J. theor. Biol. (1996) **182,** 303–309

Compensatory Nearly Neutral Mutations: Selection without Adaptation*

DANIEL L. HARTL† AND CLIFFORD H. TAUBES‡

† *Department of Organismic and Evolutionary Biology, Harvard University, 16 Divinity Avenue, and ‡ Department of Mathematics, Harvard University, One Oxford Street, Cambridge, MA* 02138, *U.S.A.*

One implication of Kacser's analysis of complex metabolic systems is that mutations with small effects exist as a consequence of the typically small flux control coefficient relating enzyme activity to the rate of a metabolic process. Although a slightly detrimental mutation is somewhat less likely to become fixed by chance than a slightly favorable mutation, mutations that are slightly detrimental might be expected to be more numerous than favorable mutations owing to the previous incorporation of favorable mutations by a long history of natural selection. The result is that, as Ohta has pointed out, a significant fraction of mutations that are fixed in evolution are slightly detrimental. In the long run, the fixation of detrimental mutations in a gene increases the opportunity for the occurrence of a compensatory favorable mutation, either in the same gene or in an interacting gene. On a suitably long timescale, therefore, every gene incorporates favorable mutations that compensate for detrimental mutations previously fixed. This form of evolution is driven primarily by natural selection, but it results in no change or permanent improvement in enzymatic function.

Introduction

Every family of proteins related by common ancestry is a mosaic of residues that show conservation and those that show variation. Shared sequence motifs serve to identify new members of protein families or subfamilies. It is generally accepted that conserved sequence motifs define a structural framework that is essential for protein folding or function. Although some amino acids may be conserved across the members of a protein family merely by chance, conservation is improbable if sufficient time has elapsed since the proteins last shared a common ancestor. The theory that conserved sequence motifs result from structural constraints is supported by the striking observation that, in some cases, the conserved sequences can substitute for one another across widely separated phylogenetic lineages. On the other hand, because each conserved amino acid residue has a fixed identity, the theory of structural constraints is a theory of non-evolution rather than a theory of evolution. The subject of evolution is divergence and, hence, non-identity.

If conserved sequence motifs tell us about structural constraints, what can we learn from sequence variation? When proteins in a family differ in function, the differences must be caused by a subset of the variable amino acids. The more difficult case to interpret is when protein function is conserved. Structural variation is then observed as polymorphisms within species or divergence between species. To what extent are such differences important in adaptive evolution? The absolutist views are usually presented as the neutral theory and the selection theory. The neutral theory asserts that most amino acid polymorphisms within species, and most amino acid replacements between species, are functionally equivalent in their effects on Darwinian fitness (Kimura, 1983). Strictly speaking, neutral mutations have $Ns = 0$, where N is the effective population number and s is the selection coefficient. Because Ns

* This paper is dedicated to the memory of Henrik Kacser.

† Author to whom correspondence should be addressed.

is continuous in both N and s, the equality $Ns = 0$ is unlikely to be realized exactly. Ohta proposed that $Ns \approx 0$ is a more realistic hypothesis, with $s < 0$ because most mutations are harmful (Ohta, 1973, 1992). The implication of her theory, called the nearly neutral theory of molecular evolution, is that most amino acid replacements result from the fixation of slightly detrimental mutations. In contrast, the selection theory asserts that Darwinian selection acting upon favorable mutations is the driving force behind most gene substitutions.

The neutral and selection theories are usually represented as antithetical, but this is an artificial dichotomy. The common feature underlying both processes is the mutational landscape defined by the relative proportions of favorable and unfavorable mutations and the distribution of the magnitude of their effects. To the extent that the mutational landscape may change as a result of adaptive evolution, the relative importance of random genetic drift and selection is changeable and historically contingent. In a previous paper, we invoked metabolic control theory (Kacser & Burns, 1973, 1981) to argue that natural selection can create a condition in which most mutations with small effects on protein structure are selectively nearly neutral (Hartl *et al*., 1985). This seeming paradox is a consequence of long continued natural selection for increased enzyme activity resulting in metabolic saturation and an ever diminishing effect on fitness of any small change in enzyme activity. Hence, the more highly adapted a molecule already is, the greater the chance that any new mutation will be nearly neutral.

