Towards a theory of evolutionary adaptation

Daniel L. Hartl[1] & Clifford H. Taubes[2]

[1]*Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138, USA (Phone: 617-496-3917; Fax 617-496-5854; e-mail d-hartl@harvard.edu);* [2]*Department of Mathematics Harvard University, One Oxford Street, Cambridge, MA 02138, USA*

Abstract

Most theoretical models in population genetics fail to deal in a realistic manner with the process of mutation. They are consequently not informative about the central evolutionary problem of the origin, progression, and limit of adaptation. Here we present an explicit distribution of phenotypes expected in an ensemble of populations under a mutation-selection-drift model that allows mutations with a distribution of adaptive values to occur randomly in time. The model of mutation is a geometrical model in which the effect of a new mutation is determined by a random angle in n dimensional space and in which the adaptive value (fitness) of an organism decreases as the square of the deviation of its phenotype from an optimum. Each new mutation is subjected to random genetic drift and fixed or lost according to its selective value and the effective population number. Time is measured in number of fixation events, so that, at any point in time, each population is regarded as genetically homogeneous. In this mutation-selection-drift model, among an ensemble of populations, the equilibrium average phenotype coincides with the optimum because the distribution of positive and negative deviations from the optimum is symmetrical. However, at equilibrium, the mean of the absolute value of the deviation from the optimum equals $\sqrt{n/8Ns}$, where n is the dimensionality of the trait space, N is the effective population size, and s is the selection coefficient against a mutation whose phenotype deviates by one unit from the optimum. Furthermore, at equilibrium, the average fitness across the ensemble of populations equals $1 - (n + 1)/8N$. When n is sufficiently large, there is a strong mutation pressure toward the fixation of slightly deleterious mutations. This feature relates our model to the nearly neutral theory of molecular evolution.

Introduction

Almost every theoretical model in population genetics can be classified into one of two major types. In one type of model, mutations with stipulated selective effects are assumed to be present in the population as an initial condition. The selective effects may be constant, frequency dependent, density dependent, and so forth; or they may vary either systematically or randomly in time. These models are not very realistic in their treatment of mutation because, in the real world, the mutations on which natural selection acts are not present as an initial condition but appear at random points in time. The second major type of models does allow mutations to occur at random intervals in time,

but the mutations are assumed to be selectively neutral or nearly neutral. The neutral models are also not very realistic in their treatment of mutation. Although real mutations do appear randomly in time, they are not all selectively neutral or nearly neutral. The classical selection models were pioneered by Fisher, Haldane, and Wright; the classical neutral models by Wright and Kimura. Examples of both classes of models are discussed at some length in Hartl and Clark (1997).

While not realistic in their treatment of mutation, both types of classical models have been very useful within their respective domains of applicability. Almost everything known about the theoretical implications of natural selection has come from the analysis of classical selection models, and almost every-

thing presently understood about random genetic drift is based on the analysis of classical neutral models.

The classical models have been so useful that there has been little or no impetus to develop alternatives. Both types of models are ideally suited to address the issue of genetic variation: How much genetic variation is there and how is it maintained? It is therefore no coincidence that population genetics has been dominated by this issue, with experimental observations eliciting new models and the new models stimulating further observations (Lewontin, 1974; 1991).

But genetic variation, while it is an important problem in evolutionary biology, is not the only problem. Another problem, at least equal in importance to genetic variation, is that of the origin, progression, and limit of adaptation. These issues have been discussed since Darwin's time. How do adaptations originate? How do they become progressively better? What are the limits of adaptation?

One feature of classical population genetic models is that they are not well-suited to address the problem of the origin, progression, and limit of adaptation. For these issues, a different type of approach is needed, in which mutations are assumed to arise sequentially in time and in which each new mutation has a favorable, neutral, or detrimental effect on the trait under selection according to some continuous distribution of selective effects. One difficulty in developing such models is that nothing is known about the distribution of selective effects of newly arising mutations. The statement that 'most mutations are harmful' is of no help because it is (1) incorrectly stated, and (2) possibly untrue. As to its being incorrectly stated, the only empirically verifiable statement along these lines is that 'most mutations that have a detectable effect on phenotype are harmful.' This leaves open the question whether there are a large number of mutations that have no detectable effect on phenotype and that are selectively neutral or nearly neutral. The undetectable latter may outnumber the detectable former, and so the unqualified statement that 'most mutations are harmful' might be empirically false.

