Proc. Natl. Acad. Sci. USA
Vol. 79, pp. 550–554, January 1982
Genetics

Genetic variation in natural populations: Problem of electrophoretically cryptic alleles

(evolution/population genetics/polymorphism/peptide analysis/charge-state model)

FRANCISCO J. AYALA

Department of Genetics, University of California, Davis, California 95616

Contributed by Francisco J. Ayala, September 28, 1981

ABSTRACT **Electrophoretic studies have shown that the average frequency of heterozygous loci per individual is about 12% in *Drosophila* and other invertebrates and about 6% in vertebrates. It is estimated that only about two-thirds of all amino acid substitutions change net electric charge; hence, a large fraction of all genetic variation may be undetected by electrophoresis. Peptide mapping of 11 independent alleles coding for alcohol dehydrogenase in *Drosophila melanogaster* has uncovered one cryptic variant; thus, the frequency of electrophoretically cryptic variation is apparently low, about 9% in this sample. Nevertheless, with a simple model it is shown that this degree of cryptic variation, if it is typical of other loci, would substantially change our current estimates of genetic variation: the average heterozygosity would increase from about 12% to about 25% for invertebrates and from about 6% to 21% for vertebrates.**

A variety of techniques—including sequential electrophoresis and heat or urea denaturation—have been used by various investigators to detect electrophoretically cryptic variation. These techniques appear to be less effective than peptide mapping for detecting cryptic variation, but, like peptide mapping, they suggest that standard electrophoresis may detect most of the protein variation present in natural populations.

The charge-state model of protein variation proposes that the "alleles" detected by electrophoresis are extremely diverse classes consisting of many electrophoretically cryptic alleles. The alcohol dehydrogenase peptide-mapping results are inconsistent with the charge-state model.

Electrophoretic techniques have uncovered enormous amounts of genetic variation in natural populations: in most sexually reproducing organisms, an individual is heterozygous at 5–20% of its loci (1). It has been proposed that electrophoresis detects allelic variants only when the encoded proteins differ in net electric charge and that only about 33% of all possible amino acid substitutions (or 25% of all single-base replacements in the translated DNA) change the net electric charge of a protein (2). Whether or not these claims are correct, it remains an open question how much genetic variation exists in natural populations above the estimates obtained by electrophoresis.

Another open issue concerns the processes that maintain the pervasive genetic variation already known. Kimura and Ohta (3) have proposed that most of the protein variation observed in natural populations is adaptively neutral—the polymorphisms represent a transitional phase in molecular evolution going from fixation for one allele to fixation for another adaptively equivalent allele as a consequence of random drift. The model presented in (3) assumes that the number of possible alleles is effectively infinite, an assumption that is not met by electrophoretically detected variation. A modified version of the neutralist hypothesis is the "charge-state model," which postulates that there is a limited number of detectable allelic states, differentiated from each other by one or more electric charges, and that each detectable allelic class consists of a large number of electrophoretically "cryptic" alleles, all of which code for proteins with identical electrophoretic mobility but have different amino acid sequences (2, 4–8). According to the charge-state model, no allele within a given electrophoretic class (or "electromorph") is expected to exist in high frequency, but rather there should be many alleles, each in low frequency.

I herein advance a simple model in order to explore the effect of electrophoretically cryptic variants on current estimates of genetic variation. Then, I examine the available empirical data and conclude that (*i*) cryptic variation seems to be considerably less extensive than it is claimed by the calculations quoted above, but (*ii*) nevertheless the resulting increase in the estimates of genetic variation in natural populations is not trivial. Finally, I argue that the existing evidence is inconsistent with the charge-state model.

