Neutral mutation hypothesis test

SIR — McDonald and Kreitman[1] claim that adaptive mutations are largely responsible for the evolution of alcohol dehydrogenase (Adh) because, according to their calculations, in the Adh gene the ratio of nonsynonymous to synonymous substitutions between three *Drosophila* species (7:17) is much larger than the ratio (2:42) within species. However, their test has at least the following problems.

First, McDonald and Kreitman neglect the between-species variation at a nucleotide site, if that site is polymorphic in any of the three species. Both this counting rule and the one we discuss below tend to underestimate between-species variation. For example, at position 1,590 all the alleles in *D. melanogaster* have T, half the alleles in *D. simulans* have T and the other half have C, and all the alleles in *D. yakuba* have C. At this position, they assigned a single polymorphic synonymous substitution within *D. simulans* but no substitution between species. However, under their assumption of monomorphism in the ancestral species, an additional synonymous substitution must have occurred between *D. melanogaster* and *D. yakuba*.

Second, they fail to consider the fact

REPLACEMENT (NONSYNONYMOUS) AND SYNONYMOUS DIFFERENCES IN THE CODING REGION OF THE ADH GENE

Type of difference	Within species			Between species*		
				mel vs *sim*	*mel* vs *yak*	*sim* vs *yak*
	mel	*sim*	*yak*			
Replacement (*R*)	0.71	0.00	0.00	2.58 (2.22)	6.58 (6.22)	6.00 (6.00)
Synonymous (*S*)	5.54	4.60	3.16	10.67 (5.60)	28.10 (23.75)	24.29 (20.41)
R/*S* ratio	0.13	0.00	0.00	0.24 (0.40)	0.23 (0.26)	0.25 (0.29)

*Numbers in parentheses denote between-species differences after subtraction of within-species differences.

mel, *D. melanogaster*; *sim*, *D. simulans*; *yak*, *D. yakuba*. (The 12 alleles in *D. yakuba* in Table 2 of ref.1 probably represent 24 alleles, and this has been assumed in our analysis.)

that a polymorphism in a species does not necessarily imply that the polymorphism arose within that species. In particular, because *D. simulans* and *D. melanogaster* are closely related sibling species, a polymorphism in one species may also indicate that the ancestral species was polymorphic, and that polymorphism was subsequently lost in one of the descendent species.

Third, their decision as to whether or not a site is polymorphic is highly sensitive to the number (n) of sequences sampled and n is only 6 for *D. simulans*, yet their test neglects sampling errors. Thus, even if one ignores all other problems, their test is still not acceptable. If the numbers of fixed and polymorphic nonsynonymous substitutions are 6 and 3 instead of 7 and 2, then Fisher's exact test gives a probability of 7% (nonsignificant). Thus, the conclusion is sensitive to sampling errors.

A simple way to overcome these problems is as follows. First, compute the number of substitutions per synonymous site (K_s) and per nonsynonymous site (K_A) between every pair of sequences[2], both within- and between species. Second, for each species, compute the average within-species K_A and K_S values and the K_A/K_S ratio. Third, compute the average between-species K_A, K_S and K_A/K_S values for every pair of species. Fourth, compare the between-species K_A/K_S ratios with the within-species ratios. In the present case, however, the sequences from *D. yakuba* were obtained by directly sequencing the polymerase chain reaction products, so that the two alleles in an individual at the ADH locus cannot be clearly defined when they differ by more than one nucleotide. For this reason, we use the following slightly less rigorous method. We compute the numbers of synonymous and nonsynonymous nucleotide differences (S and R, respectively) between every pair of sequences at each of the sites listed in Table 1 of ref. 1. For instance, consider position 816. In *D. melanogaster*, five alleles have T and seven have G. Because the difference between T and G at this position is synonymous, $R = 0$ for every sequence pair, and the average S value is $35/66 = 0.53$, because $5\times7 = 35$ out of $12\times11/2 = 66$ possible pairwise comparisons have the G⟷T difference. In *D. simulans*, all six alleles have T. Thus, between *D. melanogaster* and *D. simulans*, $R = 0$ and the average S is $7\times6/12\times6 = 0.58$. The R and S values are then summed over all sites (see table).

From the total R and S values, we find that the R/S ratios between species are remarkably similar among the three comparisons: 0.24, 0.23 and 0.25, the average being 0.24. If we assume that the ratio found in the between-species comparisons also holds for the within-species comparisons, then the expected within-species R values are 1.33, 1.10, and 0.74 for *D. melanogaster*, *D. simulans*, and *D. yakuba*, respectively. The differences between the expected values and the observed ones (0.71, 0 and 0) are not large enough to be statistically significant, for the variance of R is given by $R(n+1)/[3(n-1)] + 2R^2(n^2+n+3)/[9n(n-1)]$ (ref. 3).