In the long run, the deterioration in fitness resulting from the accumulation of detrimental mutations must be balanced by the fixation of compensatory mutations (Ohta & Tachida, 1990). In this sense, the continuance of natural selection at the molecular level depends on random genetic drift. In an enzyme-coding gene, for example, the chance fixation of detrimental mutations creates the opportunity for an offsetting favorable mutation that affects either the structure of the same protein or that of another macromolecule with which it interacts. Furthermore, in a metabolic system near saturation, as well as in a system with an intermediate optimum, most compensatory mutations merely restore, partially or completely, the level of enzyme function to that of the ancestral molecule. There will usually be no overall increase in absolute fitness. On this treadmill, each molecule evolves adaptively but does not improve.

Probability of Fixation

An important theoretical consideration is that the probability of ultimate fixation of a new mutant allele (one present in a single copy in a diploid population) is given approximately by $2s/(1 - e^{-4Ns})$, where s is the selection coefficient of the heterozygote, assuming no dominance and small s (Crow & Kimura, 1970). When N is large compared with s, then $u(s) \approx 2s$. This approximation is the basis of the conventional wisdom that the probability of fixation of even a favorable mutation is small. For example, if $s = 0.01$ (a 1% selective advantage), then $u(s) \approx 1/50$. Although not incorrect, the generalization is misleading. The relevant comparison is not the absolute probability of fixation of a new mutant allele but rather the probability of fixation compared with that of another allele (for instance, a selectively neutral allele) in the same population. Because $u(0) = 1/2N$ in an ideal population, if we take $N = 10^6$ in the above example, then $u(0.01)/u(0) \approx 4 \times 10^4$. In other words, although a mutation with $s = 0.01$ is not very likely to be fixed, it is still 40 000 times more likely to be fixed than a neutral allele in the same population.

Figure 1(a) shows the curve $u(s)/u(0)$ on a logarithmic scale for various values of Ns; the relation is $u(s)/u(0) = 4Ns/(1 - e^{-4Ns})$, which depends only on the product Ns. The curve is plotted for Ns from -1 to $+1$ because, for reasons outlined below, we consider small selection coefficients (those on the order of $1/N$) the most relevant. Note that, over this range of parameter values, a favorable allele is not much more likely to become fixed, relative to a neutral allele, than is a detrimental allele. For example, $u(0.5)/u(0) = 2.31$ whereas $u(-0.5)/u(0) = 0.31$; hence, a favorable allele with $Ns = 0.5$ has only a 7.4-fold greater chance of being fixed than a detrimental allele with $Ns = -0.5$. A direct comparison of fixation probabilities of newly arisen favorable and detrimental mutations is given in Fig. 1(b), in which the ratio $u(s)/u(-s)$ is plotted as a function of Ns. Even for Ns as large as 1, the ratio is about 50; and for Ns as small as 0.1, the ratio is about 1.5.

Fisher's Model of Adaptation

With regard to the role of detrimental vs. favorable mutations in adaptive evolution, a good deal of insight can be gained into the problem by considering the geometrical analysis of Fisher (1930). Fisher's argument is that the statistical requirements of evolutionary adaptation, in which each biological component must conform to others in a large number

