevertheless, the model predicts that large variation in the total number of traits n may translate into limited differences in n_e values, a pattern fully consistent with our observations. It may be argued the other way around that all species have a very large number of traits but that higher organisms tend to have weaker phenotypic correlations, resulting in a higher n_e (see eq. 8, limit $n \to \infty$). This hypothesis seems less parsimonious but remains to be tested.

DISCUSSION

Predicting how the distribution of mutation fitness effects should vary across species is a key issue in evolutionary biology but has received little attention so far. In particular, no study has been proposed to estimate and explain variation in $f(s)$ across species. In this paper, we proposed a model of mutation fitness effects, that extends Fisher's (1930) geometric approach to take into account (co)variation between traits in the effect of both mutation and selection (into a multivariate Gaussian framework). The shape of the predicted distribution depends on the number of phenotypic traits under selection (n). We predicted how n should influence the three first moments of $f(s)$, and we tested these predictions using gene number as a surrogate measure of phenotypic complexity n. A comparison of data collected from several species confirmed all three predicted trends: as n increases, the average s increases (Fig. 2), while the second and third moments of $f(s)$ decrease (Fig. 3). These results show that phenotypic complexity has a strong influence on the distribution of s. Our survey is based on a limited amount of data. However, we find that different experiments for the same species are consistent with each other except for VSV, where the direct measure of $E(s)$ differs markedly with BM estimates. More studies are required to obtain consistent estimates of mutation fitness effects in viruses, but it is probable that the direct estimate is more reliable (S. F. Elena, pers. comm.), in which case VSV indeed represents a striking exception in the observed correlation. Alternatively, the quantitative genetic framework used here might also not apply to very small viral genomes. Next, we show that these empirical $f(s)$ are gamma-like (Fig. 4), as expected under our model (Fig. 1), and that their shape measures an effective number of traits, n_e. Although rather indirect and qualitative, the results of this review validate three clear predictions of the model, each of which could have been invalidated by the data. Therefore, this review provides a first test of fitness landscape models, which calls for further confrontation with empirical data.

In our model, this effective number of traits is reduced

relative to n when trait heterogeneity increases. We use this finding to measure how variation in $f(s)$ across species translates into variation in rates of adaptation; that is, to predict the resulting cost of complexity. We predict only modest decrease of this rate from lower to higher organisms in the range of species for which empirical $f(s)$ are available (Fig. 5). This moderate decrease confirms the existence of a cost of complexity but minimizes its quantitative importance. Finally, we use random matrix theory to determine how $f(s)$ should vary with the number of traits if mutational and selective covariance matrices are drawn randomly. We find that under this simple model, trait heterogeneity scales with the number of traits such that even organisms with a very large number of traits exhibit a skewed and gamma-like $f(s)$ and pay a very low cost to complexity.

Comparison with Previous Theoretical Results

In a recent paper, Waxman and Welch (2005) proposed a model of mutation that accounts for selective interactions between traits (i.e., an arbitrary matrix **S** in our framework). We briefly compare the two approaches. Waxman and Welch's approach allows for the effect of biased mutation effects on phenotypes, which is neglected in ours. Mutational bias may be included in our analysis, but the results rely on some extra parameters which may be difficult to measure empirically. Our prediction of $f(s)$ based on deleterious mutations and a measure of s_o also neglects the correlation between the z_i and λ_i across traits i (see online Appendix 1), which are not neglected in Waxman and Welch (2005). This was done to propose results in terms of measurable quantities, but analytic arguments and simulations (G. Martin and T. Lenormand, unpubl. data) suggest that this correlation may have quantitatively limited impact on our results.

In addition to selective correlations, our model also accounts for mutational correlations, which are ignored in Waxman and Welch (2005). However, we show in Appendix 1 (available online) that our more general model can be reduced to a simpler model with spherically symmetrical mutation and heterogeneous selection across traits (with selection effects equal to the λ_i, the eigenvalues of $\mathbf{S}\cdot\mathbf{M}$), as is assumed in Waxman and Welch (2005). We therefore believe that their results should hold in our context.