THE MODEL

Let x_i represent the frequencies of electrophoretically detectable alleles at the k locus; $\Sigma x_i = 1$. Assuming Hardy–Weinberg equilibrium, the values of the parameters that measure electrophoretic variation are

$$\text{frequency of homozygotes: Hom} = \Sigma x_i^2$$

$$\text{frequency of heterozygotes: Het} = 1 - \Sigma x_i^2$$

$$\text{effective number of alleles: } n_e = 1/\Sigma x_i^2.$$

First, I shall assume that each electrophoretic allele is actually a class consisting of two alleles, with frequencies a and b, coding for protein sequences that differ by at least one amino acid; $a + b = 1$. I shall also assume for simplicity that the frequencies of the two alleles in each electrophoretic class are the same for all electromorphs. The values of the parameters that measure *total variation*, electrophoretic as well as cryptic, are:

$$\text{Hom}' = (a^2 + b^2)\,\Sigma x_i^2$$

$$\text{Het}' = 1 - (a^2 + b^2)\,\Sigma x_i^2$$

$$n_e' = 1/(a^2 + b^2)\,\Sigma x_i^2 = n_e/(a^2 + b^2).$$

The increase in genetic variation due to cryptic variants may be measured in two ways. First, one may obtain the difference between total heterozygosity and electrophoretic heterozygosity:

$$\text{Het}' - \text{Het} = [1 - (a^2 + b^2)]\,\Sigma x_i^2. \qquad [1]$$

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviation: ADHase, alcohol dehydrogenase.

Second, one may obtain the ratio of the effective numbers of alleles:

$$n_e'/n_e = [n_e/(a^2 + b^2)]/n_e = 1/(a^2 + b^2). \quad [2]$$

Eq. 2 measures the increase in genetic variation better than Eq. 1 for two (related) reasons. The first reason is that Eq. 2 reflects more accurately the actual increase in variation. This can be shown with an example. Consider the change from two alleles, each with a frequency of 0.5, to four alleles, each with a frequency of 0.25. The increase in heterozygosity is from 0.50 to 0.75, and the difference is Het′ − Het = 0.25, or a 50% increase over the original value. But going from two alleles with equal frequencies to four alleles with equal frequencies is a doubling of the genetic variation. This is reflected in the ratio between the effective numbers of alleles: $n_e'/n_e = 4/2 = 2$.

The second reason is that the increase in heterozygosity achieved by a given amount of cryptic variation is a function of the preexisting amount of genetic variation. For any given a and b, the difference Het′ − Het (Eq. 1) will be greatest, namely $1 - (a^2 + b^2)$, when a locus has no electrophoretically detectable variation ($\Sigma x_i^2 = 1$), and the greater the electrophoretic variation at a locus, the smaller the difference will be. This bias is not present in Eq. 2: the value of n_e'/n_e depends only on the frequency of the cryptic alleles.

The assumption that there are only two cryptic alleles in each electrophoretic class may be abandoned by assuming instead that each electromorph consists of one allele with frequency a and many other alleles, each with very low frequency. This would occur if many different amino acid sequences exist within each electrophoretic class, but [as is the case for alcohol dehydrogenase (ADH) in *Drosophila melanogaster*] one is quite common in the population. Represent the frequency of the rare cryptic variants as b_j, so that $\Sigma b_j = b$; $a + b = 1$. The values of the parameters that measure total variation are:

$$\text{Hom}'' = (a^2 + \Sigma b_j^2)\,\Sigma x_i^2 \simeq a^2\,\Sigma x_i^2$$

$$\text{Het}'' = 1 - (a^2 + \Sigma b_j^2)\,\Sigma x_i^2 \simeq 1 - a^2\,\Sigma x_i^2$$

$$n_e'' = 1/(a^2 + \Sigma b_i^2)\,\Sigma x_i^2 \simeq n_e/a^2$$

$$\text{Het}'' - \text{Het} = [1 - (a^2 + \Sigma b_j^2)]\,\Sigma x_i^2 \simeq (1 - a^2)\,\Sigma x_i^2$$

$$n_e''/n_e \simeq 1/a^2.$$

As before, the smaller the amount of electrophoretically detectable variation, the greater the difference between the two heterozygosities. The maximum possible value of this difference will be $1 - a^2$, when the locus is monomorphic by electrophoresis.

Other models are possible. For example, it might be the case that within each electrophoretic class there are several-to-many cryptic alleles, all in about equal frequencies. However, the ADH data given below suggest that there is one common allele within each electrophoretic class.