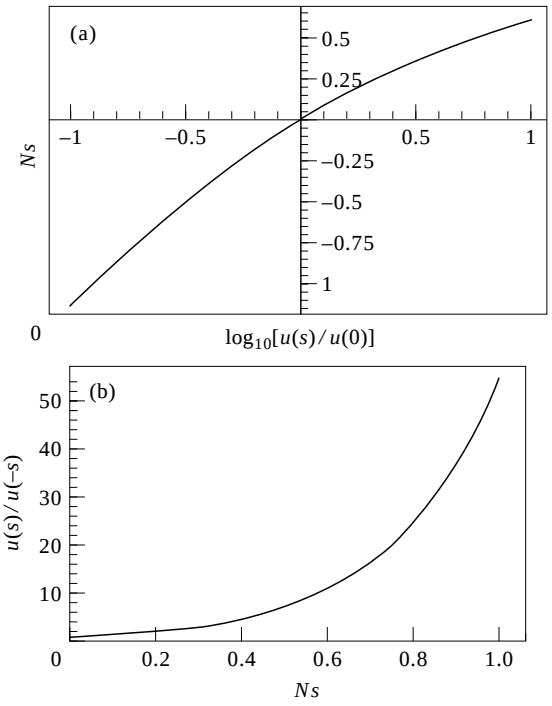

FIG. 1. Relative probability of fixation. Ns is the product of the selection coefficient s and the effective population number N. (a) Fixation probability $u(s)$ of a semi-dominant mutation with selection coefficient s, relative to that of a neutral mutation $u(0)$. Note that the abscissa is logarithmic. (b) Fixation probability of a semi-dominant mutation with selection coefficient s, relative to that of a detrimental mutation with selection coefficient $-s$.

of different respects, are analogous to those governing the closeness with which a point in space, which he calls A, approaches a fixed point O. All possible adaptations superior to A are included inside a sphere passing through A and centered at O. In this situation, if the point A is moved a fixed distance, r, in any direction by a new mutation, then the new mutation will be beneficial if the new location lies within the sphere and deleterious otherwise. If d is the diameter of the sphere, then between the limits $r = 0$ and $r = d$, the probability of improvement in the position of the point A by random movement equals $(1/2)[1 - (r/d)]$. The probability that the new mutation is favorable therefore decreases linearly from 1/2 to 0 as r increases from 0 to d.

In a large number, n, of dimensions, the probability that a change of magnitude r is favorable conforms to the right-hand tail of the standard normal distribution integrated from $r\sqrt{n}/d$, where r is again the magnitude of the change and $\sqrt{n}/d$ is a standard unit of adaptation (Fisher, 1930, and see below). Fisher notes that:

> the higher the adaptation the smaller will this standard [unit] be, and consequently the smaller the probability that a change of a given magnitude shall effect an improvement. The situation may be expressed otherwise by supposing changes of a given magnitude to occur at random in all directions, and comparing the rates of progress caused by two opposite selective agencies, one of which picks out and accumulates all changes which increase the adaptation, and another which similarly picks out and accumulates all which diminish it. For changes very small compared to the standard, these two agencies will be equally effective, but, even for changes of only one-tenth of the standard, the destructive selection is already 28 percent more effective than the selection favoring adaptation (Fisher, 1930, p. 39–40).

He further asserts that the ratio of the value "destroying adaptation" to that "building it up" for $r = 1/2$, 1, 2, and 3 times the standard unit equals 3.5, 13, 236, and 7852, respectively.

The source of Fisher's ratio of the value "destroying adaptation" to that "building it up" may be made clear with reference to Fig. 2. The points O and A separated by a distance $d/2$ are as defined by him, and r and θ are the distance and angle that the point A moves as a result of a new mutation. The new distance d' along the dashed line is given by $(d^2/4 + r^2 - rd\cos\theta)^{1/2}$, and thus the change in the distance to O, considered as a function of the angle θ, is $\Delta(\theta) = d/2 - (d^2/4 + r^2 - rd\cos\theta)^{1/2}$.

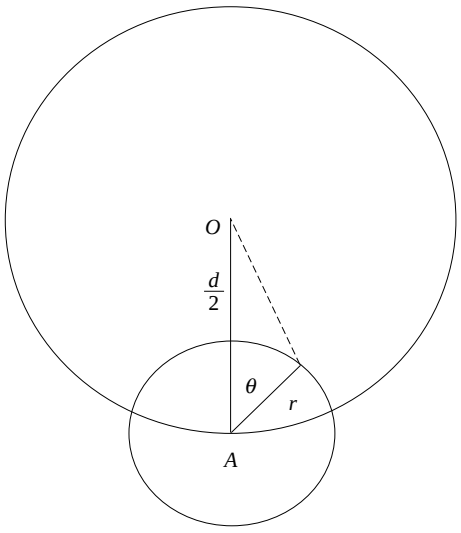

FIG. 2. Fisher's model of adaptation in two dimensions. The quantitative measure of adaptation is the closeness with which the point A approaches the fixed point O.