Although the issue of assessing the distribution of selective effects of newly arising mutations is an important epistemological problem, it is also a critical consideration in any theory concerning the origin, progression, and limit of adaptation. The theory of metabolic control of flux through biochemical pathways, pioneered by Kacser and collaborators (Kacser & Burns, 1973, 1981; Keightley & Kacser, 1987), offers one possible starting point for modeling the effects of new mutations. (Kacser's seminal work is commemorated in a special issue of the Journal of Theoretical Biology, October 7, 1996, Volume 182, Number 3). A mutational model based on metabolic control theory is implicit in the seemingly paradoxical concept of the evolution of selective neutrality, which deals with one aspect of the limits of adaptation (Hartl, Dykhuizen & Dean, 1985).

Another approach to the distribution of effects of new mutations is based on geometrical considerations, originally suggested by Fisher (1930). He used this model primarily to show that the probability that a mutation is favorable decreases with the magnitude of its effect, approximately according to the area under the tail of a normal distribution. Fisher's formulation, characteristically, was very short on detail. In a recent paper we have filled in the blanks (Hartl & Taubes, 1996).

In the present paper, we again use Fisher's approach to model the distribution of selective effects of newly arising mutations. Each new mutation affects the phenotype of a continuous trait according to the magnitude of a randomly chosen angle in n-dimensional space. Each affects fitness in proportion to the square of the deviation of the phenotype from an optimum value. The ultimate fixation or loss of each new mutation is also a random process governed by the selective effect of the new mutation and the effective size of the population.

For this mutation-selection-drift model, we derive an approximation that is valid when the number of dimensions is sufficiently large. The approximation leads to an expression for the mean phenotype of the continuous trait at the selection-mutation-drift equilibrium, which is a function of the number of dimensions, the population size, and the magnitude of the fitness effects. Surprisingly, the mean equilibrium phenotype does not depend on the 'step size', or average magnitude, of the mutational effects. Equally surprising is the finding that the mean equilibrium fitness depends only on the number of dimensions and the effective population size.

Fisher's mutation model

Fisher's mutation model can be described with reference to Figure 1. He proposes by analogy that the phenotype of an organism with respect to some trait can be represented as the length of a line segment, x, which measures the distance by which the phenotype
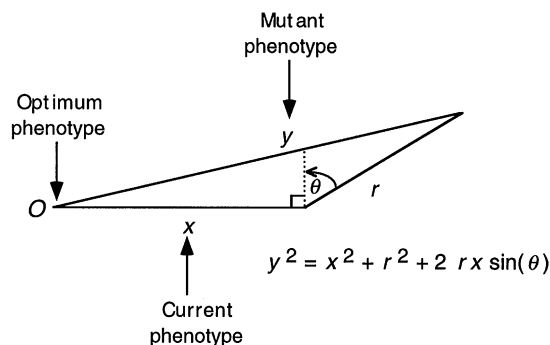
Figure 1. Geometrical model of mutation. The parameter θ is a random angle in n dimensions, and the step length, r, of each mutation is a constant. The value x is the current phenotype, y that of a new mutation.

Figure 2. Distribution of a random angle θ in n dimensions.

Figure 3. Effect of increasing step size, r, on the distribution of phenotypes due to new mutant alleles. In these examples, the current phenotype is $x = 2$, $n = 16$, and $N = 50$ with $s = 0.02$. After many fixation events in an ensemble of populations, the equilibrium value of x (x_0) is 1.414, indicated by the arrow. (The equilibrium is calculated from Equation 17).

departs from an optimum, represented by the point O. The effect of a mutation on the phenotype is determined by the magnitude of a random angle, θ, which can range from $-\pi/2$ to $+\pi/2$, and a fixed 'step size', r. The phenotype of an organism carrying the new mutation is equal to the length of the line segment, y, which is given by

$$y^2 = x^2 + r^2 + 2rx\sin(\theta) \ . \tag{1}$$

The angle θ depicted in Figure 1 differs from that defined in our previous paper by the quantity $-\pi/2$ (Hartl & Taubes, 1996). At this point it would be appropriate to symbolize x as $|x|$ and y as $|y|$ because both are defined as line segment lengths and so are necessarily positive. For ease of representation, we will continue to use x and y but emphasize that they are absolute deviations from the optimum phenotype, not signed deviations.