RESULTS

The amount of electrophoretically cryptic variation at a given locus could be estimated by obtaining the amino acid sequence of each electrophomorph from a sufficiently large number of individuals. Clearly, this is not presently feasible because the required investment of time and other resources is enormous. Peptide mapping, or "fingerprinting," of protein digests is a less expensive method, although it may not give full resolution (because some amino acid substitutions in insoluble peptides or in large ones may be undetected).

ADH (alcohol:NAD$^+$ oxidoreductase, EC 1.1.1.1) from *D. melanogaster* has been examined in our laboratory by fingerprinting (9, 10). Two electromorphs, ADH^F (or, simply, F for "fast") and ADH^S (or S for "slow"), occur both with high frequencies in natural populations of *D. melanogaster*. Strains were made homozygous by descent, so that each strain would carry only one allele sampled from a natural population. A total of 11 strains were examined. Six strains carried the F electromorph: four were derived from flies collected in Napa County, California; one was derived from flies in Redwood City, California; and one was from Cedar Rapids, Iowa. The other five strains carried the S electromorph and were all derived from the Napa County population. Tryptic digests of purified enzyme from each strain were fingerprinted.

The two relevant results are as follows (9, 10). First, the difference between the two electromorphs involves in all cases the same substitution, threonine in ADH^F to lysine in ADH^S. This substitution occurs in residue 192. (The total sequence consists of 254 residues; see ref. 11). Second, only 1 electrophoretically cryptic substitution has been detected among the 11 allele products: the 5 S strains give identical peptide maps; among the F strains, those from Redwood City, from Cedar Rapids, and three from Napa County are identical to each other but differ from the fourth Napa County strain by at least one amino acid substitution.

A sample of 11 allele products is small. Moreover, there may be more variation than has been detected by peptide mapping because the peptide maps contain only about 50% of the ADH sequence; the other half of the molecule is insoluble in the buffers used (12). Be that as it may, the frequency of electrophoretically cryptic variants in the sample is 1/11 = 0.091. This is the observed value of b in the model proposed above; $a = 0.909$. The 95% binomial confidence limits of these estimates are $b = 0.0023$–0.4045 and $a = 0.5955$–0.9977.

Table 1 shows the increase in total genetic variation obtained when the frequency of cryptic variants is as observed for ADH. The table uses three values of a (and b): the observed frequency and each of the two confidence bounds. The calculations are made under the two assumptions considered above: (*i*) when each electromorph consists of only two cryptic alleles and (*ii*) when there is one common cryptic allele but many rare ones. It can be seen that the results under these two assumptions are fairly similar whenever the frequency, a, of the common allele is high (0.90 or greater). In general, the lower the frequency of a, the greater the increase in variation obtained under the second assumption.

The frequencies of the two electromorphs, ADH^F and ADH^S, in natural populations of *D. melanogaster* are about 0.7 and 0.3. Thus, for electrophoretically detectable variation, Het $\simeq 0.420$ and $n_e \simeq 1.72$. If we use, as the best guess, $a = 0.91$, we obtain under assumption (*i*): Het′ = 0.516, $n_e' = 2.07$; and

Table 1. Total genetic variation as a function of the frequency a of the most common cryptic allele*

	Total variation		Increase in variation	
a	Hom′	Het′	Het′ − Het	n'_e/n_e
With only two alleles per electromorph:				
0.9977	$0.995\,\Sigma x_i^2$	$1 - 0.995\,\Sigma x_i^2$	$0.005\,\Sigma x_i^2$	1.005
0.9091	$0.835\,\Sigma x_i^2$	$1 - 0.835\,\Sigma x_i^2$	$0.165\,\Sigma x_i^2$	1.198
0.5955	$0.518\,\Sigma x_i^2$	$1 - 0.518\,\Sigma x_i^2$	$0.482\,\Sigma x_i^2$	1.930
With many alleles per electromorph:				
0.9977	$0.995\,\Sigma x_i^2$	$1 - 0.995\,\Sigma x_i^2$	$0.005\,\Sigma x_i^2$	1.005
0.9091	$0.826\,\Sigma x_i^2$	$1 - 0.826\,\Sigma x_i^2$	$0.174\,\Sigma x_i^2$	1.210
0.5955	$0.354\,\Sigma x_i^2$	$1 - 0.354\,\Sigma x_i^2$	$0.645\,\Sigma x_i^2$	2.825

* x_i is the allelic frequency and n_e is the effective number of alleles for electrophoretically detectable variants.

under assumption (*ii*): Het″ = 0.521, n_e'' = 2.09. The two assumptions yield fairly similar estimates, but in both cases there is a nontrivial increase in genetic variation—a difference of about 0.10 in heterozygosity and a 21% increase in the effective number of alleles.