On the other hand, in a space of n dimensions, the probability that a random change in the position of point A decreases the distance to O is given by

$$\frac{\int_0^{\theta'} \sin^{n-2}\theta d\theta}{\int_0^{\pi} \sin^{n-2}\theta d\theta} \tag{1}$$

where $\cos\theta' = r/d$. These integrals can be evaluated by the method of steepest descent because, for large n, the integrand is concentrated near $\theta = \pi/2$. The evaluation entails changing the variable to $x = \sqrt{n}\cos\theta$ and writing $\sin^n\theta \approx e^{-x^2/2}$ for θ near $\pi/2$. Therefore, when n is large, the integral in (1) is essentially

$$\int_{\eta}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \tag{2}$$

where $\eta = r\sqrt{n}/d$. (The function in (2), incidentally, is the one plotted in Fisher's fig. 3 on p. 40.)

As regards adaptation, Fisher's average value of a mutation "building up" adaptation can be interpreted as the average distance by which a new mutation moves a point toward O, given that the distance to O is decreased. This quantity is given by the ratio

$$\frac{\int_0^{\theta'} \Delta(\theta) \sin^{n-2}\theta \, d\theta}{\int_0^{\theta'} \sin^{n-2}\theta \, d\theta} \tag{3}$$

Using Taylor's theorem to approximate $\Delta(\theta) \approx d\eta/n(x-\eta)$, where $\cos\theta = x/\sqrt{n}$, the ratio in (3) can again be evaluated by the method of steepest descent. For large n,

$$\frac{d\eta}{n} \frac{\int_{\eta}^{\infty} (x-\eta) \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx}{\int_{\eta}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx}. \tag{4}$$

Similarly, the average value "destroying adaptation" can be interpreted as the average distance by which a new mutation moves a point away from O, given that the distance to O is increased. Evaluation of a ratio of integrals analogous to that in (3) leads to

$$\frac{d\eta}{n} \frac{\int_{-\infty}^{\eta} (\eta - x) \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx}{\int_{-\infty}^{\eta} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx}. \tag{5}$$

The denominator in (1) and (2) is the probability of a favorable or an unfavorable mutation, respectively, and so the average favorable and unfavorable effects, weighted by their frequencies, equal the numerators of the two expressions. Therefore, Fisher's ratio is the ratio of the numerators of (5) to (4), or

$$\frac{\int_{-\infty}^{\eta} (\eta - x) e^{-x^2/2} dx}{\int_{\eta}^{\infty} (x - \eta) e^{-x^2/2} dx} \tag{6}$$

which evaluates to 3.53, 13.00, 236.55, and 7851.23 for $\eta = 1/2$, 1, 2, and 3 times the standard unit. A plot of the ratio (Fig. 3) makes for a bleak evolutionary scenario. (Note that the ordinate is scaled as the logarithm.) The agency that picks out and accumulates all changes that diminish adaptation is apparently much stronger than the agency that similarly picks out and accumulates all changes which increase it.

On the other hand, in a finite population, the agency that accumulates mutations is selection and random genetic drift. Mutations that are detrimental are accumulated less efficiently than mutations that are beneficial. For a new detrimental mutation of selective value $-Ns$, compared with one with a selective value $+Ns$, the ratio of fixation probabilities

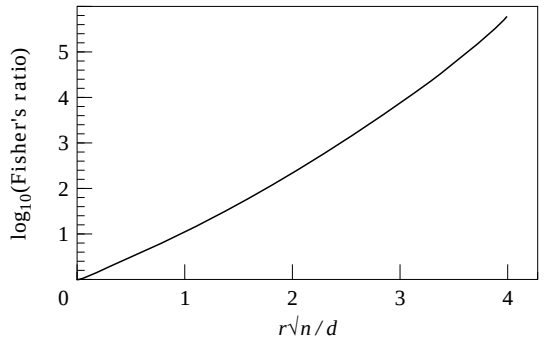

FIG. 3. Evaluation of the expression (6), or Fisher's ratio. Note that the scale of the ordinate is logarithmic.