In a space of n dimensions, a random angle θ between $-\pi/2$ and $+\pi/2$ has a probability density given by

$$Z\cos^{n-2}(\theta) \tag{2}$$

where Z is a constant chosen so that $\int Z\cos^{n-2}(\theta)d\theta = 1$ when the range of integration is $-\pi/2$ to $+\pi/2$. Some examples for various values of n are illustrated in Figure 2. As n becomes large, the distribution of θ becomes increasingly concentrated around $\theta = 0$ and, as shown below, it approaches a normal distribution. An intuitive explanation of Equation (2) is that, on a sphere of n dimensions, each tiny patch on the surface is equally likely to include the spot to which a new mutation changes the phenotype, assuming that mutations in all
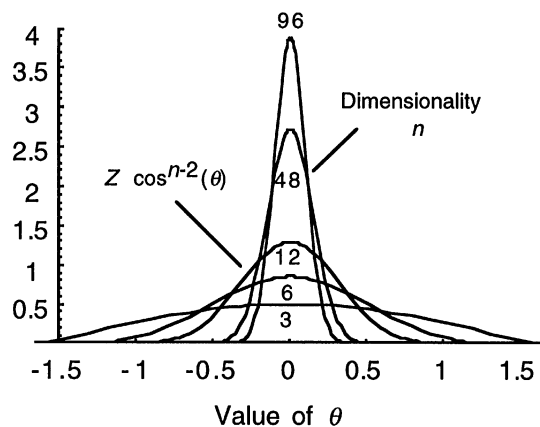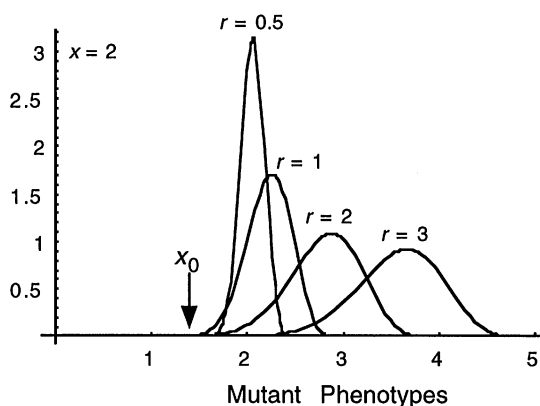
directions are equally likely. The number of such patches near the poles is much smaller than the number near the equator, hence θ has a distribution that is relatively concentrated near 0; and larger values of n increase the concentration near 0. The fact that $\cos^{n-2}(\theta)$ is the appropriate measure follows from geometrical considerations.

The distribution of θ in Equation (1) can be transformed into a distribution of y using Equation (2) and its derivative. The resulting explicit equations are not very illuminating, but some numerical examples are plotted in Figure 3. The distribution of y is approximately symmetric, but the mean is strongly dependent on the step size, r, more precisely on the magnitude of r relative to the current position, phenotype x. When r

is sufficiently small relative to x, favorable mutations are quite frequent; in the limit as $r \to 0$, the probability of a favorable mutation goes to 1/2 (Fisher, 1930). On the other hand, because r is small, most of the favorable mutations also have small effects. As r increases relative to x, the probability of a favorable mutation decreases also. For the specific values of r in Figure 3, the probabilities of a favorable mutation for $r = 0.5$, 1, 2, and 3 are 0.308, 0.159, 0.023, and 0.001, respectively (Hartl & Taubes, 1996).

Fixation filtration

Not all new mutations are incorporated into the population, but undergo a sort of 'filtering' resulting from the combined effects of selection and random genetic drift. Once a new mutation has occurred, and its selective effect is known, this part of the process is a classical selection-drift process. In a diploid population of constant size N, the probability of ultimate fixation of a single copy of a new mutation with selection coefficient s given by

$$2s/(1 - e^{-4Ns}) \qquad (3)$$

(Crow & Kimura, 1970). Equation (3) is valid for detrimental mutations ($s < 0$) as well as for favorable mutations ($s > 0$). The selection coefficient s is that for the heterozygous genotype, and semidominance is assumed. As $s \to 0$, L'Hôpital's rule implies from Equation (3) that the probability of fixation of a single copy of a neutral mutation is $1/2N$.

