Fisher's microscope and Haldane's ellipse

Article (Unspecified)

vol. 166, no. 4 the american naturalist october 2005

Fisher's Microscope and Haldane's Ellipse

D. Waxman* **and J. J. Welch†**

Centre for the Study of Evolution, School of Life Sciences, University of Sussex, Brighton BN1 9QG, Sussex, United Kingdom

Submitted March 13, 2005; Accepted June 6, 2005; Electronically published August 5, 2005

abstract: Fisher's geometrical model was introduced to study the phenotypic size of mutations contributing to adaptation. However, as pointed out by Haldane, the model involves a simplified picture of the action of natural selection, and this calls into question its generality. In particular, Fisher's model assumes that each trait contributes independently to fitness. Here, we show that Haldane's concerns may be incorporated into Fisher's model solely by allowing the intensity of selection to vary between traits. We further show that this generalization may be achieved by introducing a single, intuitively defined quantity that describes the phenotype prior to adaptation. Comparing the process of adaptation under the original and generalized models, we show that the generalization may bias results toward either larger or smaller mutations. The applicability of Fisher's model is then discussed.

Keywords: geometrical model, beneficial mutations, adaptation.

The classic arguments in favor of micromutationalism—the doctrine that mutations of very small phenotypic effect are the most likely to contribute to adaptation—are given by Fisher (1930). To support this position, Fisher offered an analogy, comparing the effect of a mutation to the "mechanical adaptation of an instrument, such as a microscope." He claimed that "it is sufficiently obvious that any large derangement will have a very small probability of improving the adjustment," while in the case of the smallest possible alterations, "the chance of improvement should be almost exactly half" (Fisher 1930, pp. 37–38). In addition to his verbal analogy, Fisher introduced a mathematical model of natural selection acting on multiple quantitative traits. In the model, each trait has an optimal value, and the fitness of a complete phenotype is jointly determined from the distances of the different traits from their optimal values. Fisher showed that under these assumptions, mutations that affect all traits have a probability of being beneficial that increases rapidly as their size decreases (see below for a definition of size). When the size of mutations approaches 0, their probability of being beneficial achieves the maximal value of 1/2.

Two years later, Haldane (1932) questioned the generality of Fisher's conclusion. He pointed out that Fisher's calculation relies on the assumption that traits act independently to determine fitness. In most real cases of adaptation, however, traits are likely to interact such that a change in a given trait might be beneficial in one phenotypic context and deleterious in another. "For example," Haldane wrote, "an increase in pigmentation in an animal might be disadvantageous unless balanced by an increase in the capacity of its liver for storing vitamin D during sunny weather" (Haldane 1932, p. 175). Haldane's criticism is particularly apposite, since Fisher explicitly set out to examine "complex" adaptations involving many integrated traits. However, Haldane did not explore his suggestion mathematically, and so it remained unclear to what extent Fisher's conclusions were threatened by his suggestion.

A more serious challenge to Fisher's advocacy of micromutationalism came more than 50 years later in the work of Kimura, who did not question the assumptions of Fisher's model or the accuracy of his calculations. Instead, he pointed out that the quantity that is most relevant to evolutionary adaptation is not the probability that a mutation is beneficial but rather the probability that it is both beneficial and achieves fixation (Kimura 1983, pp. 150–156; see also Orr 1998). This is because only those beneficial mutations that achieve permanence in the population can be said to contribute to adaptation. Kimura showed that under Fisher's model, mutations of an intermediate size have the greatest overall probability of being beneficial and fixing. This is because small-effect mutations, though most likely to be beneficial, typically have very small effects on fitness and hence are most likely to be lost via genetic drift. By contrast, large-effect mutations are less susceptible to drift but are less likely to be beneficial.

* E-mail: d.waxman@sussex.ac.uk.

† E-mail: j.j.welch@sussex.ac.uk.

Am. Nat. 2005. Vol. 166, pp. 447–457.

In the last decade, many further studies have used Fisher's model. Some of these studies have continued the investigation of the size of mutations contributing to adaptation (Hartl and Taubes 1998; Orr 1998, 1999; Burch and Chao 1999), while others have treated topics such as drift load (Hartl and Taubes 1996; Peck et al. 1997; Poon and Otto 2000), hybridization (Barton 2001), and evolutionary rates (Orr 2000; Welch and Waxman 2003). Like Kimura, all of these authors have used Fisher's model in its original form.

In this study, we modify Fisher's model to take into account the concerns expressed by Haldane (1932). We then examine the extent to which this modification affects the conclusions of Fisher (1930), Kimura (1983), and, by implication, those of the more recent studies. Let us begin by presenting the model in a general manner that can be applied both to Fisher's original treatment and to our extended form.

The Geometrical Model

Fisher's model treats natural selection as acting on a set of n quantitative (i.e., continuously varying) traits. The relevant phenotype of any individual is geometrically described as a point in an n-dimensional space of traits. Such a point is conveniently represented by an n-dimensional vector of trait values: $\mathbf{z} = (z_1, z_2, \ldots, z_n)$.

Consider a mutation that may alter the value of all n traits. The mutational changes on the n traits are denoted by $\mathbf{r} = (r_1, r_2, \ldots, r_n)$, and the mutated phenotype is given by $\mathbf{z} + \mathbf{r} = (z_1 + r_1, z_2 + r_2, \ldots, z_n + r_n)$. The size of a mutation can now be defined as the magnitude of the mutational change:

$$\|\mathbf{r}\| \equiv \sqrt{r_1^2 + r_2^2 + \ldots r_n^2}, \tag{1}$$

provided that all r_i are measured in the same units. More generally, an appropriate choice of measurement scale for the traits is crucial to make $\|\mathbf{r}\|$ a meaningful measure of mutational size. For some sets of traits (see, e.g., Cheverud 2001), a scale may naturally present itself. However, even for an arbitrary collection of traits (including, e.g., pigmentation and capacity to store vitamin D), a reasonable dimensionless scale can be chosen, for example, by defining each r_i as a proportion of some reference phenotype or by measuring in units of the mutational or environmental standard deviation of the trait.

Denoting the fitness of phenotype $\mathbf{z}$ by $W(\mathbf{z})$, we define the selection coefficient of a mutation, $\mathbf{r}$, via $s = [W(\mathbf{z} + \mathbf{r}) - W(\mathbf{z})]/W(\mathbf{z})$. It then follows that the mutation is beneficial if $s > 0$. The quantity that interested Fisher was the probability that a mutation of given size, say r, is beneficial; that is, the fraction of mutations with $\|\mathbf{r}\| = r$, for which $s > 0$. To express this quantity succinctly, we introduce the random variable Q, defined by

$$Q = \ln(1 + s) = \ln\left[\frac{W(\mathbf{z} + \mathbf{r})}{W(\mathbf{z})}\right]. \tag{2}$$

A random mutation is then beneficial only if Q is positive. Let $\psi(q; r)$ denote the probability density of Q when mutations have size r. In other words, $\psi(q; r)dq$ is the probability that Q lies in the infinitesimal range q to $q + dq$ for mutations with $\|\mathbf{r}\| = r$. The probability that a mutation of size r is beneficial can now be written as

$$P_{\text{ben}}(r) = \int_0^\infty \psi(q; r)dq. \tag{3}$$

Kimura (1983) calculated the probability that a mutation of size r is both beneficial and achieves fixation. This is denoted by $P_{\text{fix}}(r)$. In a large Wright-Fisher population, the fixation probability of a newly arising mutant is approximately $1 - \exp(-2s)$ for $s > 0$ and 0 otherwise (see, e.g., Kimura 1983). In terms of the random variable Q, this probability is $1 - \exp[-2(e^Q - 1)]$, which is well approximated by $2Q$ when selection is not too strong, such that values of Q are typically small. In this case, $P_{\text{fix}}(r)$ is simply

$$P_{\text{fix}}(r) \simeq 2\int_0^\infty q\psi(q; r)dq. \tag{4}$$

Equations (3) and (4) show that the calculation of either $P_{\text{ben}}(r)$ or $P_{\text{fix}}(r)$ reduces to the problem of calculating the distribution, $\psi(q; r)$. To find this distribution, we need to know not just the fitness function but also the probability of occurrence of all possible mutations of given magnitude.