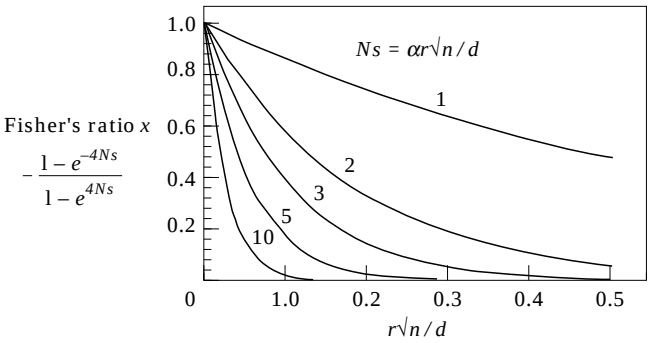

FIG. 4. Fisher's ratio in expression (6) corrected for the differing probabilities of fixation of unfavorable and favorable mutations. Ns is assumed to be proportional to the magnitude of any mutational change, in standard units. The numbers on the curves are the values of the parameter α, which is a scaling factor for selective constraints. Increasing α implies increasing constraints.

is $-(1-e^{4Ns})/(1-e^{-4Ns})$. Fisher's ratio should be modified to take this into account, but how should we interpret $r\sqrt{n/d}$ in terms of Ns? It may be reasonable to assume that Ns is proportional to $r\sqrt{n/d}$, say, $Ns = \alpha r\sqrt{n/d}$. The parameter α is not a characteristic of a population but rather of a gene at a particular time in its evolutionary history. Genes that are more constrained have larger values of α, which decreases the likelihood of fixation of detrimental mutations even if they have relatively small effects. Figure 4 shows Fisher's ratio, modified to take into account the differing fixation probabilities. The destructive force is substantially weakened. Indeed, the modified ratio is always smaller than one, and it tends to zero rather quickly if α is at all large. The picture is much more favorable for adaptation because random genetic drift does not pick out and accumulate detrimental mutations but rather eliminates all but those with small effects.

Detrimental and Favorable Mutations

There is one important issue left to consider, which is the ratio of unfavorable to favorable mutations. With regard to favorable mutations with large effects, orthodox evolutionary thinking since Darwin has regarded them as extremely rare except, perhaps, for times of major environmental change when a previously unfavorable mutation may become favorable. To use Fisher's geometrical analogy, for most adaptations the point A is already quite close to point O owing to the past history of natural selection. In the nearly neutral theory of evolution (Ohta, 1992), a key premise is that most gene substitutions involve a mutation whose effect is on the order of the reciprocal of the effective population number. The reasoning is that detrimental mutations that cause a large decrease in fitness have a negligible probability of being fixed by chance, and that most favorable mutations with large effects should have been fixed already.

An independent argument, based on general metabolic considerations due to Kacser and collaborators (Kacser & Burns, 1973, 1981; Keightley & Kacser,1987), also suggests that most gene substitutions should involve mutations with small effects (Hartl *et al*., 1985). The hyperbolic curve in Fig. 5(a) is the expected relation between the reaction velocity or flux through a linear metabolic pathway and the activity of any of the enzymes in the pathway when the enzymes have classical Michaelis–Menten kinetics (Kacser & Burns, 1973). Such hyperbolic curves are observed in numerous metabolic pathways (Kacser & Burns, 1979). Activity is the total enzyme activity, and so it expressly includes the genetic regulatory network affecting the expression of the enzymes. The plateau results in a saturation effect in which the change in flux for a given small change in enzyme activity decreases as a function of the activity. When the enzyme activity is large, a small change in activity