The value of the selection coefficient for a new mutation must depend on the phenotypic value x, of the original genotype, and y, of the new mutant. Let us suppose that the absolute fitness of an organism decreases as the square of the deviation from an optimum, conveniently taken as 0. The fitness of an organism with phenotype x can therefore be written as $c - sx^2$ and that of an organism with phenotype y as $c - sy^2$. Therefore, the difference in fitness of an organism with phenotype y, relative to an organism of phenotype x, is given by

$$s_{xy} = -s(y^2 - x^2) \qquad (4)$$

where the fitnesses are in malthusian parameters (Crow & Kimura, 1970). In a population in which the current phenotype is x, the probability of fixation of a single copy of a new mutation with phenotype y is given by

$$2s_{xy}/(1 - e^{-4Ns_{xy}}) . \qquad (5)$$

Expressions (1), (2), and (5) can be combined to yield an explicit equation for the probability density of the phenotype, y, of new mutations that happen to be fixed, given an initial population homozygous for an allele with phenotype x. Some examples of the distributions are shown in Figure 4. In these examples, an ensemble of populations initially fixed for an allele yielding a phenotype of $x = 0.8$ undergoes mutation, and the new mutations also have a step length of $r = 0.8$. The distribution of phenotypes among the mutants is given in each curve labeled 'new mutations'. The new mutations are then 'filtered' according to the fixation process in Equation (5), and the resulting distribution of phenotypes among the populations that have fixed for a new allele is given in each curve labeled 'fixed mutations'. In panel A, $n = 4$; and in panel B, $n = 16$. The value of Ns has been changed accordingly, from $Ns = 1$ in panel A to $Ns = 4$ in panel B, to keep the equilibrium at a value of x at 0.707 (shown below in Equation 17). Although random genetic drift is much less important in panel A than in panel B, there are still a significant number of fixations that move the population farther from the origin (as well as farther from the equilibrium value).

Two features of the process represented in Figure 4 need to be emphasized:

- Each distribution of phenotypes is a distribution among an ensemble of populations, not the distribution of phenotypes within any one population. At any one point in time, each population in the ensemble is genetically homogeneous, and the time required for fixation, when fixation occurs, is assumed to be short relative to the waiting time for a new mutation destined to be fixed.
- The time scale of the mutation-selection-drift process is measured in number of fixations rather then number of generations. In every population of the ensemble, each time a new mutant allele becomes fixed, the clock ticks one additional unit of time.

Approximation for high dimensionality

Although the exact expressions for the curves in Figure 4 are cumbersome, there is a convenient approximation when the dimensionality, n, is sufficiently large. In this case, Equation (2) converges to a normal prob-
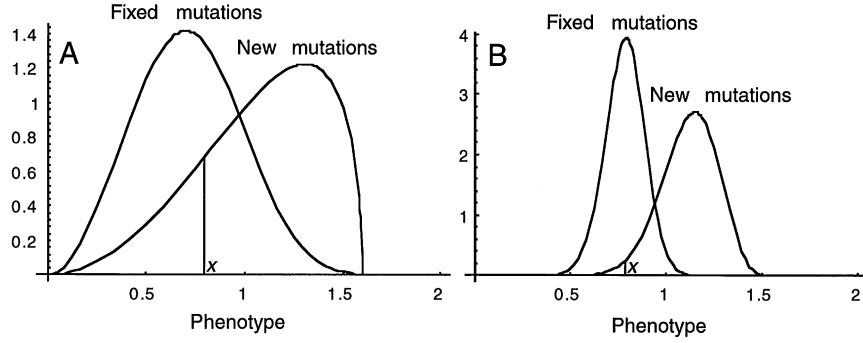
Figure 4. Distribution of phenotypes among all new mutations and among all mutations that are fixed, given a current phenotype of $x = 0.8$ and a step size of $r = 0.8$. The other parameters are: (A) $n = 4$, $N = 50$, $s = 0.02$, $Ns = 1$; (B) $n = 16$, $N = 200$, $s = 0.02$, $Ns = 4$. In both cases the equilibrium value of x is 0.707, calculated from Equation 17.
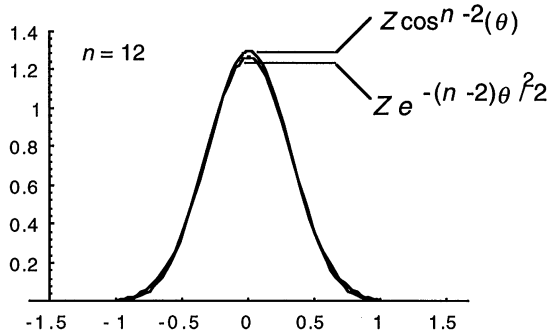

Figure 5. The distribution of a random angle in n dimensions converges rapidly to a normal distribution, here illustrated for $n = 12$.

ability density. The reason is that, when n is large, θ becomes concentrated around $\theta = 0$, and then

$$\cos(\theta) \approx 1 - \theta^2/2 \approx e^{-\theta^2/2}$$

so that the distribution of θ is given by

$$Z\cos^{n-2}(\theta) \approx Ze^{-(n-2)\theta^2/2} . \tag{6}$$

This approximation is already excellent for $n = 12$, as shown in Figure 5.