At this stage, we make one of the key simplifying assumptions of this work. We assume that mutations are equally likely to occur in all "directions" in the n-dimensional phenotypic space; this is equivalent to assuming a distribution of mutational changes that is radially symmetric (i.e., it depends only on the magnitude of $\mathbf{r}$, namely $\|\mathbf{r}\|$). This assumption, which is common to most previous work on Fisher's geometrical model, is made for analytical tractability. In particular, it allows us to calculate $P_{\text{ben}}(r)$ and $P_{\text{fix}}(r)$ for any given fitness function using simple procedures from " geometric probability" (Kendall and Moran 1963). We return to this assumption and its justification in "Discussion."

Normal Approximation

Of course, evaluating the integrals in equations (3) and (4) is much easier if $\psi(q; r)$ takes a simple form. It is shown below that a simple approximation often does hold when the number of traits under selection, n, is large. In this case, for some fitness functions of interest, $\psi(q; r)$ is approximately normal and takes the form

$$\psi(q; r) \simeq \frac{1}{\sqrt{2\pi v}} \exp\left[-\frac{(q-\mu)^2}{2v}\right], \tag{5}$$

where μ and v are the mean and variance of Q. The great benefit of the normal approximation is that any quantity that depends on Q can be expressed solely in terms of μ and v (as calculated for a particular fitness function). In fact, $P_{\text{ben}}(r)$, which is dimensionless, depends only on the ratio

$$\rho = -\frac{\mu}{\sqrt{v}}. \tag{6}$$

This follows from substituting equation (5) into equation (3), and making the change of variables $t = (q - \mu)/v^{1/2}$ yields

$$P_{\text{ben}}(r) \simeq \frac{1}{\sqrt{2\pi}} \int_{\rho}^{\infty} e^{-t^2/2} dt. \tag{7}$$

Similarly, the probability of a beneficial mutation with magnitude r reaching fixation is found by using equation (5) in equation (4), yielding

$$P_{\text{fix}}(r) \simeq \sqrt{\frac{2v}{\pi}} \int_{\rho}^{\infty} (t - \rho) e^{-t^2/2} dt. \tag{8}$$

The above results for $P_{\text{ben}}(r)$ and $P_{\text{fix}}(r)$ can both be expressed in terms of the error function (Abramowitz and Stegun 1970), if desired.

To make the results for $P_{\text{ben}}(r)$ and $P_{\text{fix}}(r)$ meaningful, we must of course first demonstrate that, for some fitness functions of interest, the distribution $\psi(q; r)$ is indeed approximately normal and, second, calculate the key quantities μ and v. This is done in the following sections.

Fisher's Spherically Symmetric Fitness Function

In Fisher's and Kimura's analyses, it was assumed that all traits are under stabilizing selection of identical intensity. In particular, it was assumed that the fitness of a phenotype is a monotonically decreasing function of its Euclidean distance from the optimal phenotype. Geometrically, this corresponds to a "fitness landscape" that is spherically symmetric. "Surfaces" of constant fitness are hyperspheres (i.e., circles when $n = 2$, spheres when $n = 3$, …) that are centered on the optimal phenotype. If we choose to measure each trait in such a way that its optimal value is 0, then the optimal phenotype will lie at the coordinate origin, $\mathbf{z} = \mathbf{0} \equiv (0, 0, \ldots, 0)$. Fitness is then a function of $\|\mathbf{z}\| \equiv (z_1^2 + z_2^2 + \ldots z_n^2)^{1/2}$.

Under these assumptions, a mutation is beneficial if it results in a phenotype that is closer to the origin, in trait space, than the parental phenotype; that is, if $\|\mathbf{z} + \mathbf{r}\| < \|\mathbf{z}\|$. As such, Fisher's result does not depend on the rate at which fitness declines with distance from the optimal phenotype, and it applies to any spherically symmetric fitness function of stabilizing form. However, it is convenient to specify a fitness function and then work with the quantity Q of equation (2). This facilitates comparison with the general case treated below and is necessary for the calculation of the fixation probability, $P_{\text{fix}}(r)$. With this in mind, we choose the following spherically symmetrical fitness function:

$$W_0(\mathbf{z}) = \exp(-\bar{\sigma}\|\mathbf{z}\|^2) \equiv \exp\left(-\bar{\sigma}\sum_{j=1}^{n} z_j^2\right), \tag{9}$$

where $\bar{\sigma}$ is the common (nonnegative) intensity of selection on all traits. We use the overbar, indicating a mean, for consistency with following sections. Of course, when the intensity of selection is identical for all traits, the mean intensity of selection coincides with the intensity of selection on any single trait.

In what follows, we use the subscript 0 to indicate that a quantity applies only to the spherically symmetric fitness function of equation (9). For this fitness function, the quantity Q is given by

$$Q_0 = -\bar{\sigma}\|\mathbf{z} + \mathbf{r}\|^2 + \bar{\sigma}\|\mathbf{z}\|^2 \equiv -\bar{\sigma}\sum_{j=1}^{n} (2z_j r_j + r_j^2). \tag{10}$$

The distribution of this variable, namely $\psi_0(q; r)$, can be calculated exactly (see, e.g., Kimura 1983); however, more readily comprehensible results are obtained by noting that for $n \gg 1$, the normal approximation of equation (5) applies, with the mean and variance of Q_0 are given by

$$\mu_0 = -r^2\bar{\sigma}, \tag{11}$$

$$v_0 = \frac{4r^2\bar{\sigma}^2\|\mathbf{z}\|^2}{n}. \tag{12}$$

Appendix A derives the normal approximation of $\psi_0(q; r)$, and the curve of intermediate thickness of figure