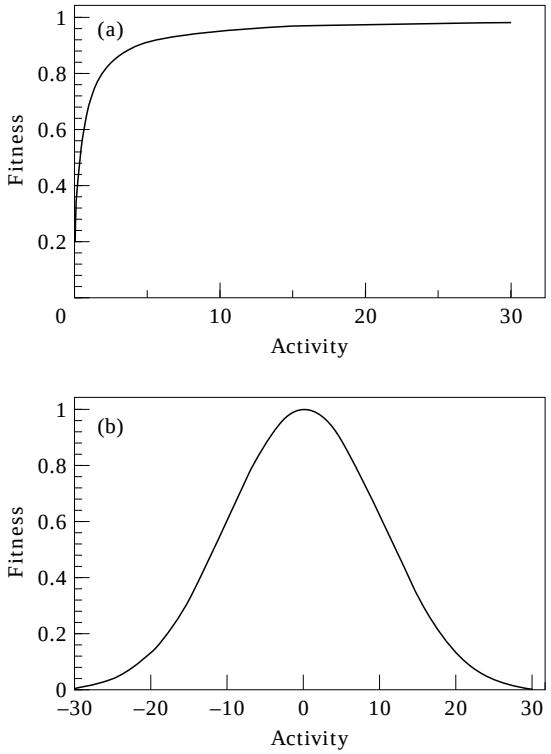

FIG. 5. Fitness as a function of flux through a metabolic pathway implied by Kacser's model of metabolic control (Kacser & Burns, 1973, 1979). (a) When fitness increases with flux, scaled in such a way that fitness equals one when activity equals 30 (Hartl *et al*., 1985). (b) When fitness is maximized at an intermediate optimum flux (Dean *et al*., 1988).

results in an exceedingly small change in flux. The curve in Fig. 5(a) is standardized to the form $J = E/(1 + E)$, where the flux J is the ordinate and the enzyme activity E is the abscissa. With this scaling, the effect of a change in activity diminishes according to $dJ/dE = 1/(1 + E)^2$. More generally, each enzyme in a metabolic pathway has an associated flux control coefficient, defined at steady state under a given set of conditions as the proportional change $(dJ/J)/(dE/E) = 1/(1 + E)$. For the enzymes in a segment of any complex metabolic pathway, under quite general conditions, a summation principle holds in which the sum of the control coefficients must sum to unity (Kacser & Burns, 1973; Dean, 1994). Hence, not every enzyme in a metabolic pathway can have a large control coefficient, and the greater the value for any one enzyme, the smaller the sum of the values for the remaining enzymes.

To the extent that flux through metabolism can serve conceptually as a surrogate for fitness, the implication of Fig. 5(a) for molecular evolution is that a small change in enzyme activity produces a disproportionately small change in fitness (Hartl *et al*., 1985; Dean *et al*., 1986). The magnitude of the effect depends on the initial value of the enzyme activity. The realized value of activity for the standardized hyperbola in Fig. 5(a) has been estimated for 14 enzymes in *Drosophila* based on the frequency of null alleles in natural populations, assuming mutation-selection balance and partial dominance (Hartl *et al*., 1985). The estimated values ranged from 383 to 5880 with a median of 1203, which correspond to control coefficients with respect to fitness of 2.6×10^{-3}, 1.7×10^{-4}, and 8.3×10^{-4}, respectively. For these values, a decrease in the enzyme activity of 1% results in a decrease in fitness of 2.6×10^{-5}, 1.7×10^{-6}, and 8.4×10^{-6}, respectively. Because of the small control coefficient, even a change as large as 1% in enzyme activity is expected to have a very small effect on fitness.

One may well object that fitness is not always proportional to metabolic flux and hence that the saturation effect in Fig. 5(a) is not universal. A natural alternative is to suppose that fitness, expressed as a function of flux, has an intermediate optimum of the sort in Fig. 5(b). The selective effect of a given change in enzyme activity then depends not only on the control coefficient of the enzyme but also on the shape of the optimization curve. Unless the optimization curve is sharp and narrow, there must exist a region around the optimum within which a small change in flux results in approximately a proportional change in fitness. The size of the region is determined by the details of the optimization curve but, within this region, a small change in the activity of an enzyme with a small flux control coefficient will also result in a small change in fitness (Dean *et al*., 1988).

Among mutations with very small effects, Fisher convinced himself that favorable and detrimental mutations should be about equally frequent, based on the following analogy:

> The conformity of these statistical requirements [for adaptation] with common experience will be perceived by comparison with 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 affected by the maker or the operator, the chance improvement should be almost exactly one-half (Fisher, 1930, p. 40–41).