The link between the two models and results can be made simply by considering the simplest situation common to both models: initial genotype at the optimum, unbiased, and uncorrelated effects of mutation on phenotypes. The most remarkable fact arising from this comparison is that both approaches yield very similar expressions for the effective number of dimensions, n_e. When the initial phenotype is at the optimum ($\mathbf{z}_o = 0$), both predict the same reduction of n_e relative to n, when traits are heterogeneous: $n_e = n/(1 + \mathrm{CV}(\lambda)^2)$. This can be seen by comparing equation 27 of Waxman and Welch (2005) and equation (3) of the present paper: $\mathrm{CV}(\lambda)^2 = f_{z,\sigma}$ in Waxman and Welch's notations, when $\mathbf{z} = 0$. When not at the optimum, our definition of n_e differs from that of Waxman and Welch (2005): ours is based on $f(s)$ alone, whereas that of Waxman and Welch (2005) is based on the rate of adaptation, which depends on both $f(s)$ and the distance to the optimum. Overall, these similarities in the predicted effect of phenotypic correlations on n_e give strong support to the idea that they have a very large impact on $f(s)$ and on rates of adaptation in general.

The Structure of Phenotypic Interactions

We find that random and independent mutational and selective covariance matrices (**M** and **S**) generate distributions of mutation fitness effects that are consistent with the available empirical $f(s)$. However, this agreement does not rule out that mutational and selective covariance matrices be in fact nonrandom and/or interdependent. We used random matrices to exhibit a simple situation in which the observed pattern is predicted. However, we believe that random matrix theory, which has proven fruitful in the analysis of complex systems in physics and finance (Forrester et al. 2003), is a promising avenue to analyze models of phenotypic interactions with many traits, or selective interactions with many genes (e.g., distribution of epistatic interactions Bonhoeffer et al. 2004; Sanjuan et al. 2004b).

Why and when should we expect mutational and selective covariance matrices to be nonrandom and mutually dependent? We can discuss two extreme situations that reflect the range of possibilities. In the first situation, $\mathbf{M} \propto \mathbf{S}^{-1}$, (where $\mathbf{S}^{-1}$ is the inverse of **S**) such that the traits under the strongest selective pressure exhibit the lowest mutational variance. Consequently, $\mathbf{S}\cdot\mathbf{M} \propto \mathbf{I}$, and all λ_i are equal, so that $n_e = n$, even if traits are very heterogeneous within both **M** and **S**. This situation would be theoretically expected under strong canalization (Rice 1998). In the opposite situation, $\mathbf{M} \propto \mathbf{S}$, such that traits under strong selection exhibit the highest mutational variance. Consequently, $\mathbf{S}\cdot\mathbf{M} \propto \mathbf{M}^2$, so that $\mathrm{CV}(\lambda_i)$ is maximized and $n_e \ll n$. This situation, on the contrary, would be expected under decanalization (Rice 1998). The moderate increase in n_e among species of presumably large variation in phenotypic complexity (or at least gene number) could thus be explained if more complex organisms tend to show less canalization. Obviously, more work is needed to determine how selection can shape mutational and selective covariance matrices.

Small n_e Values

Even assuming strong correlations, the very small n_e values obtained for *E. coli* ($n_e < 0.5$) cannot be explained by our model. In addition, in *E. coli*, unlike in the VSV, the estimates of $E(s)$ from two independent MA studies and from a direct measure are consistent (Fig. 1). Therefore, an overestimation of $\mathrm{CV}(s)$ (hence an underestimation of n_e) due to imprecision in $E(s)$ estimates in this species seems unlikely at first glance. However, the agreement between BM and direct estimates of $E(s)$ in *E. coli* is in fact surprising, because the former are biased upward relative to the latter proportionately to $1 + \mathrm{CV}(s)^2$ (Lynch et al. 1999). Given the large values of $\mathrm{CV}(s)$ reported for this species (Table 2), we would expect direct measures of $E(s)$ to be smaller than the corresponding BM estimates (whereas this effect should be limited in yeast or *Drosophila*, where $\mathrm{CV}(s)$ is much smaller). It is possible that a direct deleterious effect of transposition biased upward the estimates of $E(s)$ in the direct measure based on single TE insertions—such an effect has been reported in a later study

using the same lines (Remold and Lenski 2001). This could then lead to the unexpected agreement between BM and direct estimates of $E(s)$ in *E. coli*. In any case, correcting for such overestimation of $E(s)$ could only lead to an even smaller n_e estimate than the one reported here. More generally, if $CV(s)$ decreases with complexity, as suggested by our survey, the BM estimates of $E(s)$ should particularly overestimate $E(s)$ in lower organisms. As a consequence, the increase of $E(s)$ with gene number should be more radical than the one we report here, based mostly on BM estimates.

