Letter to the Editor

The Effect of Constraint on the Rate of Evolution in Neutral Models With Biased Mutation

One of the central axioms of the neutral theory of molecular evolution is that the rate of evolution of a sequence is inversely related to its importance (Kimura and Ohta 1974; Kimura 1983). The reason is simply that important sequences have few potential neutral alleles. This prediction was made assuming that the pattern of mutation is uniform. However, as I will show below there are biased mutation patterns under which a reduction in the number of potential neutral alleles can lead to an increase in the rate of evolution. The approach taken is to consider the rate of evolution at single sites, comparing the rates at sites which have two, three or four neutral alleles (with all other alleles being very deleterious). Note that a site with one neutral allele does not undergo substitution.

Following Wright (1969, chapter 3) let us consider a single site in a DNA molecule at which n alleles can segregate. Normally n will have a maximum value of four corresponding to the four bases which can occupy a site. Let the frequency of the ith allele be $f_i$, the probability of a mutation from allele i to j be $U_{ij}$ and the probability of fixing a newly arising j mutant in a population of i be $P_i$. Let $K_j = 2NU_{ij}P_i$ where N is the population size of a diploid organism. If we assume $NU_{ij} \ll 1$, then $K_j = U_{ij}$ (Kimura 1968). Furthermore the population will generally be monomorphic for one of the alleles and the expected frequency of allele i ($\bar{f}_i$) (i.e., the relative amount of time for which the population is fixed for i) can be obtained from a consideration of the flux between the n alleles: i.e., by solving $n - 1$ simultaneous equations of the form:

$$\Delta f_i = -f_i \sum_{j \neq i} K_{ij} + \sum_{j \neq i} f_j K_{ji} = 0.$$  \hspace{1cm} (1)

The average rate of substitution at the site is then

$$R_s = \sum_i \bar{f}_i \sum_{j \neq i} K_{ij}.$$  \hspace{1cm} (2)

The equilibrium frequencies for two, three and four allele systems are given by Wright (1969, chapter 3). As one might expect the expressions for $R_3$ and $R_4$ are complex and do not yield readily to further analysis. However, there is an informative simplification that can be made. If we consider the mutation pattern symmetrical about allele 4 such that $U_{12} = U_{21} = U_{13} = U_{31} = U_{23} = U_{32} = U$, $U_{14} = U_{41} = U_{24} = U_{42} = U_{34} = U_{43}$ and $U_{44} = U_4$, the rates of evolution in the following systems of two, three and four neutral alleles become:

$$R_4 = \frac{6U_4(U + U_4)}{3U_4 + U_4}$$  \hspace{1cm} (3a)

$$R_3(*,*,4) = \frac{2U_4(U + 2U_4)}{2U_4 + U_4}$$  \hspace{1cm} (3b)

$$R_3(*,*,*) = 2U$$  \hspace{1cm} (3c)

$$R_4(*,*) = U$$  \hspace{1cm} (3d)

where * refers to one of alleles 1, 2 or 3. Using the expressions in (3) it is simple to show that

$$R_3(*,*,*) > R_4 \text{ when } U > 3U_4.$$  \hspace{1cm} (4a)

$$R_4(*,*) > R_4 \text{ when } U > \frac{6U_4U_4}{U_4 - 3U_4}.$$  \hspace{1cm} (4b)

and $U_4 - 3U_4 > 0$

$$R_3(*,*,4) > R_3(*,*,*) \text{ when } U > 4U_4.$$  \hspace{1cm} (4c)

In other words, there are mutation patterns under which a reduction in the number of potential neutral alleles at a site leads to an increase in the average rate of evolution. Note, however, how biased the mutation patterns must be for this to occur. Essentially the elimination of allele 4 in each case leads to an increase in the time for which the site is occupied by alleles 1, 2 and 3. So if the mutation rate between alleles 1, 2 and 3 is high, elimination of allele 4 increases the overall flux. As an extreme example, consider a mutation pattern which is very biased so the site is almost permanently fixed for allele 4. Clearly there can be little substitution. Removal of allele 4 from the system allows the other alleles to occupy the site alternately thereby increasing the rate of substitution at the site. Of course eliminating an allele from a system also removes several mutation pathways (e.g., those between alleles 1 and 4). This explains why the conditions under which three-allele systems evolve faster than four-allele systems are rather less stringent than those for other comparisons.

It turns out that some of the relationships can be applied to mutation patterns in general, not just those in which $U_{12} = U_{21}, \ldots, \text{etc.}$, by using average mutation rates. Thus if we define
alleles may be brought about by an increase in the different alleles may be removed from different sites: site. However, a three-allele site must contain allele 4 to form the three-allele site. Of course the mutation patterns must be very biased for this to occur; although if such patterns did exist they would overcome the partial exclusivity of inequalities 4a and 4c.

Hence in both sequences and sites, a reduction in the number of potential neutral alleles can increase the rate of evolution. However, this does depend on there being extreme bias in the mutation pattern. Although it is unclear what mutation patterns one might expect to find in the natural world, the very limited data available suggest that the patterns are not very biased (Gojobori, Li and Graur 1982; Li, Wu and Luo 1984).

The present findings suggest some caution should be exercised in deducing the action of positive natural selection when the rate of substitution at nonsilent sites exceeds that at silent sites, especially when few sites are involved or there is extreme codon bias; codon bias being an indicator of biased mutation.

I am very grateful to John Brookfield, Eddy Holmes, Andy Leigh Brown, Ian Hastings, Bill Hill and W.-H. Li for helpful discussions and criticism. I also thank the Science and Engineering Research Council for their financial support.

LITERATURE CITED