One does not need to accept the figure of one-half as exact to appreciate the evolutionary implication of nearly neutral mutations. Since all the curves in Fig. 4 go to one as $r\sqrt{n/d}$ decreases to zero, any appreciable fraction of slightly detrimental mutations implies that some of them are destined to be fixed unless α is very large. The result is that a molecule evolves in a pattern of continual fixation of one or more slightly detrimental mutations interspersed with one or more compensatory favorable mutations. A molecule subjected to such a process over a long period of evolutionary time is always evolving but never improving. Except for the occasional strongly favored mutation, each substitution, by itself, has an effect on fitness that is of the same order of magnitude or smaller than the reciprocal of the effective population number. Accumulated over time, the aggregate effects of the substitutions may be substantial but, in the absence of special experiments, the effects are hidden from the observer owing to their being compensated by offsetting changes either within the same molecule or in macromolecules with which it interacts.

Discussion

We have argued that, unless a gene is highly constrained, a significant fraction of nucleotide

substitutions should be slightly detrimental. In the long run, these substitutions are compensated by favorable mutations. Given sufficient time, this process appears almost inevitable, barring strong selective constraints, but it has some perverse implications. First, it suggests that, on a long enough timescale, most genes undergo selectively driven nucleotide substitutions, though not owing to adaptation to external conditions but rather to compensation for deleterious mutations previously incorporated into the gene. This process makes it difficult to interpret statistical evidence for natural selection as being in any way related to the Darwinian notion of adaptive evolution (Brookfield & Sharp, 1994). A second implication is that a gene whose function remains unchanged may nevertheless undergo selectively driven nucleotide substitutions, hence one should be able to find evidence that genes that carry out the same function in different species are not functionally fully equivalent in the sense of being able to carry out complete genetic complementation in interspecific transformation experiments. A third implication of the process is that, to the extent that many of the compensatory changes will take place at different sites in the same gene, chimeric proteins containing regions from each of two homologous proteins in different species will be found to function less effectively than either of the parental proteins.

It is of some satisfaction to enumerate counterintuitive expectations in a paper dedicated to the memory of Henrik Kacser. His scientific life was guided by the hope of identifying simple and general principles by which features common to complex systems and living organisms could be understood. He enjoyed unexpected findings. Like the White Queen, he could sometimes believe as many as six impossible things before breakfast. A few years ago, he sent one of us (DLH) a reprint of a paper entitled ''A universal method for achieving increases in metabolite production'' (Kacser & Acerenza, 1993). Enclosed was this handwritten note:

> Here is my latest paper. The problem, which has been a constant thorn in my flesh, has turned out to be rather simple. Perhaps now the biotechnologists will pay attention to my 20 year struggle to convince them that only by taking a systemic approach can you hope to understand (and predict) the consequences of genetic interaction.

We gratefully acknowledge the support of grants from the National Institutes of Health (DLH) and the National Science Foundation (CHT). Thanks also to Professor A. W. F. Edwards for his thoughts on Fisher's model of adaptation.

REFERENCES

Brookfield, J. F. Y. & Sharp, P. M. (1994). Neutralism and selectionism face up to DNA data. *Trends Genet*. **10,** 109–111.

Crow, J. F. & Kimura, M. (1970). *An Introduction to Population Genetics Theory*. New York: Harper & Row.

Dean, A. M. (1994). Fitness, flux and phantoms in temporally variable environments. *Genetics* **136,** 1481–1495.

Dean, A. M., Dykhuizen, D. E. & Hartl, D. L. (1986). Fitness as a function of β-galactosidase activity in *Escherichia coli*. *Genetics* **48,** 1–8.

Dean, A. M., Dykhuizen, D. E. & Hartl, D. L. (1988). Theories of metabolic control in quantitative genetics. In: *Proc*. 2*nd Int*. *Conf*. *Quantitative Genetics* (Weir, B. S., Eisen, E. J., Goodman, M. M. & Namkoong, G., ed). pp. 536–548. Sunderland, MA: Sinauer.

Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford: Oxford University Press.

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

Kacser, H. & Acerenza, L. (1993). A universal method for achieving increases in metabolite production. *Eur*. *J*. *Biochem*. **216,** 361–367.

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

Kacser, H. & Burns, J. A. (1979). Molecular democracy: who shares the controls? *Biochem*. *Rev*. **7,** 1150–1160.

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

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

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge University Press.

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. & Tachida, H. (1990). Theoretical study of near neutrality. I. Heterozygosity and rate of mutant substitution. *Genetics* **126,** 219–229.