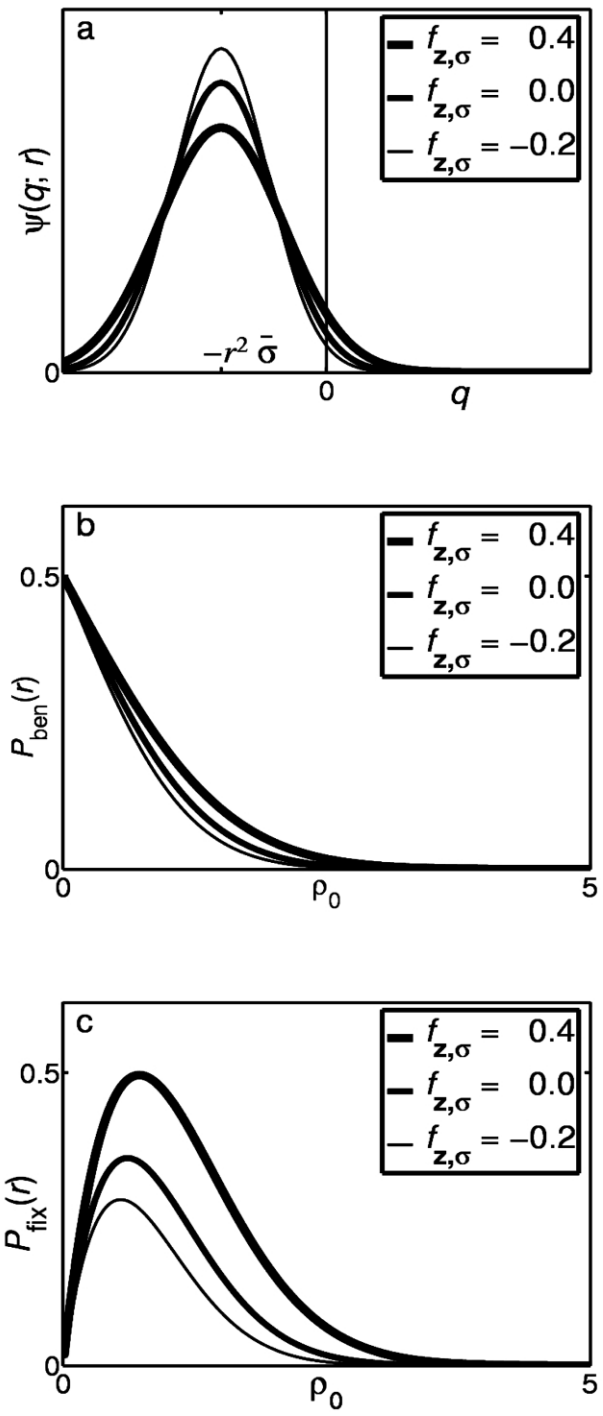

Figure 1: *a*, Plots of $\psi(q;r)$ against q, where $\psi(q;r)$ is the probability density of $Q = \ln(1+s)$ for mutations of magnitude $\|\mathbf{r}\| = r$. It is assumed that the number of traits, n, is sufficiently large that the normal approximation of equation (5) applies. Results are presented for three different values of the quantity $f_{z,\sigma}$ defined in equation (23), which captures the variation in the intensity of selection on different traits, and the extent to which the population is maladapted for particularly strongly or weakly selected traits. The curve of intermediate thickness shows results for $f_{z,\sigma} = 0$, which would hold for a symmetrical fitness landscape. The thickest and thinnest curves are for the cases $f_{z,\sigma} = 0.1 \pm 0.3$, respectively. These values were chosen to represent two cases in which selection intensity was equally variable over all traits but where the premutation phenotype resulted in a covariance term in equation (23) that was either positive (*thickest curve*) or negative (*thinnest curve*). *b*, Probability of a mutation being beneficial, $P_{ben}(r)$ (eq. [7]), is plotted against the standard magnitude of change, $\rho_0 = n^{1/2}r/(2\|\mathbf{z}\|)$, which is proportional to r. Results are plotted for the three values of $f_{z,\sigma}$ used in *a* and correspond to the areas under the probability densities of *a*, where q is positive. *c*, Equivalent plots for the probability that a mutation is beneficial and reaches fixation, $P_{fix}(r)$ (eq. [8]). We have chosen the constant of proportionality, $-2v_0/\mu_0 \equiv 8\bar{\sigma}\|\mathbf{z}\|^2/n$, to have the value 3.5 solely in order to retain the axes of *b*.

1*a* (corresponding to $f_{z,\sigma} = 0$) shows the distribution as a function of q.

Approximate normality of $\psi_0(q;r)$ follows from the fact that Q_0 is a sum of n random variables (see eq. [10]), which gives rise to a central limit type of behavior when n is large (see, e.g., Bulmer 1967). Appendix A shows that this behavior occurs despite the statistical nonindependence of the variables in the sum for Q_0 (nonindependence results from the fixed magnitude of $\mathbf{r}$). The fact that the mean of the distribution, μ_0, is always negative indicates that mutations are, on average, deleterious; this is a biologically reasonable property of Fisher's model.

Using the above results, we find that the key quantity ρ of equation (6) is

$$\rho_0 \equiv \frac{-\mu_0}{\sqrt{v_0}} = \frac{r\sqrt{n}}{2\|\mathbf{z}\|}. \tag{13}$$

The ratio ρ_0, which compares the mutation magnitude r with the quantity $\|\mathbf{z}\|/n^{1/2}$, was called the "standard magnitude of change" by Fisher (1930; also see Orr 1998), and using ρ_0 in equation (7) yields Fisher's (1930) result for the probability of a mutation of size r being beneficial. This probability, which we write as $P_{\text{ben},0}(r)$, is

$$P_{\text{ben},0}(r) = \frac{1}{\sqrt{2\pi}}\int_{\rho_0}^{\infty} e^{-t^2/2}dt \tag{14}$$

and is reproduced by the curve of intermediate thickness (corresponding to $f_{z,\sigma} = 0$) in figure 1*b*. In the plot, the rapid decline of $P_{\text{ben},0}(r)$ with ρ_0 (and hence, by eq. [13], with r) is clearly visible. Note that the intensity of selection, $\bar{\sigma}$, does not appear in ρ_0. As such, altering the value of $\bar{\sigma}$ has no affect on $P_{\text{ben},0}(r)$, confirming that Fisher's result is independent of this detail of the fitness function.

We can also use dependency on ρ_0 to express the way the fixation probability of equation (8) depends on r, when