In the table we also compute the net between-species differences using the formula $d = d_{xy} - a\,(d_x+d_y)/2$, where d_{xy} denotes the differences between species X and Y, and d_x denotes the within-species differences in species X. Note that both the synonymous and replacement differences between *D. yakuba* and either of the other two species are much larger than the within-species values; for example, $S>20.0$ versus $S<6.0$. This is in sharp contrast to McDonald and Kreitman's calculation that only 17 synonymous substitutions had occurred between species whereas 42 had occurred within species. For the net between-species differences the R/S ratios are now 0.40, 0.26 and 0.29 with an average of 0.32 (we take the simple average, though the last two ratios are more favourable for our argument, and would be more reliable, for they are based on larger R and S values). With this ratio, the expected within-species R values are 1.77, 1.47 and 1.01 for *D. melanogaster*, *D. simulans*, and *D. yakuba*, respectively. The total expected R value (4.25) for the three species differs from the total observed value (0.71) by 3.56, which is smaller than 1.96 times the standard error 1.86, which is an underestimate, for it neglects the positive evolutionary correlations between Rs and the errors in estimating the expected R/S ratio and the within-species S value. Therefore, the neutral hypothesis cannot be rejected, even under unfavourable assumptions.

A more powerful test is to use Watterson's theory[4]. For example, for *D. melanogaster*, $n = 12$ and $k = 14$ (number of segregating sites in the sample) for synonymous sites and so from formula (1.4a) of Watterson we obtain $\theta_S = 4.64$. If $R/S = 0.32$ $\theta_R = 1.48$ for replacement (nonsynonymous) mutations and from formulas (1.4a) and (1.4b) we obtain the mean and variance of k as 4.48 and 7.91 for nonsynonymous sites. The corresponding mean and variance are 3.52 and 7.00 for *D. simulans* and 5.76 and 9.57 for *D. yakuba*. The expected k values are not significantly different from the observed values (2, 0 and 0). If we assume that the three species are independent and pool the data together, then the total k is 13.76 ± 4.95. Taken at face value, this expected k differs significantly from the observed value 2 (at the 2% level). However, the significance is uncertain because the estimate of s.e. neglects all the errors mentioned in the preceding paragraph. If $R/S = 0.24$, then k becomes 10.3 ± 4.04, which is not significantly different from the observed value if the above-mentioned errors are included in the estimation of s.e.

In conclusion, it is not clear as to whether the ADH data can be taken as evidence against the neutral hypothesis.

DAN GRAUR
Department of Zoology, Tel Aviv University, Ramat Aviv 69978, Israel

WEN-HSIUNG LI*
Center for Demographic and Population Genetics, University of Texas, PO Box 20334, Houston, Texas 77225, USA

* To whom correspondence should be addressed.

SIR – Comparing nucleotide sequences of the alcohol dehydrogenase (Adh) gene within and between three species of *Drosophila*, McDonald and Kreitman[1] concluded that the number of nonsynonymous (amino-acid altering) nucleotide substitutions is significantly greater than expected under neutral pro-

ceases and that the excess substitutions result from fixation of selectively advantageous mutations. This conclusion is based on a statistical test of the prediction that under neutrality "the ratio of replacement [nonsynonymous] to synonymous fixed substitutions should be the same as the ratio of replacement to synonymous polymorphisms." We believe that there are subtle but serious problems in McDonald and Kreitman's reasoning.

The prediction as derived from the neutral theory can be more precisely stated as a null hypothesis as follows: for strictly neutral mutations, the ratio of nonsynonymous to synonymous substitutions between species is equal to the ratio of nonsynonymous to synonymous substitutions between alleles within species. A proper test of this hypothesis requires estimates of the number of nucleotide substitutions per site that have occurred during the divergence of alleles within species and between species. However, the *G*-test used by McDonald and Kreitman does not compare rates of nucleotide substitution but instead is based on the ratio of synonymous to nonsynonymous sites that have been classified as polymorphic within or fixed between species. Such a classification of sites is arbitrary because fixation (or polymorphism) at a site in a population is a transient event in evolution. In addition, the number of polymorphic and fixed sites depends, to an unknown extent, on both number of sequences examined as well as the number of species studied. This second point is illustrated, for example, by site 816 in Table 1 of ref. 1, which would have been a fixed site if only *D. simulans* and *D. yakuba* were studied, but was classified as a polymorphic site because it was polymorphic in *D. melanogaster*. Although these effects may have only minor influence on the ratio of nonsynonymous to synonymous sites that are polymorphic or fixed in a sample, there is virtually no mathematical theory for predicting the distribution of the number of polymorphic and fixed sites across species under the neutral hypothesis. (Hey recently studied[3] the sampling distribution of fixed differences between two species.) Finally, since both the number of fixed and polymorphic sites are subject to large stochastic errors[4,5], the stochastic variances of these quantities should be taken into account in a valid test of the difference between the two ratios.