The probability of finding the population at a distance y after a fixation, given that it was initially at a distance x, is given by the probability of a new mutation yielding a phenotype of y times the probability of such a mutation being fixed. It therefore follows from Equations (5) and (6) that the probability of the transition from y to x is given by

$$p[y(\theta)|x] = Z' \cdot s_{xy} \cdot (1 - e^{-4Ns_{xy}})^{-1} \cdot e^{-n\theta^2/2} \tag{7}$$

where s_{xy} is given in Equation 4 and where Z' is a normalization constant chosen to make

$$\int p[y(\theta)|x]d\theta = 1 .$$

Equilibrium distribution

To consider the ramifications of Equation (7), write

$$s_{xy} = -s[r^2 + 2rx\sin(\theta)] . \tag{8}$$

Because the factor $exp(-n\theta^2/2)$ in Equation (7) is very much peaked around $\theta = 0$ when n is large, we may approximate $\sin(\theta) \approx \theta$. With this substitution, Equation (8) becomes

$$s_{xy} = -2srx(\theta + r/2x) . \tag{9}$$

The assumption that θ is concentrated around 0 is valid only if

$$r^2 << n/Ns . \tag{10}$$

On the one hand, if Ns is too large, the selection process dominates and the population is driven towards the origin O in Figure 1. On the other hand, if r is too large, relative to x, then all angles move the phenotype away from O.

Next, replace the variable θ with the variable

$$\zeta = \theta + r/2x .$$

Then Equation (7) becomes

$$p[y(\zeta, x)|x] = Z'' \cdot \zeta \cdot e^{8Nsrx\zeta} - 1)^{-1} \cdot e^{nr\zeta/2x} \cdot e^{-n\zeta^2/2} . \tag{11}$$

In terms of the variable ζ, the condition $\zeta > 0$ means that a new mutation has a phenotype $y > x$, whereas $\zeta < 0$ means that a new mutation has a phenotype $y < x$. The constant Z'' is again a normalization constant.

Now we define the equilibrium phenotype x_0 to be the distance x such that, when $x = x_0$, the population is equally likely to become fixed for either a favorable or an unfavorable new mutation. It follows that, for $x = x_0$,

$$\int_{\zeta > 0} p[y(\zeta, x_0)|x_0]d\zeta = \int_{\zeta < 0} p[y(\zeta, x_0)|x_0]d\zeta .$$
(12)

To analyze Equation (12) further, it is convenient to make the substitutions

$$K = 4Nsrx$$
$$L = nr/2x$$
(13)

which, using Equation (14), reduce the equilibrium condition in Equation (12) to

$$\int_{\zeta > 0} \zeta \cdot (e^{2K\zeta} - 1)^{-1} \cdot e^{L\zeta} \cdot e^{-n\zeta^2/2}d\zeta =$$
$$\int_{\zeta < 0} \zeta \cdot (e^{2K\zeta} - 1)^{-1} \cdot e^{L\zeta} \cdot e^{-n\zeta^2/2}d\zeta .$$
(14)

A certain symmetry in Equation (14) allows its almost immediate solution. First multiply the left-hand integrand by $1 = exp(-K\zeta)/exp(-K\zeta)$; then multiply the right-hand integrand by $1 = exp(K\zeta)/exp(K\zeta)$ and bring the right-hand side to the left by changing ζ to $-\zeta$. With these substitutions, the condition in Equation (14) reduces to

$$\int_{\zeta > 0} \zeta \cdot (e^{K\zeta} - e^{-K\zeta})^{-1} \cdot$$
$$(e^{(L-K)\zeta} - e^{-(L-K)\zeta}) \cdot e^{-n\zeta^2/2}d\zeta = 0 .$$
(15)