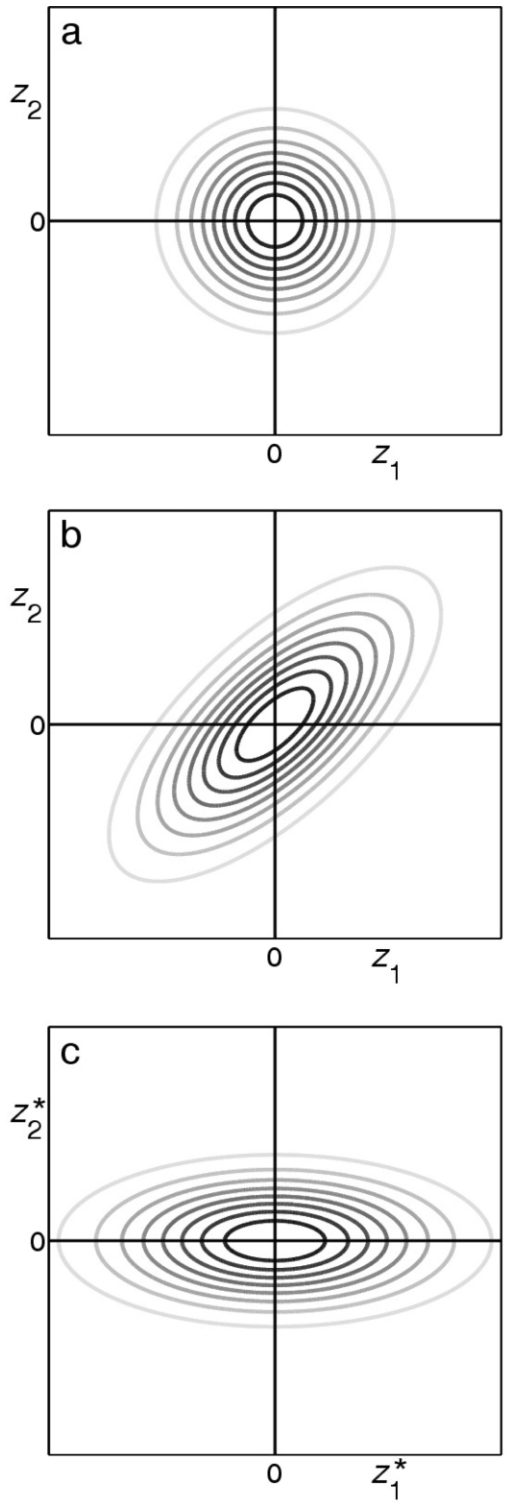

Figure 2: Surfaces of constant fitness are shown as gray curves for three fitness landscapes with $n = 2$ phenotypic characters, namely z_1 and z_2. Lighter curves indicate lower fitness. The globally optimal phenotype lies at the coordinate origin ($z_1 = 0$, $z_2 = 0$). *a*, Example of the spherically symmetrical fitness landscape of equation (9). *b*, Example of an elliptical landscape of equation (17). *c*, Landscape of *b*, with the axes rotated. This is a "diagonalization" of the fitness function of equation (17) and leads to the result of equation (18). The axes now represent "compound traits" that are linear combinations of the traits labeling the axes in *a* and *b*.

the symmetrical fitness function is assumed. From equation (8), we have

$$P_{\text{fix},0}(r) \simeq \sqrt{\frac{2v_0}{\pi}} \int_{\rho_0}^{\infty} (t - \rho_0) e^{-t^2/2} dt$$

$$\propto \frac{\rho_0}{\sqrt{2\pi}} \int_{\rho_0}^{\infty} (t - \rho_0) e^{-t^2/2} dt, \tag{15}$$

where the constant of proportionality, $8\bar{\sigma}\|\mathbf{z}\|^2/n$, is independent of r. The probability $P_{\text{fix},0}(r)$ of equation (15) is plotted as the curve of intermediate thickness (corresponding to $f_{z,\sigma} = 0$) in figure 1*c*. Unlike $P_{\text{ben},0}(r)$, which declines monotonically with ρ_0, the probability $P_{\text{fix},0}(r)$ has a maximum at an intermediate value of ρ_0 given by $\rho_0 \simeq 0.61$. From this and equation (13), it is trivial to calculate the size of mutation most likely to reach fixation. This "optimal" mutation size is given by

$$r_{\text{opt},0} \simeq 1.22\|\mathbf{z}\|/\sqrt{n}. \tag{16}$$

Again, this result is independent of the intensity of selection (although $\bar{\sigma}$ does appear in the constant of proportionality of eq. [15], with stronger selection leading to higher values of $P_{\text{fix},0}(r)$ for mutations of all sizes).

We note here that Kimura (1983) used the approximation $P_{\text{fix},0}(r) \propto \rho_0 P_{\text{ben},0}(r)$ in place of equation (15). Numerical comparison with equation (15) shows that Kimura's approximation is quite inaccurate; nevertheless, it was adequate for his purpose, which was to point out the qualitative difference in the way $P_{\text{ben}}(r)$ and $P_{\text{fix}}(r)$ vary with r.

Generalized Fitness Function

Let us now generalize Fisher's model in order to incorporate the concerns of Haldane (1932). We assume a smooth fitness landscape and a population in the vicinity of a fitness optimum, again located arbitrarily at the origin. A general fitness function, $W(\mathbf{z})$, that conforms to these assumptions may be well approximated by performing a Taylor expansion of log fitness around the optimum and then excluding all deviations from the optimum that are of higher order than quadratic (see, e.g., Apostol 1974, pp. 361–362). This produces a function of the form

$$W(\mathbf{z}) \simeq \exp\left(-\sum_{i,j=1}^{n} M_{ij} z_i z_j\right), \tag{17}$$

where $M_{ij} = -(1/2)\partial^2 \ln [W(\mathbf{z})]/\partial z_i \partial z_j|_{\mathbf{z}=0}$ is a positive definite, symmetric matrix. This is the kind of fitness function envisaged by Haldane (1932, p. 175), and it produces surfaces of constant fitness that are generally ellipses when $n = 2$ and the higher-dimensional analogs of an ellipse when $n > 2$.

To show how this fitness function embodies trait interactions, figure 2*b* depicts a two-dimensional example and compares it with its symmetrical counterpart (fig. 2*a*). Note that, in figure 2*b*, if $z_1 < 0$, then negative values of z_2 confer higher fitness than do positive values; conversely, if $z_1 > 0$, then positive values of z_2 confer higher fitness. Nothing equivalent can ever occur with a spherically symmetric fitness function (fig. 2*a*).

Rather than work directly with the fitness function of equation (17), we introduce n new traits ("compound traits") that are linear combinations of the original traits (the z_i's). Denoting these compound traits as z_i^* ($i = 1, 2, \ldots, n$), we write $z_i^* = \sum_{j=1}^{n} O_{ij} z_j$, where the O_{ij} make up an $n \times n$ square matrix. In terms of the z_i^*'s, fitness takes the form $\exp(-\sum_{i,j,k,l=1}^{n} M_{ij} O_{ik} O_{jl} z_k^* z_l^*)$. If the matrix O_{ij} is an appropriately chosen orthogonal matrix, then the matrix M_{ij} is "diagonalized" in the sense that $M_{kl}^* = \sum_{i,j=1}^{n} M_{ij} O_{ik} O_{jl}$ is a matrix whose only nonzero elements lie on the diagonal ($k = l$) and are given by σ_k ($k = 1, 2, \ldots, n$). We can then write the fitness function as $\exp(-\sum_{k,l=1}^{n} M_{kl}^* z_k^* z_l^*) \equiv \exp(-\sum_{k=1}^{n} \sigma_k z_k^{*2})$. Accordingly, we define

$$W_{\mathrm{G}}(\mathbf{z}^*) = \exp\left(-\sum_{k=1}^{n} \sigma_k z_k^{*2}\right). \tag{18}$$

This result indicates that equation (17) is formally equivalent to a fitness function where independent stabilizing selection acts on n compound traits (fig. 2*c*). Indeed, besides the definition of the traits, the only difference from the symmetrical fitness function of equation (9) is that the intensity of stabilizing selection, represented by the σ_k, may differ for each (compound) trait. Furthermore, the distribution of mutations, because it has been taken to be spherically symmetric, is completely unaffected by the above "diagonalization" (which is simply a rotation in the n-dimensional space of traits).

It is now clear that to incorporate the concerns of Haldane (1932), all that is necessary is to allow the strength of stabilizing selection to vary between traits. Haldane was, of course, aware that his generalized model could be expressed in a form resembling equation (18), and he pointed out that the compound traits involved may appear to be rather artificial, generally consisting of combinations of apparently disparate properties of an organism. (In Haldane's own example, referred to above, a single compound trait would involve measures of both pigmentation and the capacity of the liver to store vitamin D.) Nevertheless, this artificiality does not alter any of the conclusions reached here. As such, to reflect the formal similarity, in what follows we will use the notation $\mathbf{z}$ to refer to a collection of traits, regardless of whether these are compound traits, such as those that appear in equation (18).