Electrophoretic studies of 20 or more loci in many organisms give average heterozygosities per locus of about Het = 0.12 (n_e = 1.14) for *Drosophila* and other invertebrates, and Het = 0.06 (n_e = 1.06) for vertebrates such as humans. It is not presently known what the amount of electrophoretically cryptic variation detected by peptide mapping might be for the various loci. There seems to be no *a priori* reason why amino acid substitutions that do not entail electric charge changes would be more or less frequent in electrophoretically polymorphic loci than in electrophoretically monomorphic loci. In order to gain some insight about the effects that electrophoretically cryptic variation may have on current estimates of genetic variation, I shall assume that the frequency of cryptic alleles observed at the *Adh* locus (i.e., a = 0.909, b = 0.091) is typical for other loci. Then, the estimates of total genetic variation would be as follows (see Table 1): for *Drosophila* and other invertebrates, n_e' = 1.14 × 1.2 = 1.37 and Het′ = 0.27; for vertebrates, n_e' = 1.27 and Het′ = 0.21. Thus, the effective number of alleles per locus is 20% greater; the average heterozygosity becomes more than double for invertebrates and more than triple for vertebrates. The increase in heterozygosity is, of course, largely due to the contribution of cryptic variation in electrophoretically monomorphic or near-monomorphic loci, which change from zero or near-zero to 18% heterozygosity when the frequency of cryptic rare variants (b) is 0.10.

The calculations just made show that, if electrophoretically cryptic variants occur at most or all loci, our current estimates of genetic variation in natural populations may increase substantially. This would occur even if the cumulative frequency of the additional variants at each locus is relatively low—for example, of the order of 10% as assumed above.

DISCUSSION

A variety of techniques have been applied in recent years, mostly in *Drosophila*, in order to uncover variation undetected by electrophoresis. The techniques most extensively used are sequential electrophoresis, heat denaturation, and urea denaturation.

Electrophoretic studies of genetic variation usually use a single set of experimental conditions in the assay of a given enzyme. The method of sequential electrophoresis consists of applying a variety of conditions to a given enzyme. The conditions most often varied are gel concentration and buffer pH; typically from 6–10 different sets of conditions are used. Electromorphs that have identical mobility under a set of conditions may be distinguishable when the conditions are changed.

The species most extensively examined by sequential electrophoresis is *Drosophila pseudoobscura*. Table 2 summarizes some of the results obtained (some data have not been included because the statistics used in the table could not be calculated). The increase in variation detected ranges from zero, at a number of loci, to 76% at the *Est*-5 locus (n_e'/n_e = 1.76). In general, the more heterozygous a locus, the greater the increase in variation appears to be. This rule, however, has exceptions. For example, although *Adh* in *D. melanogaster* is highly heterozygous, no additional variation has been detected by sequential electrophoresis (17). For the 13 loci shown in Table 2, there is an average increase of 12% (n_e'/n_e = 1.12) in the amount of variation, with heterozygosity, designated hereafter simply as H, increasing by 0.04. The average heterozygosity for the 13 loci is greater than it has been observed in studies of *D. pseudoobscura* that include many loci ($\overline{H} \approx 0.10$). Thus, if the increase in variation detected by sequential electrophoresis is proportional to H, then the overall increase for random samples of electrophoretic loci would be even less than the averages given in Table 2. Be that as it may, the increase detected by sequential electrophoresis is relatively small.