Because the integrand in Equation (15) is either strictly positive, when $K > L$, or else strictly negative, when $K < L$, Equation (15) can be satisfied if and only if $K = L$. From the definitions in (13), $K = L$ implies that

$$x^2 = n/8Ns$$
(16)

and so the equilibrium value of the phenotype, x_0, is given by

$$x_0 = \sqrt{n/8Ns} .$$
(17)

We emphasize again that x_0 is the mean of the absolute value $|x|$ of the deviation of phenotype from the optimum. Taking the sign of the deviation into account,

the mean deviation is zero because the distribution of the signed deviations is symmetrical around the optimum. Furthermore, each individual population does not come to a stable equilibrium with an absolute deviation of which (17) is the mean. Rather, each population evolves continuously, fixing a sequence of favorable and unfavorable alleles determined, at each step, according to the probability given in Equation (7). At any instant in time, each population is genetically homogeneous. The equilibrium in Equation (17) refers to the mean phenotype of an ensemble of otherwise identical populations evolving independently under mutation-selection-drift. It can also be shown that the mean squared absolute deviation, $E(x_0^2)$, is given by $E(x_0^2) = (n + 1)/8Ns$.

Mean fitness at equilibrium

Because the fitness of an organism with phenotype x is given by $c - sx^2$, it follows that, across an ensemble of populations, the mean fitness at equilibrium is proportional to

$$1 - sE(x_0^2) = 1 - s[(n+1)/8Ns] = 1 - [(n+1)/8N]$$
(18)

which, perhaps surprisingly, does not depend on the constant s that determines the decrease in fitness as a quadratic function of distance from the origin (Figure 1). (The rate of approach to the equilibrium does, however, depend on s.) The average fitness in (18) is measured as a Malthusian parameter (Hartl & Clark, 1997); because both n and N are positive, the maximum average fitness equals 1. Equation (18) shows that the average equilibrium fitness increases as a function of N and decreases as a function of n. The implication is that, when n is large, genetic systems of higher dimensionality are less likely to be exquisitely adaptable (in the sense of evolving ever closer to the optimum) than systems of lower dimensionality. The reason seems to be that, the greater the value of n, the more tightly concentrated is θ around 0, and the less likely a favorable mutation is to occur.

In view of the important role of the dimensionality n in Equation (18), it is a pity that Fisher (1930), his insight notwithstanding, does not expatiate his views on whether the dimensionality has a concrete biological interpretation. Reading between the lines, one gets the sense that he expects n to be quite large, at least in most cases: 'The representation in three dimensions is evidently inadequate; for even a single organ, in cas-
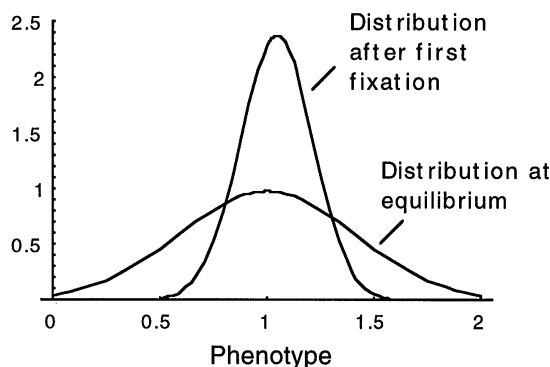
Figure 6. Distribution of phenotypes among an ensemble of populations after one fixation and at equilibrium, assuming that each population begins at the equilibrium phenotype of $x = 1$. The optimum phenotype is $x = 0$ and the parameters are $s = 0.03$, $n = 12$, and $N = 50$.

es in which we know enough to appreciate the relation between structure and function, as is, broadly speaking, the case with the eye in vertebrates, often shows this conformity in many more than three respects' (Fisher, 1930, p. 39). In writing this, Fisher was obviously thinking of the dimensionality in terms of morphology and physiology. In modern evolutionary genetics, it would be of very great interest to develop an interpretation of n appropriate to the number of loci affecting a quantitative trait or to the number of subunits in sequences of either amino acids or nucleotides, for such an interpretation would connect Fisher's model to other theories of molecular evolution in an evidently straightforward manner.