Results for the Generalized Fitness Function

Let us now derive the quantities of interest for the generalized fitness function $W_{\mathrm{G}}(\mathbf{z})$ of equation (18). Throughout, the subscript G denotes a result applying only to this fitness function.

First, the key quantity Q is found to be

$$Q_{\mathrm{G}} = -\sum_{i=1}^{n} \sigma_i (2 z_i r_i + r_i^2). \tag{19}$$

In appendix B, it is shown that the large n normal approximation of equations (5), (7), and (8) still applies in this general case. Given that a normal approximation holds for both the symmetric and generalized fitness functions (eqq. [9], [18]), any difference between the two cases must lie solely in the mean and the variance of the variable Q. For the general case, equation (18), these quantities are found to be

$$\mu_{\mathrm{G}} = -r^2 \bar{\sigma}, \tag{20}$$

$$v_{\mathrm{G}} = 4 r^2 \overline{\sigma^2 \mathbf{z}^2} + O\left(\frac{r^4}{n}\right), \tag{21}$$

where overbars denote a mean over all traits, hence $\bar{\sigma} = n^{-1} \sum_{j=1}^{n} \sigma_j$ and $\overline{\sigma^2 \mathbf{z}^2} = n^{-1} \sum_{j=1}^{n} \sigma_j^2 z_j^2$. In equation (21), we shall neglect the $O(r^4/n)$ term that is, for practical purposes, always smaller than the term present by a factor $O(n^{-1})$ (see app. B).

Comparing the mean and variance of equations (20) and (21) with those of equations (11) and (12) reveals the relationship between the general and symmetrical cases. Clearly, the two expressions for μ are identical, so $\mu_{\mathrm{G}} = \mu_0$, but the relationship between v_{G} and v_0 is more complex. With $\mathrm{Cov}(\sigma^2, \mathbf{z}^2)$ denoting the covariance $n^{-1} \sum_{j=1}^{n} (\sigma_j^2 - \overline{\sigma^2})(z_j^2 - \|\mathbf{z}\|^2/n)$, the relationship between v_{G} and v_0 can be clarified by expressing equation (21) in the expanded form, $v_{\mathrm{G}} = 4 r^2 [\|\mathbf{z}\|^2 \overline{\sigma^2}/n + \mathrm{Cov}(\sigma^2, \mathbf{z}^2)]$. This expanded expression can be written as

$$v_G = v_0 \times (1 + f_{\mathbf{z},\sigma}), \tag{22}$$

where

$$f_{\mathbf{z},\sigma} = \frac{\mathrm{Var}(\sigma)}{\bar{\sigma}^2} + \frac{\mathrm{Cov}(\sigma^2, \mathbf{z}^2)}{\bar{\sigma}^2 \|\mathbf{z}\|^2/n}. \tag{23}$$

The consequence of equation (22) is that results for the general fitness function can be derived by making the substitution $v_0 \to v_G \equiv (1 + f_{\mathbf{z},\sigma}) \times v_0$ in equations (14) and (15), which were derived for the symmetrical fitness function. Using the subscripts G and 0 to distinguish results for the two landscapes, we find that after some simplification, we can write

$$P_{\mathrm{ben,G}}(r) = P_{\mathrm{ben},0}\left(\frac{r}{\sqrt{1 + f_{\mathbf{z},\sigma}}}\right), \tag{24}$$

$$P_{\mathrm{fix,G}}(r) = (1 + f_{\mathbf{z},\sigma}) P_{\mathrm{fix},0}\left(\frac{r}{\sqrt{1 + f_{\mathbf{z},\sigma}}}\right). \tag{25}$$

Figure 1 plots these probabilities and the corresponding distributions of Q_G for three different values of $f_{\mathbf{z},\sigma}$.

Several conclusions follow immediately from the simple results of equations (24) and (25). First, from equation (24), the probability that a mutation of magnitude r is beneficial in a generalized fitness landscape is equal to the equivalent probability for a mutant of magnitude $r/(1 + f_{\mathbf{z},\sigma})^{1/2}$ in a symmetric landscape. Second, from equation (25), the size of mutations most likely to reach fixation is

$$r_{\mathrm{opt,G}} = r_{\mathrm{opt},0}\sqrt{1 + f_{\mathbf{z},\sigma}}, \tag{26}$$

where $r_{\mathrm{opt},0}$ is the corresponding quantity for a symmetrical landscape (eq. [16]). Third, and of less relevance here, it is possible to show that both $P_{\mathrm{ben,G}}(r)$ and $P_{\mathrm{fix,G}}(r)$ will increase with $f_{\mathbf{z},\sigma}$ for mutations of any magnitude.

Together, equations (24) and (26) show that the arguments in favor of micromutationalism will be further weakened if it is generally true that $f_{\mathbf{z},\sigma} > 0$, because then the equations predict that larger mutations are more likely to contribute to adaptation than would be the case under Fisher's original model. Conversely, arguments for micromutationalism are strengthened if $f_{\mathbf{z},\sigma} < 0$. What value, then, might $f_{\mathbf{z},\sigma}$ typically have? To answer this question, consider the two terms of equation (23). The first term measures the variance in the intensity of selection on different traits; as such, it will always be positive and may be appreciable. The second term includes the covariance between z_i^2, a measure of the maladaptation of each (compound) trait, and σ_i^2, the intensity of selection on that trait. As such, if a population is particularly maladapted for particularly strongly selected traits, then this term will be positive; conversely, if the population is particularly maladapted for particularly weakly selected traits, then it will be negative. If, as seems probable, environmental changes are equally likely to affect strongly and weakly selected traits, then this term will be close to 0 on average. In this case, the always positive (variance) term will dominate, and $f_{\mathbf{z},\sigma} > 0$ will hold. However, it is possible that in some cases, the negative contribution from the second term may exceed the positive contribution from the first term, in which case $f_{\mathbf{z},\sigma} < 0$ would hold.

The dependence of the quantities $P_{\mathrm{ben,G}}(r)$ and $P_{\mathrm{fix,G}}(r)$ on details of the premutation phenotype (as described by the quantity $f_{\mathbf{z},\sigma}$) also has the implication that these probabilities cannot be expressed solely as functions of premutation fitness. To see this, it is helpful to visualize a fitness landscape with $n = 2$ traits, which are under different intensities of selection, that is, with $\sigma_1 \neq \sigma_2$.

Figure 3 shows such a landscape, with two (parental) phenotypes of equal fitness represented as solid circles and mutations as circles centered on these phenotypes. The proportion of the "mutational circles" visible indicates (see fig. 3) that the probability of a mutation being beneficial can be different even for different parental phenotypes with identical fitness.