A possible explanation to the very small n_e estimate in *E. coli* could be that we considered universal pleiotropy. Modularity may affect the distribution of s (Wagner and Altenberg 1996; Welch and Waxman 2003). However, as noted in the introduction, with both mutational and selective covariances, such ''modularity'' requires having matching blocks in both **S** and **M**, such that mutations in a given module only affect fitness in the same module. Such ''matching blocks'' modularity could explain low values of n_e among species (i.e., not only in *E. coli*), and lead to a small value in this species. For example, with m exactly equivalent modules, the three first moments of s are obtained by simply replacing n by n/m in equation (1) , which can help to explain the very low n_e values. However, the limited cost of complexity that we predict across species in our survey is due to the limited variation in n_e across species rather than to n_e values themselves. To explain such limited variation with modularity, a very specific relationship between the number of modules and the total number of traits would be required. A more general model including modules of variable sizes and trait heterogeneity within modules would be necessary to compare their relative impact on $f(s)$. Such a model could be developed by considering a set of covariance matrices describing the mutational phenotypic effects of each module. Weighting each matrix by the probability of a mutation in the corresponding module would yield the net effect of all modules and could be summarized with a single matrix **M**.

Evolution of Complex Organisms

The relationship between gene number and phenotypic complexity might be rather weak, beyond coarse phylogenetic divisions (for discussion see Otto and Yong 2002). However, at the phylogenetic scale of our study, gene number provides, at least, an intuitive measure of complexity by ranking viruses, unicellular and multicellular organisms. In any case, beyond our interpretation in terms of phenotypic complexity, our survey shows that gene number is a good predictor of differences in $f(s)$: increasing gene number results in larger average deleterious effects and in approximately gamma $f(s)$ with increasing shape parameter, VSV being an exception (Fig. 4).

Because the term ''complexity'' may have various meanings according to authors, it is important to recall that the definition we refer to here is a number of adaptation traits under selection, as in the Fisher-Orr approach. However, even in this context, the notion of complexity remains somewhat vague and only defined in reference to a measurable quantity; for instance, in reference to rates of adaptation. It may be more straightforward to define it in reference to $f(s)$ because other factors may influence the rate of adaptation either favoring (Orr 2000; Welch and Waxman 2003) or disfavoring adaptation in higher organisms (e.g., longer generation time and smaller population sizes or mutation rates; Lynch and Conery 2003). Complexity in reference to $f(s)$ is our effective number of traits n_e, which may be radically different from complexity as perceived from organismal organization. This idea of an ''effective dimensionality'' (Orr 1998, 2000; Barton and Keightley 2002) has already been put forward; we intended here to provide a formal analysis of the influence of explicit biological assumptions on this quantity. We found that n_e and consequently, predicted rates of adaptation, differ little among species, and we showed that increasing the number of traits can have almost no effect on the rate of adaptation if they are not independent. If mutational and selective covariances are drawn randomly, the outcome is even more extreme: when the number of traits is large, n_e (as both $f(s)$ and rates of adaptation) is determined primarily by phenotypic correlations and tends to a finite limit as the number of trait increases. We therefore expect that $f(s)$ in more ''complex'' organisms should be similar to $f(s)$ in fruit flies or nematodes in our survey. In any case ''complexity,'' as defined by a number of traits under selection (adaptation traits, defined in the introduction), may not pose such an evolutionary paradox as previously suggested.