Equilibrium distribution

Another approach, beyond the scope of this paper, has enabled a more detailed analysis of the implications of Equation (7). The details of the analysis will be published elsewhere, but the main result is as follows. For sufficiently large n, the equilibrium distribution of the mean phenotype in an ensemble of populations is given by a normal distribution with mean given by Equation (17) and variance $1/4Ns$. In particular, the equilibrium distribution of x (that is, of $|x|$) is

$$f(x) = \sqrt{\frac{2Ns}{\pi}} e^{-2Ns(x-x_0)^2} . \qquad (19)$$

An example is shown in Figure 6. An ensemble of populations, each initially at the equilibrium value of 1, undergoes mutation-selection-drift and, after one

fixation event in each population, yields the distribution of phenotypes after one fixation. This process continues in each population until the ensemble eventually attains the equilibrium distribution. Once again, the equilibrium distribution is an ensemble property, as each population in the ensemble is genetically homogeneous.

Equilibrium ensemble distributions are illustrated for some specific examples in Figure 7. In panel A, n is fixed at $n = 12$ while Ns increases from 0.1 to 5. Two things happen as Ns increases. First, the variance of the distribution decreases, from 2.5 for $Ns = 0.1$ to 0.25 for $Ns = 1$. At the same time, the mean of the ensemble comes closer to the optimum phenotype. Both of these features are caused by stronger selection and/or decreasing random genetic drift as Ns increases.

In panel B, Ns is held constant at $Ns = 1$ while the dimensionality n increases from 12 to 96. In this case, while the variance of the ensemble distribution stays the same, the equilibrium mean of the ensemble moves progressively farther away from the optimum phenotype. This feature of the process is caused by the greater concentration of θ around 0 as n increases. Although the variance of the equilibrium distribution does not change with increasing n, the coefficient of variation does. In particular, the coefficient of variation of the equilibrium distribution is $\sqrt{(2/n)}$, which is a function only of n.

Discussion

The mutation-selection-drift model presented here is in the spirit of the simulation model of Bürger and Lande (1994), who considered a quantitative trait. In their model, the fitness of an organism decreases as a Gaussian function of the deviation of the organism's phenotype from an optimum, and if the deviation is not too large, this is equivalent to the quadratic selection model in our Equation (4). They assume a distribution of new mutations whose effects are symmetrically distributed with mean 0 in a finite diploid population of size N. Their simulations indicate that, whereas the effective number of loci influencing the trait may be quite large, the number of segregating alleles per locus is typically rather small, and a significant fraction of the loci may be monomorphic. Furthermore, the authors find that the mean phenotype across an ensemble of populations is distributed as an almost perfect normal distribution with a mean that coincides with the phenotypic optimum. Superficially, this seems in
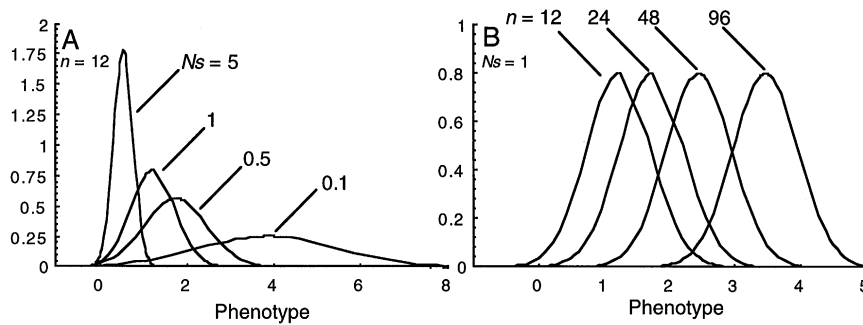
Figure 7. Equilibrium distributions of phenotypes among an ensemble of populations (A) as a function of *Ns* and (B) as a function of *n*.

contradiction to our result that the equilibrium average absolute deviation from the optimum phenotype is nonzero. However, the measure of phenotype in the Bürger-Lande model is not the average absolute deviation from the optimum but the deviation itself, which may be positive or negative. Averaged across the entire ensemble, the positive and negative deviations cancel, and the average equilibrium deviation is 0. This is true in our model as well, but the scale on which we measure phenotype is the *absolute* deviation from the optimum, in which the positive and negative deviations do not cancel, and so the equilibrium mean of $|x|$ equals $\sqrt{(n/8Ns)}$ as indicated in Equation 17.

It is also interesting to note that, in our model, the variance of $|x|$ around 0 is given by $(n + 1)/8Ns$. The simulations of Bürger and Lande indicate an extrordinary correlation between the variance in phenotype and the ratio of n/N, where, in their case, n is the 'effective' number of loci. (The effective number of loci corrects for the fact that not all segregating loci contribute equally to the trait). In the simulations of Bürger and Lande, the correlation coefficient between the phenotypic variance and n/N equals 0.999. The strong dependence of the variance on n/N seems to suggest an underlying congruence between the models. It is as if each of the n effective loci in the Bürger-Lande model were equivalent to one dimension in trait space in our model, but we have not examined this speculative interpretation in detail.