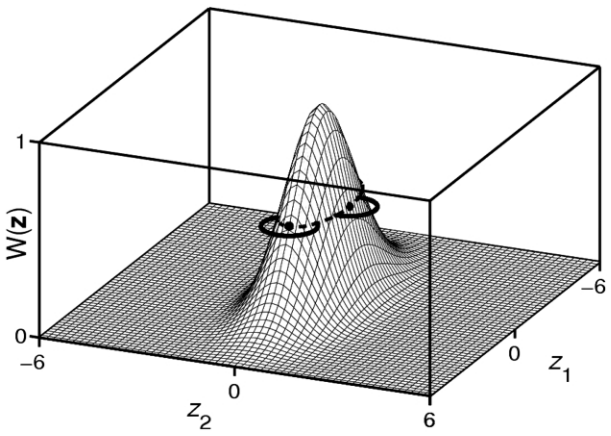

Figure 3: Fitness function $W_G(\mathbf{z})$ of equation (18), for $n = 2$ traits, is plotted as a function of the trait values z_1 and z_2. Both traits are under stabilizing selection, with intensities $\sigma_1 = 0.3$ and $\sigma_2 = 1.7$. The set of $\mathbf{z}$ values corresponding to $W_G(\mathbf{z}) = 1/2$ is plotted as a dashed line. Mutational changes of equal magnitude are illustrated as two partially visible circles. These are centered on two sets of parental trait values (*solid circles*) that both lead to identical parental fitnesses (i.e., fitness 1/2). The part of the mutational circles that are visible correspond to deleterious mutations, since the trait values of the mutant offspring would, if projected onto the fitness surface, lie at a lower fitness than that of the parent. The two mutational circles have significantly different proportions visible, indicating that there can be an appreciable range of variation in the proportion of adaptive mutations for a given fitness value.

Discussion

In this work, we have generalized the geometrical model of Fisher (1930) to take into account concerns first expressed by Haldane (1932). Haldane pointed out that Fisher's model relies on the assumptions that each trait acts independently to determine fitness and that all are under stabilizing selection of equal intensity. We have shown that Haldane's concerns can be dealt with solely by allowing the intensity of selection to vary between traits, although the relevant "traits" may have to be redefined. We have shown further that, within the framework of a normal approximation, the complications introduced by this generalization amount to the introduction of a quantity $f_{\mathbf{z},\sigma}$, equation (23). This quantity describes the variation in the intensity of selection between traits and the extent to which the premutation phenotype is maladapted for particularly strongly or weakly selected traits. Figure 1 shows that expressions for the proportion of beneficial mutations and the fixation probability of new mutations are not qualitatively altered by this change. However, conclusions about the size of mutations contributing to adaptation are affected. We have shown that conclusions will be biased in favor of larger mutations when $f_{\mathbf{z},\sigma} > 0$ and in favor of smaller mutations when $f_{\mathbf{z},\sigma} < 0$. We have also suggested that, following a random environmental change, $f_{\mathbf{z},\sigma}$ is most likely to be positive.

The problem of incorporating variable intensities of selection into Fisher's model has been investigated by previous authors. In particular, Rice (1990) gave an alternative approximation for $P_{\text{ben,G}}(r)$ (see his eqq. [3], [8], [9]). However, Rice's result is difficult to relate directly to our equation (24), because it is expressed in terms of $n - 1$ measures of the curvature of the fitness surface rather than in terms of phenotype and selection intensity, as used in this work. Orr and Coyne (1992) suggested that selection of variable intensity might be incorporated into Fisher's model by defining an "effective number of traits" (see also Orr 2000). The results above show that, when n is large, this can indeed be done by replacing n with

$$n_e = \frac{n}{1 + f_{\mathbf{z},\sigma}} \tag{27}$$

in equation (12) and related expressions. Orr and Coyne (1992) also asserted that variation in the intensity of selection would always decrease the effective number of traits and so weaken arguments in favor of micromutationalism. The results above show that this is typically true (since we can expect $f_{\mathbf{z},\sigma} > 0$ to hold more often than not), but it is not always the case. It is also true that n_e will change in value during a bout of adaptive evolution, even if the fitness landscape remains constant. This is because $f_{\mathbf{z},\sigma}$ depends on details of the phenotype and so will vary with each substitution that occurs.

Whitlock et al. (2003) used computer simulation to investigate how variation in the strength of selection affects the level of drift load experienced by a population that is evolving under Fisher's model (cf. Poon and Otto 2000). These authors found that results were not qualitatively changed, which accords with the conclusions reached here. Finally, it must be acknowledged that, although we have generalized Fisher's model somewhat, it remains highly idealized (Orr 2001, 2005*a*, 2005*b*).

Perhaps the most serious simplification involves the distribution of mutational changes. We have assumed that this distribution is spherically symmetric, but for any real biological system, this is most unlikely to be the case. Indeed, we can expect single mutations of any given size to be more common in some phenotypic directions than others (e.g., Santiago et al. 1992; Mackay 1996; Keightley and Ohnishi 1998). For some conceivable distributions, such mutational biases might be removable by applying transformations and scalings (similar to those applied here to the fitness function). However, even when such a transformation is possible, there are three objections that could be raised to its use in the present context. First, any approximations involved in the transformation (such as truncating a Taylor expansion) are likely to be least accurate for those mutations—the largest—that are of the greatest interest. Second, because the distribution of mutations is likely to vary with the premutant phenotype, maintaining a spherically symmetrical distribution would require a fresh transformation after each fixation event. Third, contriving a spherically symmetric distribution in this way is, in effect, redefining the "size" of a mutation in terms of its probability of occurrence; in the present context, such a redefinition is both counterintuitive and may lead to circularity.

While these particular objections do not apply to the transformations applied to the fitness surface, this too is an idealization: recall that each (compound) trait was assumed to be under symmetrical stabilizing selection with a single fixed optimum. Relaxing any of these assumptions is sure to alter the results presented (see, e.g., Rice 1990; Williams 1992, ch. 5; Fisher 2000, p. 302; Barton 2001; Orr 2005*b*).

Even if complications such as those mentioned above can be addressed, any attempt to use Fisher's model to make detailed quantitative predictions (relating, say, to quantitative trait locus studies; Lynch and Walsh 1998) faces formidable difficulties. These include problems with making the required measurements (e.g., Whitlock et al. 1995; Keightley et al. 2000) and more fundamental difficulties with identifying individual traits (Barton 1998; Wagner 2001). This is particularly important because, in

Fisher's scheme, the definition of mutation size refers its magnitude across all traits (eq. [1]) and not to its effect on a single focal trait (Orr 1999; Griswold and Whitlock 2003).

Despite this skepticism, Fisher's model still has important uses as a guide to intuition and, perhaps, as a phenomenological model (see, e.g., Burch and Chao 1999; Orr 2000, 2005*b*; Barton and Keightley 2002). One important role will be to encourage careful examination of the predictions from other modeling frameworks that rely on the a priori assumption that mutations have vanishingly small size. These include Fisher's own infinitesimal model (Fisher 1918; Turelli and Barton 1994) and the more recent adaptive dynamics approach (Geritz et al. 1998; Barton and Polechová, forthcoming; Waxman and Gavrilets, forthcoming).

Acknowledgments

It is a pleasure to thank J. Feng for helpful discussions and H. A. Orr for helpful correspondence. We also thank two anonymous reviewers for their constructive suggestions. The research of D.W. was supported by the Leverhulme Trust and that of J.J.W. by the Biotechnology and Biological Sciences Research Council.

APPENDIX A

Normal Approximation of the Distribution of Q_0

In this appendix, we determine the normal approximation of the probability density of the random variable Q_0, which is defined in equation (10). With $\delta(\cdot)$ denoting a Dirac δ function, the exact probability density of Q_0 is

$$\psi_0(q;r) = E[\delta(q - Q_0)] = E\left\{\delta\left[q + \bar{\sigma}\sum_{j=1}^{n}(2z_j r_j + r_j^2)\right]\right\},$$

where $E[\ldots]$ denotes an average over $\mathbf{r}$'s of fixed magnitude (r) with a radially symmetric distribution.