Conclusions

Our model and analysis is an attempt to predict distributions of mutation fitness effects based on explicit biological hypotheses and to validate it with empirical data. We also intended to show that the Fisher-Orr geometric approach may not be as unrealistic as it is sometimes suggested, provided phenotypic correlations are accounted for. However, important limitations remain (apart from the issue on phenotypic modularity discussed above). First, frequency-dependent or disruptive selection cannot be taken into account. While this may not be a problem to predict $f(s)$ in the context of laboratory studies where each line's fitness is assayed individually, it may limit the generality of predictions on the adaptation of natural populations. Second, predictions far from the optimum may be less robust, as we outlined in the introduction. However, such predictions can still be made in this situation (eq. A3 in online Appendix 1 shows the predicted effect of maladaptation on moments of $f(s)$ in our model). These predictions could be tested, by considering, for example, the effect of stress on the distribution of mutation fitness effects. Finally, we note that it remains difficult to critically test Fisher's model based on empirical data because of the lack of alternative models. One possibility would be to compare rates of adaptation predicted using extreme value theory (Orr 2002) or Fisher's model, but this is beyond the scope of this paper.

Our results suggest that mutation fitness effect distributions have a predictable shape and variation from lower to higher organisms, and that phenotypic landscape models may capture this variation. Furthermore, our approach suggests that distributions of deleterious mutation effects can be used to predict the distribution of beneficial ones for a given environmental change, which is open to further empirical tests.

This opens a wide array of perspectives, as these distributions may be important to a large diversity of questions in evolutionary genetics.

Acknowledgments

We thank P. Jarne, Y. Michalakis, S. F. Elena, D. Waxman, and M. Kirkpatrick. We also thank R. Korona for kindly providing the estimates of selection coefficients in the yeast from Wloch et al. (2001). This work was supported by an Action Concertée Incitative grant from the French Ministry of Research to TL.

Literature Cited

Avila, V., and A. Garcia-Dorado. 2002. The effects of spontaneous mutation on competitive fitness in *Drosophila melanogaster*. J. Evol. Biol. 15:561–566.

Baer, C. F., F. Shaw, C. Steding, M. Baurngartner, A. Hawkins, A. Houppert, N. Mason, M. Reed, K. Sinnonelic, W. Woodard, and M. Lynch. 2005. Comparative evolutionary genetics of spontaneous mutations affecting fitness in rhabditid nematodes. Proc. Natl. Acad. Sci. USA 102:5785–5790.

Barton, N., and L. Partridge. 2000. Limits to natural selection. BioEssays 22:1075–1084.

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

Bataillon, T. 2000. Estimation of spontaneous genome-wide mutation rate parameters: whither beneficial mutations? Heredity 84:497–501.

———. 2003. Shaking the ''deleterious mutations'' dogma? Trends Ecol. Evol. 18:315–317.

Bennetzen, J. L., C. Coleman, R. Y. Liu, J. X. Ma, and W. Ramakrishna. 2004. Consistent over-estimation of gene number in complex plant genomes. Curr. Opin. Plant Biol. 7:732–736.

Bonhoeffer, S., C. Chappey, N. T. Parkin, J. M. Whitcomb, and C. J. Petropoulos. 2004. Evidence for positive epistasis in HIV-1. Science 306:1547–1550.

Burch, C. L., and L. Chao. 1999. Evolution by small steps and rugged landscapes in the RNA virus *phi*6. Genetics 151: 921–927.

Charlesworth, B., and D. Charlesworth. 1998. Some evolutionary consequences of deleterious mutations. Genetica 103:3–19.

Charlesworth, B., H. Borthwick, C. Bartolome, and P. Pignatelli. 2004. Estimates of the genomic mutation rate for detrimental alleles in *Drosophila melanogaster*. Genetics 167:815–826.

Chavarrias, D., C. Lopez-Fanjul, and A. Garcia-Dorado. 2001. The rate of mutation and the homozygous and heterozygous mutational effects for competitive viability: a long-term experiment with *Drosophila melanogaster*. Genetics 158:681–693.

Clarke, B., and W. Arthur. 2000. What constitutes a ''large'' mutational change in phenotype? Evol. Dev. 2:238–240.

Drake, J. W., B. Charlesworth, D. Charlesworth, and J. F. Crow. 1998. Rates of spontaneous mutation. Genetics 148:1667–1686.

Elena, S. F., and R. E. Lenski. 2003. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. Nat. Rev. Genet. 4:457–469.

Elena, S. F., and A. Moya. 1999. Rate of deleterious mutation and the distribution of its effects on fitness in vesicular stomatitis virus. J. Evol. Biol. 12:1078–1088.

Elena, S. F., L. Ekunwe, N. Hajela, S. A. Oden, and R. E. Lenski. 1998. Distribution of fitness effects caused by random insertion mutations in *Escherichia coli*. Genetica 103:349–358.