Electromorphs with identical electrophoretic mobility may differ in thermostability. If the differences are shown by genetic tests to be associated with the locus itself, they may reflect different amino acid sequences in the encoded polypeptides. Satisfactory data exist for four enzymes in a single species, *D. melanogaster* (Table 3). All four loci are highly polymorphic, but the additional variation detected ranges from none (*αGpdh*) to 54% (*Est-6*). On the average, $\overline{H}$ increases from 0.410 to 0.485 and the effective number of alleles is 18% greater. The increase in variation is somewhat larger than with sequential electrophoresis, but this should be taken under advice, given that the species and the set of loci are different in the two studies.

Loukas *et al.* (21) have assayed by urea denaturation eight loci in *D. subobscura*. Differences in sensitivity to urea treatment appear to be associated with each locus under examination rather than with the genetic background. The results are sum-

Table 2. Increase in genetic variation observed by sequential electrophoresis in U.S. populations of *D. pseudoobscura*

	Standard conditions		All conditions		Increase in variation		
Locus	Het	n_e	Het′	n'_e	Het′ − Het	n'_e/n_e	Ref.
Est-5	0.645	2.8	0.798	4.9	0.153	1.76	(13)
Pt-8	0.51	2.04	0.55	2.22	0.04	1.09	(14)
*Ao**	0.499	2.0	0.584	2.4	0.085	1.20	(15)
Xdh	0.436	1.8	0.628	2.7	0.192	1.50	(15)
Odh	0.082	1.09	0.098	1.11	0.016	1.02	(15)
Hex-1	0.077	1.08	0.077	1.08	0.00	1.00	(16)
Pt-7	0.06	1.06	0.09	1.10	0.03	1.03	(14)
Pt-6	0.04	1.04	0.04	1.04	0.00	1.00	(14)
Mdh	0.00	1.00	0.00	1.00	0.00	1.00	(15)
Hex-2	0.00	1.00	0.00	1.00	0.00	1.00	(16)
Hex-6	0.00	1.00	0.00	1.00	0.00	1.00	(16)
Hex-7	0.00	1.00	0.00	1.00	0.00	1.00	(16)
αGpdh	0.00	1.00	0.00	1.00	0.00	1.00	(13)
Average	0.181	1.38	0.221	1.65	0.040	1.12	

* Labeled *Adh-6* in the original paper.

Table 3. Increase in genetic variation observed by heat denaturation in *D. melanogaster*

	Electrophoresis		Electrophesis plus heat		Increase in variation		
Locus	Het	n_e	Het′	n'_e	Het′ − Het	n'_e/n_e	Ref.
Est-6	0.48	1.93	0.67	2.98	0.19	1.54	(18)
Adh	0.48	1.94	0.50	2.00	0.02	1.03	(19)
Pgm	0.36	1.57	0.45	1.82	0.09	1.16	(20)
αGpdh	0.32	1.47	0.32	1.47	0.00	1.00	(19)
Average	0.410	1.73	0.485	2.06	0.075	1.18	

marized in Table 4. There is some association between the amount of electrophoretic variation and the increase in variation detected by urea, although *Pept-1* is a notable exception. The average increase in heterozygosity (0.077) is about the same as the average increase detected by thermostability in *D. melanogaster*, but the increase in effective number of alleles is greater (25% vs. 18%). It has been suggested that the variation uncovered by urea denaturation may be the same as that uncovered by thermostability tests (21, 22).

Comparison of the effectiveness of the various methods to uncover electrophoretically cryptic variation is handicapped by the fact that the available data are for different sets of loci in different organisms. However, the *Adh* locus of *D. melanogaster* has been assayed by sequential electrophoresis and heat denaturation and by peptide mapping. The results are compared in Table 5. Using sequential electrophoresis, Kreitman (17) failed to detect any electrophoretically cryptic variants in a sample of 96 allelic products (although he could differentiate some variants first identified by thermostability). Sampsell (19) examined by heat denaturation 4436 allelic products and obtained a small increase in variation. Although the allelic sample studied by fingerprinting is much too small, this technique appears to have considerably greater power for detecting electrophoretically cryptic variation than the two other techniques. Yet, the various techniques give consistent results in the sense that they all indicate that conventional electrophoresis may actually detect most of the protein variation present in natural populations (see also ref. 23). Although, as pointed out, an increase of 20% in the effective number of alleles is not trivial, it is not likely to have drastic consequences for most evolutionary considerations.