We suggest that as a general test of the null hypothesis of equal ratios one should evaluate the average nucleotide substitutions for all pairwise comparisons of sequences within and between species (π and d_A, respectively)[6]. Our test uses the frequency of nucleotides in the sample, accounts for multiple substitutions at sites, and incorporates stochastic errors in the evolutionary process. We first reanalysed McDonald and Kreitman's data for *D. melanogaster* and *D. yakuba*. We estimated d_A and the average π for the two species for both synonymous and nonsynonymous sites using the Nei–Gojobori method[6]. The standard errors of these quantities were computed by taking into account stochastic variances[6,7]. The results obtained were $d_A = 14.20 \pm 3.71$ and $\pi = 2.40 \pm 1.01$ per 100 synonymous sites and $d_A = 1.1 \pm 0.44$ and $\pi = 0.06 \pm 0.05$ per 100 nonsynonymous sites. Therefore, the ratio of nonsynonymous to synonymous substitutions is 0.077 ± 0.037 between species and 0.026 ± 0.024 within species. Although the former ratio is higher than the latter, the difference is not statistically significant ($Z = 1.2$; $P > 0.2$). (The smaller ratio within species could be partly due to excess synonymous substitutions caused by the balanced polymorphism of the *F* and *S* alleles in *D. melanogaster*[8].) The analysis for three species, which is somewhat more complicated, gives essentially the same results. Thus, these results do not support the conclusion that there is a significant excess of nonsynonymous substitutions resulting from adaptive fixation of mutations.

THOMAS S. WHITTAM
MASATOSHI NEI
Institute of Molecular Evolutionary Genetics,
Department of Biology,
Pennsylvania State University
University Park,
Pennsylvania 16802, USA

MCDONALD AND KREITMAN REPLY — Graur and Li, and Whittam and Nei, point out that a polymorphism in one species may have arisen in an ancestral species; that different polymorphic sites have different allele frequencies; and that the number of polymorphisms will increase as more alleles and more species are sampled. For the neutral model of molecular evolution that we tested, all of these phenomena will affect neutral replacement substitutions and neutral synonymous substitutions equally. Therefore they do not affect the validity of our test of the neutral model, and they are not alternatives to adaptation as explanations for the ratio of replacement to synonymous substitutions being much greater for fixed differences than for polymorphisms at the *Adh* locus in three species of *Drosophila*[1].

The authors also question our method for counting a site which has one nucleotide fixed in one species, a different nucleotide fixed in a second species, and both nucleotides polymorphic in a third species. They suggest that we count such a site as one fixed difference and one polymorphism, rather than just as one polymorphism. Any rule for classifying substitutions as fixed or polymorphic will affect neutral replacement and neutral synonymous substitutions equally, and it is only important to apply the same rule to both. We choose to count as fixed substitutions only those that are fixed in every species in which they appear. This is because under the alternative hypothesis of adaptive fixation of replacement mutations, a replacement substitution that is adaptive in one species, and thus rapidly becomes fixed, will either be adaptive or maladaptive in other species, and thus is unlikely to be polymorphic in any species. There is also a practical reason for counting each substitution only once, rather than trying to estimate the number of times that the substitution has gone to fixation. Estimating the number of fixation events would require an accurate estimate of the species' phylogeny; we think it is an advantage of our test that it requires no such estimate.

The authors of the above letters suggest tests based on the gene diversity between and within species, rather than the numbers of fixed and polymorphic substitutions. These tests use estimates of the stochastic variance, which is the variation in gene diversity among loci resulting from the different coalescent times (times to the most recent common ancestor) of different loci[3,7]. For a single locus, however, replacement substitutions are intermingled with synonymous substitutions, and thus replacement and synonymous substitutions have the same coalescent time. Therefore only sampling variance needs to be considered. It would be interesting to see the results of a test using gene diversity that used the sampling variance, which is much smaller than the stochastic variance. Whatever the outcome of such a diversity-based test, however, our substitution-counting test remains valid. We suspect that any diversity-based test will be more complicated, will require more assumptions, and will be less statistically powerful than our method for detecting adaptive protein evolution.

JOHN H. MCDONALD
MARTIN KREITMAN
Department of Ecology and Evolutionary Biology,
Princeton University,
Princeton, New Jersey 08544, USA

1. McDonald, J. H. & Kreitman, M. *Nature* **351**, 652–654 (1991).
2. Li, W.-H., Wu, C.-I. & Luo, C.-C. *Molec. biol. Evol.* **2**, 150–174 (1985).
3. Tajima, F. *Genetics* **105**, 437–460 (1983).
4. Watterson, G. A. *Theor. Pop. Biol.* **7**, 256–276 (1975).
5. Hey, J. *Genetics* **128**, 831–840 (1991).
6. Nei, M. *Molecular Evolutionary Genetics* (Columbia University Press, New York, 1987).
7. Takahata, N. & Nei, M. *Genetics* **116**, 325–344 (1985).
8. Kreitman, M. in *Evolution at the Molecular Level* (eds Selander, R. K., Clark, A. G. & Whittam, T. S.) 204–221 (Sinauer, Sunderland, Massachusetts, 1991).