Estes, S., P. C. Phillips, D. R. Denver, W. K. Thomas, and M. Lynch. 2004. Mutation accumulation in populations of varying size: the distribution of mutational effects for fitness correlates in *Caenorhabditis elegans*. Genetics 166:1269–1279.

Fisher, R. A. 1930. The genetical theory of natural selection. Oxford Univ. Press, Oxford, U.K.

Forrester, P. J., N. C. Snaith, and J. J. M. Verbaarschot. 2003. Developments in random matrix theory. J. Physics A 36: R1–R10.

Garcia-Dorado, A., C. Lopez-Fanjul, and A. Caballero. 1999. Properties of spontaneous mutations affecting quantitative traits. Genet. Res. 74:341–350.

Gerrish, P. J., and R. E. Lenski. 1998. The fate of competing beneficial mutations in an asexual population. Genetica 103: 127–144.

Ihaka, I., and G. Robert. 1996. R: a language for data analysis and graphics. J. Comput. Graphic. Stat. 5:299–314.

Imhof, M., and C. Schlotterer. 2001. Fitness effects of advantageous mutations in evolving *Escherichia coli* populations. Proc. Natl. Acad. Sci. USA 98:1113–1117.

Joseph, S. B., and D. W. Hall. 2004. Spontaneous mutations in diploid *Saccharomyces cerevisiae*: more beneficial than expected. Genetics 168:1817–1825.

Keightley, P. D. 1994. The distribution of mutation effects on viability in *Drosophila melanogaster*. Genetics 138:1315–1322.

———. 2004. Comparing analysis methods for mutation-accumulation data. Genetics 167:551–553.

Keightley, P. D., and A. Eyre-Walker. 2000. Deleterious mutations and the evolution of sex. Science 290:331–333.

Keightley, P. D., and M. Lynch. 2003. Toward a realistic model of mutations affecting fitness. Evolution 57:683–685.

Keightley, P. D., and O. Ohnishi. 1998. EMS-induced polygenic mutation rates for nine quantitative characters in *Drosophila melanogaster*. Genetics 148:753–766.

Keightley, P. D., E. K. Davies, A. D. Peters, and R. G. Shaw. 2000. Properties of ethylmethane sulfonate-induced mutations affecting life-history traits in *Caenorhabditis elegans* and inferences about bivariate distributions of mutation effects. Genetics 156: 143–154.

Kibota, T. T., and M. Lynch. 1996. Estimate of the genomic mutation rate deleterious to overall fitness in *E-coli*. Nature 381: 694–696.

Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. E. Hill, A. Hoang, P. Gibert, and P. Beerli. 2001. The strength of phenotypic selection in natural populations. Am. Nat. 157:245–261.

Korona, R. 2004. Experimental studies of deleterious mutation in *Saccharomyces cerevisiae*. Res. Microbiol. 155:301–310.

Lande, R. 1980. The genetic covariance between characters maintained by pleiotropic mutations. Genetics 94:203–215.

Loewe, L., V. Textor, and S. Scherer. 2003. High deleterious genomic mutation rate in stationary phase of *Escherichia coli*. Science 302:1558–1560.

Lyman, R. F., F. Lawrence, S. V. Nuzhdin, and T. F. C. Mackay. 1996. Effects of single P-element insertions on bristle number and viability in *Drosophila melanogaster*. Genetics 143: 277–292.

Lynch, M., and J. S. Conery. 2003. The origins of genome complexity. Science 302:1401–1404.

Lynch, M., and B. Walsh. 1998. Correlations between characters. Pp. 629–655. Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland, MA.

Lynch, M., J. Blanchard, D. Houle, T. Kibota, S. Schultz, L. Vassilieva, and J. Willis. 1999. Perspective: Spontaneous deleterious mutation. Evolution 53:645–663.

Mackay, T. F. C., R. F. Lyman, and M. S. Jackson. 1992. Effects of P-Element insertions on quantitative traits in *Drosophila melanogaster*. Genetics 130:315–332.

Mathai, A. M., and S. B. Provost. 1992. Quadratic forms in random variables. Marcel Dekker, New York.

Mukai, T., S. I. Chigusa, L. E. Mettler, and J. F. Crow. 1972. Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster*. Genetics 72:335–355.