We can write $\sum_{j=1}^{n}(2z_j r_j + r_j^2) = 2\|\mathbf{z}\| r\cos\theta + r^2$, where θ is the angle between $\mathbf{z}$ and $\mathbf{r}$. Adopting spherical polar coordinates in n dimensions, we find

$$\psi_0(q;r) = \frac{\int_0^{\pi}\sin^{n-2}\theta\,\delta(q + 2\bar{\sigma}\|\mathbf{z}\| r\cos\theta + \bar{\sigma}r^2)d\theta}{\int_0^{\pi}\sin^{n-2}\theta\,d\theta},$$

where all angular integrals, except the one over θ, cancel between numerator and denominator. This last result can be written as

$$\psi_0(q;r) = \frac{\int_{-\pi/2}^{\pi/2}\cos^{n-2}\theta\,\delta(q + \bar{\sigma}r^2 + 2\bar{\sigma}\|\mathbf{z}\| r\sin\theta)d\theta}{\int_{-\pi/2}^{\pi/2}\cos^{n-2}\theta\,d\theta}. \quad \text{(A1)}$$

If $|q + \bar{\sigma}r^2| > 2\bar{\sigma}\|\mathbf{z}\| r$, the δ function remains 0 and results in $\psi_0(q;r)$ vanishing. When $|q + \bar{\sigma}r^2| < 2\bar{\sigma}\|\mathbf{z}\| r$, a single θ, namely $-\arcsin[(q + \bar{\sigma}r^2)/(2\bar{\sigma}\|\mathbf{z}\| r)]$, contributes to the integral and leads to

$$\psi_0(q;r) \propto \left[1 - \left(\frac{q + \bar{\sigma}r^2}{2\bar{\sigma}\|\mathbf{z}\| r}\right)^2\right]^{(n-3)/2}.$$

Assuming $n \gg 1$ and using $\ln(1 + x) \simeq x$ for $|x| \ll 1$ allow $\psi_0(q;r)$ to be approximated as

$$\psi_0(q;r) \propto \exp\left\{\frac{n-3}{2}\ln\left[1 - \left(\frac{q + \bar{\sigma}r^2}{2\bar{\sigma}\|\mathbf{z}\| r}\right)^2\right]\right\}$$
$$\simeq \exp\left[-\frac{n}{2}\left(\frac{q + \bar{\sigma}r^2}{2\bar{\sigma}\|\mathbf{z}\| r}\right)^2\right].$$

From this result, we can read off the mean and the variance of Q_0, namely $\mu_0 = -r^2\bar{\sigma}$ and $v_0 = 4r^2\bar{\sigma}^2\|\mathbf{z}\|^2/n$. These are, in fact, the exact mean and variance of Q_0.

Note that the normal approximation to $\psi_0(q;r)$ can also be directly derived from equation (A1) by making the approximation

$$\int_{-\pi/2}^{\pi/2}\cos^{n-2}\theta \to \int_{-\infty}^{\infty}\exp\left(-\frac{n\theta^2}{2}\right),$$
$$\sin\theta \to \theta. \quad \text{(A2)}$$

Formally, this approximation corresponds to treating the different r_i as statistically independent normal random variables with mean 0 and variance r^2/n.

APPENDIX B

Normal Approximation of the Distribution of Q_G

In this appendix, we outline a derivation of a normal approximation of the probability density of Q when the fitness landscape has the generalized form, equation (18). This appendix builds upon the intuition developed in appendix A with the derivation relying on n being large.

We have carried out Monte Carlo simulations, where the components of $\mathbf{z}$ were independently drawn from identical normal distributions with mean 0, and the magnitude, $\|\mathbf{z}\|$, was adjusted to have the fixed value of unity. Values of the σ_k were independently drawn from a uniform distribution over the interval 0.2–1.8. Then 10^5 mutations from a spherically symmetric distribution with fixed mag-

nitude r were generated. Our findings suggest that the normal approximation is reasonably accurate for $n \sim 10$ traits: the error in the proportion of beneficial mutations is approximately 3% (7%) when $r = 0.2$ (0.4) and more accurate for $n = 50$ traits; the error in the proportion of beneficial mutations is approximately 1% (3%) when $r = 0.2$ (0.4).

We begin by writing the probability density of Q_G as $\psi_G(q; r)$ and shall use the integral representation of a Dirac δ function $\delta(u) = (2\pi)^{-1}\int_{-\infty}^{\infty} \exp(i\lambda u)d\lambda$, where $i = (-1)^{1/2}$ and λ is an arbitrarily named integration variable. We then have

$$\psi_G(q; r) = E[\delta(q - Q_G)]$$
$$= (2\pi)^{-1}\int_{-\infty}^{\infty} d\lambda \exp(i\lambda q) E\left\{\exp\left[i\lambda \sum_{j=1}^{n} (2z_j r_j + r_j^2)\sigma_j\right]\right\},$$

where $E[\ldots]$ denotes an average over $\mathbf{r}$'s of fixed magnitude (r) with a radially symmetric distribution. Using the identity

$$\sqrt{\frac{i}{4\pi\sigma\lambda}}\int_{-\infty}^{\infty} \exp\left(-\frac{ix^2}{4\lambda\sigma} + irx\right)dx = \exp(i\lambda\sigma r^2)$$

allows us to write

$$\psi_G(q; r) = \int_{-\infty}^{\infty}\frac{d\lambda}{2\pi}dx_1 dx_2 \ldots dx_n\left(\prod_{j=1}^{n}\sqrt{\frac{i}{4\pi\sigma_j\lambda}}\right)$$
$$\times \exp\left(-i\sum_{j=1}^{n}\frac{x_j^2}{4\lambda\sigma_j} + i\lambda q\right) \quad \text{(B1)}$$
$$\times E\left\{\exp\left[i\sum_{j=1}^{n}(2\lambda z_j\sigma_j + x_j)r_j\right]\right\}.$$

We write the expectation appearing in the above equation as $E[\exp(i\sum_{j=1}^{n} b_j r_j)]$. Using spherical polar coordinates in n dimensions leads, for large n (using eq. [A2]), to

$$E\left\{\exp\left(i\sum_{j=1}^{n} b_j r_j\right)\right\} = \frac{\int_{-\pi/2}^{\pi/2}\cos^{n-2}\theta e^{-i\|\mathbf{b}\|r\sin\theta}d\theta}{\int_{-\pi/2}^{\pi/2}\cos^{n-2}\theta d\theta}$$
$$\simeq \frac{\int_{-\infty}^{\infty} e^{-n\theta^2/2}e^{-i\|\mathbf{b}\|r\theta}d\theta}{\int_{-\infty}^{\infty} e^{-n\theta^2/2}d\theta} = e^{-(\|\mathbf{b}\|^2 r^2)/2n}.$$

Using this result in equation (B1) and integrating over all x_j yields

$$\psi_G(q; r) = \int_{-\infty}^{\infty}\frac{d\lambda}{2\pi}e^{i\lambda q}\prod_{j=1}^{n}\left(1 - \frac{2i\lambda r^2\sigma_j}{n}\right)^{-1/2}$$
$$\times \exp\left(-\frac{2r^2\lambda^2}{n}\sum_{j=1}^{n}\frac{z_j^2\sigma_j^2}{1 - 2ir^2\lambda\sigma_j/n}\right). \quad \text{(B2)}$$

We approximate equation (B2) by

$$\psi_G(q; r) \simeq$$
$$\int_{-\infty}^{\infty}\frac{d\lambda}{2\pi}\exp\left\{i\lambda(q + r^2\bar{\sigma}) - \lambda^2\left[2r^2\overline{\sigma^2\mathbf{z}^2} + O\left(\frac{r^4}{n}\right)\right]\right\}, \quad \text{(B3)}$$

where $\bar{\sigma} = n^{-1}\sum_{j=1}^{n}\sigma_j$ and $\overline{\sigma^2\mathbf{z}^2} = n^{-1}\sum_{j=1}^{n}\sigma_j^2 z_j^2$. On neglecting the $O(r^4/n)$ terms in equation (B3) and carrying out the λ integral, we obtain the normal approximation of equation (5), supplemented by equations (20) and (21).