In our model, for large n, the distribution of θ is peaked near $\theta = 0$, and the distribution of mutational effects is nearly normal in distribution (Figure 3). Biological intuition suggests that new mutations with large effects should have very little chance of being favorable. Fisher makes this intuition plausible by suggesting an analogy:

Consider the mechanical adaptation of an instrument, such as a microscope, when adjusted for distinct vision. If we imagine a derangement of the system by moving a little each of the lenses, either longitudinally or transversely, or by twisting through an angle, by altering the refractive index and transparency of the different components, or the curvature, or the polish of the interfaces, it is sufficiently obvious that any large derangement will have a very small probability of improving the adjustment, while in the case of alterations much less than the smallest of those intentionally effected by the maker or operator, the chance of improvement should be almost exactly one half. (Fisher, 1930, p. 40)

Our model incorporates the equal probability feature for very small displacements because, for very large n, θ is very close to 0, and with θ very close to 0, y is very close to x provided that r is sufficiently small (Equation 1).

The reason that the mean absolute deviation does not coincide with the phenotypic optimum in our model is not that this point is unattainable. Any given population may become homozygous for a genotype yielding the optimum phenotype. The reason is that, in n dimensional space, when θ is highly concentrated near 0, the vast majority of mutations are harmful; in spite of the smaller likelihood of fixation of unfavorable mutations, there are so many of them that the mutation pressure exerts a strong pull tending to move the ensemble of populations away from the optimum. This pull is counteracted by selection, the effect of which increases as the product *Ns*. Eventually, in an ensemble of populations, there comes a time when the offsetting forces of mutation and selection come into balance, and at this equilibrium the average fitness among the ensemble is proportional to 1 - $(n + 1)/8N$, which increases

as *n* becomes small and/or *N* large (Equation 18). It is important to note that one of the key assumptions in the nearly neutral theory of molecular evolution is a relentless tendency toward the fixation of slightly detrimental mutations because of an overwhelming preponderance of slightly detrimental effects (Ohta, 1973, 1992; Ohta & Tachida, 1990). Hence, our model affords a specific geometrical intepretation of mutation that may prove useful in further understanding the nearly neutral process and its evolutiary implications.

Acknowledgements

References

Bürger, R. & R. Lande, 1994. On the distribution of the mean and variance of a quantitative trait under mutation-selection-drift balance. Genetics 138: 901–912.

Crow, J.F. & M. Kimura, 1970. An Introduction to Population Geneics Theory. Harper & Row, New York.

Fisher, R.A., 1930. The Genetical Theory of Natural Selection. Oxford University Press, Oxford.

Hartl, D.L. & A.G. Clark, 1997. Principles of Population Genetics. Sinauer Associates, Sunderland, MA.

Hartl, D.L., D.E. Dykhuizen & A.M. Dean, 1985. Limits of adaptation: The evolution of selective neutrality. Genetics 111: 655–674.

Hartl, D.L. & C.H. Taubes, 1996. Compensatory nearly neutral mutations: Selection without adaptation. J. Theoret. Biol. 182: 303–309.

Kacser, H. & J.A. Burns, 1973. The control of flux. Symp. Soc. Exp. Biol. 32: 65–104.

Kacser, H. & J.A. Burns, 1981. The molecular basis of dominance. Genetics 97: 639–666.

Keightley, P.D. & H. Kacser, 1987. Dominance, pleiotropy and metabolic structure. Genetics 117: 319–329.

Lewontin, R.C., 1974. The Genetic Basis of Evolutionary Change. Columbia University Press, New York.

Lewontin, R.C., 1991. Electrophoresis in the development of evolutionary genetics: Milestone or millstone? Genetics 128: 657–662.

Ohta, T., 1973. Slightly deleterious mutant substitutions in evolution. Nature 246: 96–98.

Ohta, T., 1992. The nearly neutral theory of molecular evolution. Annu. Rev. Ecol. Syst. 23: 263–286.

Ohta, T. & H. Tachida, 1990. Theoretical study of near neutrality: I. Heterozygosity and rate of mutant substitution. Genetics 126: 219–229.