Ohnishi, O. 1977. Spontaneous and ethyl methanesulfonate-induced mutations controlling viability in *Drosophila melanogaster*. II. Homozygous effects of polygenic mutations. Genetics 87: 529–545.

Orr, H. A. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution 52: 935–949.

———. 2000. Adaptation and the cost of complexity. Evolution 54:13–20.
———. 2001. The ''sizes'' of mutations fixed in phenotypic evolution: a response to Clarke and Arthur. Evol. Dev. 3:121–123.
———. 2002. The population genetics of adaptation: the adaptation of DNA sequences. Evolution 56:1317–1330.
———. 2003. The distribution of fitness effects among beneficial mutations. Genetics 163:1519–1526.
———. 2005a. Theories of adaptation: what they do and don't say. Genetica 123:3–13.
———. 2005b. The genetic theory of adaptation: a brief history. Nat. Rev. Genet. 6:119–127.
Otto, S. P. 2004. Two steps forward, one step back: the pleiotropic effects of favoured alleles. Proc. R. Soc. Lond. B 271:705–714.
Otto, S. P., and P. Yong. 2002. The evolution of gene duplicates. Pp. 451–483. Homology effects. Academic Press, San Diego, CA.
Poon, A., and S. P. Otto. 2000. Compensating for our load of mutations: freezing the meltdown of small populations. Evolution 54:1467–1479.
Remold, S. K., and R. E. Lenski. 2001. Contribution of individual random mutations to genotype-by-environment interactions in *Escherichia coli*. Proc. Natl. Acad. Sci. 98:11388–11393.
Rice, S. H. 1998. The evolution of canalization and the breaking of von Baer's laws: modeling the evolution of development with epistasis. Evolution 52:647–656.
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.
Rozen, D. E., J. de Visser, and P. J. Gerrish. 2002. Fitness effects of fixed beneficial mutations in microbial populations. Curr. Biol. 12:1040–1045.
Sanjuan, R., A. Moya, and S. F. Elena. 2004a. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. Proc. Natl. Acad. Sci. USA 101:8396–8401.
———. 2004b. The contribution of epistasis to the architecture of fitness in an RNA virus. Proc. Natl. Acad. Sci. USA 101: 15376–15379.
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.
Shaw, R. G., F. H. Shaw, and C. Geyer. 2003. What fraction of mutations reduces fitness? A reply to Keightley and Lynch. Evolution 57:686–689.
Stein, L. D., Z. R. Bao, D. Blasiar, T. Blumenthal, M. R. Brent, N. S. Chen, A. Chinwalla, L. Clarke, C. Clee, A. Coghlan, and many others. 2003. The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. Plos Biol. 1: 166.
Thatcher, J. W., J. M. Shaw, and W. J. Dickinson. 1998. Marginal fitness contributions of nonessential genes in yeast. Proc. Natl. Acad. Sci. USA 95:253–257.
Vassilieva, L. L., A. M. Hook, and M. Lynch. 2000. The fitness effects of spontaneous mutations in *Caenorhabditis elegans*. Evolution 54:1234–1246.
Wagner, G. P., and L. Altenberg. 1996. Perspective: Complex adaptations and the evolution of evolvability. Evolution 50: 967–976.
Ware, D., and L. Stein. 2003. Comparison of genes among cereals. Curr. Opin. Plant Biol. 6:121–127.
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.
Wloch, D. M., K. Szafraniec, R. H. Borts, and R. Korona. 2001. Direct estimate of the mutation rate and the distribution of fitness effects in the yeast *Saccharomyces cerevisiae*. Genetics 159: 441–452.
Xu, J. P. 2004. Genotype-environment interactions of spontaneous mutations for vegetative fitness in the human pathogenic fungus *Cryptococcus neoformans*. Genetics 168:1177–1188.
Zeyl, C., and J. DeVisser. 2001. Estimates of the rate and distribution of fitness effects of spontaneous mutation in *Saccharomyces cerevisiae*. Genetics 157:53–61.
Zhang, X. S., and W. G. Hill. 2003. Multivariate stabilizing selection and pleiotropy in the maintenance of quantitative genetic variation. Evolution 57:1761–1775.

Corresponding Editor: P. Phillips