The $O(r^4/n)$ terms in equation (B3) are not perfectly captured by the normal approximation of equation (A2), since deviations of the r_i from normality begin to manifest themselves at terms of this order. This is not problematic since the $O(r^4/n)$ terms are, under usual circumstances, a factor $O(n^{-1})$ down on the leading $2r^2\overline{\sigma^2\mathbf{z}^2}$ term in equation (B3). To see this, we take the σ_i independent of n and first note that we can expect $\bar{\sigma}\|\mathbf{z}\|^2 \lesssim 1$ to hold; this follows from equation (18) and the assumption that genetic load is not too high. It follows from this that typically $\|\mathbf{z}\|^2 = O(1)$ and $\overline{\sigma^2\mathbf{z}^2} \sim \bar{\sigma}^2\|\mathbf{z}\|^2/n = O(n^{-1})$. Next, we note that the formula for P_{ben} and P_{fix} in equations (7) and (8) are at all appreciable and of significance only when $\rho \lesssim 1$. For the generalized case, this occurs when $r\bar{\sigma}/[4\overline{\sigma^2\mathbf{z}^2} + O(r^2/n)]^{1/2} \lesssim 1$, and this necessarily requires $r^2 \lesssim \|\mathbf{z}\|^2/n$; that is, $r^2 = O(n^{-1})$. For large n, it is this last estimate that allows us to neglect the $O(r^4/n)$ term in equation (B3), which is then seen to be $O(n^{-3})$, compared with the leading $2r^2\overline{\sigma^2\mathbf{z}^2}$ term, which is $O(r^2/n) \equiv O(n^{-2})$.

Literature Cited

Abramowitz, M., and I. Stegun. 1970. Handbook of mathematical functions. Dover, New York.

Apostol, T. M. 1974. Mathematical analysis. 2nd ed. Addison-Wesley, Reading, MA.

Barton, N. H. 1998. The geometry of adaptation. Nature 395:751–752.

———. 2001. The role of hybridization in evolution. Molecular Ecology 10:551–568.

Barton, N. H., and P. D. Keightley. 2002. Understanding quantitative genetic variation. Nature Reviews Genetics 3:11–21.

Barton, N. H., and J. Polechová. Forthcoming. The limitations of adaptive dynamics as a model of evolution. Journal of Evolutionary Biology.

Bulmer, M. G. 1967. Principles of statistics. Oliver & Boyd, Edinburgh.

Burch, C. L., and L. Chao. 1999. Evolution by small steps and rugged landscapes in the RNA virus phi6. Genetics. 151:921–7.

Cheverud, J. M. 2001. The genetic architecture of pleiotropic relations and differential epistasis. Pages 411–433 *in* G. P. Wagner, ed. The character concept in evolutionary biology. Academic Press, San Diego, CA.

Fisher, R. A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. Transactions of the Royal Society of Edinburgh 52:399–433.

———. 1930. The genetical theory of natural selection. Oxford University Press, Oxford.

———. 2000. The genetical theory of natural selection: a complete variorum edition. Oxford University Press, Oxford.

Geritz, S. A. H., E. Kisdi, G. Meszena, and J. A. J. Metz. 1998. Evolutionary singular strategies and the adaptive growth and branching of the evolutionary tree. Evolutionary Ecology 12:35–57.

Griswold, C. K., and M. C. Whitlock. 2003. The genetics of adaptation: the roles of pleiotropy, stabilizing selection and drift in shaping the distribution of bidirectional fixed mutational effects. Genetics 165:2181–2192.

Haldane, J. B. S. 1932. The causes of evolution. Longmans Green, London.

Hartl, D., and C. H. Taubes. 1996. Compensatory nearly neutral mutations: selection without adaptation. Journal of Theoretical Biology 182:303–309.

———. 1998. Towards a theory of evolutionary adaptation. Genetica 102/103:525–533.

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.

Kendall, M. G., and P. A. P. Moran. 1963. Geometric probability. Hafner, New York.

Kimura, M. 1983. The neutral theory of molecular evolution. Cambridge University Press, Cambridge.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, MA.

Mackay, T. F. C. 1996. The nature of quantitative genetic-variation revisited: lessons from *Drosophila bristles*. Bioessays 18:113–121.

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

———. 1999. The evolutionary genetics of adaptation: a simulation study. Genetical Research 74:207–214.

———. 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. Evolution and Development 3: 121–123.

———. 2005*a*. The genetic theory of adaptation: a brief history. Nature Reviews Genetics 6:119–127.

———. 2005*b*. Theories of adaptation: what they do and don't say. Genetica 123:3–13.

Orr, H. A., and J. A. Coyne. 1992. The genetics of adaptation: a reassessment. American Naturalist 140:725–742.

Peck, J. R., G. Barreau, and S. C. Heath. 1997. Imperfect genes, Fisherian mutation and the evolution of sex. Genetics 145:1171–1199.

Poon, A., and S. P. Otto. 2000. Compensating for our load of mutations: freezing the meltdown of small populations. Evolution 54: 1467–1479.

Rice, S. H. 1990. A geometric model for the evolution of development. Journal of Theoretical Biology 143:319–342.

Santiago, E., J. Albornoz, A. Dominguez, M. A. Toro, and C. Lopez. 1992. The distribution of spontaneous mutations on quantitative traits and fitness in *Drosophila melanogaster*. Genetics 132:771–781.

Turelli, M., and N. H. Barton. 1994. Genetic and statistical analyses of strong selection on polygenic traits: what, me normal? Genetics 138:913–941.

Wagner, G. P., ed. 2001. The character concept in evolutionary biology. Academic Press, San Diego, CA.

Waxman, D., and S. Gavrilets. Forthcoming. 20 questions on adaptive dynamics: a target review. Journal of Evolutionary Biology.

Welch, J. J., and D. Waxman. 2003. Modularity and the cost of complexity. Evolution 57:1723–1734.

Whitlock, M. C., P. C. Phillips, F. B. G. Moore, and S. J. Tonsor. 1995. Multiple fitness peaks and epistasis. Annual Review of Ecology and Systematics 26:601–629.

Whitlock, M. C., C. K. Griswold, and A. D. Peters. 2003. Compensating for the meltdown: the critical effective size of a population with deleterious and compensatory mutations. Annales Zoologici Fennici 40:169–183.

Williams, G. C. 1992. Natural selection: domains, levels and challenges. Oxford University Press, Oxford.

Associate Editor: Michael C. Whitlock
Editor: Jonathan B. Losos