These results apply, of course, only to genetic variation that yields differences in the primary structure of proteins. In addition, variation exists in redundant third-codon positions and in the untranslated regions of the DNA. Comparisons between species and between homologous genes within the same genome are providing information about the rates of evolution of redundant vs. nonredundant nucleotide substitutions, of introns vs. exons, of genes vs. intergenic sequences, and so on. In order to estimate variation in natural populations, however, comparisons must be made between individuals (or between the two genomes of the same individual). Slightom *et al.* (24) have sequenced the $^{A}\gamma$ gene in the two chromosomes of one human individual. The two alleles differ in about 2% of the 1468 nucleotide pairs sequenced, but all of the differences occur in untranslated segments of the DNA. Evidence obtained from DNA hybridization suggests that heterozygosity at 2–4% of the nucleotide pairs may be typical of the whole genome (25, 26). These observations suggest that, at the DNA level, if the untranslated regions of genes are taken into account, every individual in sexually reproducing organisms is likely to be heterozygous at every locus.

The charge-state model of genetic variation postulates that most alleles are "cryptic," i.e., undetectable by electrophoresis (2, 5–8). The total number of different amino acid sequences that one would expect to find within a given electromorph depends, first of all, on the number of sites that are free to vary without undermining the functional properties of the protein, but in any case it should be large (27). If we assume, for example, that the number of sites free to vary in a protein is about 10% of the total (28), there would be some 25 such sites in *D. melanogaster* ADH. Because only about half of the molecule is present in the peptides examined by fingerprinting, it might be assumed that about 12 variable sites would be represented in the peptide maps. If each site can be occupied by any one of the 20 amino acids, the number of possible different sequences would be 20^{12}, two-thirds of which would not yield charge differences. Even though only 11 ADHs have been sampled by fingerprinting, the charge-state model predicts that each electrophoretic class should have been heterogeneous, consisting of a variety of sequences. As King (27) has written: "A prediction can . . . be made: when the time and techniques are available, it will be found that the commoner electrophoretic 'alleles' in almost every polymorphic system in large populations will be shown to consist of a number of sequentially distinct isoalleles." The model also predicts that the charge difference between the F and S electromorphs will involve various amino acid substitutions at a variety of sites.

These predictions are not borne out by the present results. First, only one cryptic allele has been found in a sample of 11 independent ADHs. Second, the charge difference between the ADH^S and ADH^F electromorphs is in every case due to the

Table 4. Increase in genetic variation observed by urea denaturation in two Greek populations of *D. subobscura**

	Electrophoresis		Electrophoresis plus urea		Increase in variation	
Locus	Het	n_e	Het′	n'_e	Het′ − Het	n'_e/n_e
Xdh	0.61	2.56	0.77	4.44	0.16	1.73
Est-5	0.60	2.52	0.73	3.74	0.13	1.48
Ao	0.59	2.41	0.69	3.19	0.10	1.32
Pept-1	0.48	1.95	0.48	1.95	0.00	1.00
Lap	0.47	1.90	0.62	2.67	0.15	1.41
Odh	0.12	1.14	0.12	1.14	0.00	1.00
Me	0.10	1.11	0.11	1.13	0.01	1.02
Acph	0.06	1.06	0.13	1.15	0.07	1.08
Average	0.379	1.83	0.456	2.42	0.077	1.25

* From ref. 21.

Table 5. Increase in genetic variation detected by three different methods at the *Adh* locus of *D. melanogaster*

Method	Het′ − Het	n'_e/n_e	Ref.
Sequential electrophoresis	0.00	1.00	(17)
Heat denaturation	0.02	1.03	(19)
Peptide mapping	0.10	1.20	

substitution of lysine for threonine. The significance of this observation is enhanced by the fact that the ADH^F electromorphs examined come from three different natural populations. And there can be little doubt that peptide mapping can detect all amino acid substitutions that alter the charge of a soluble peptide. Thus, the results of the ADH fingerprinting test—as well as those of other tests (8, 23)—are inconsistent with the charge-state model of genetic variation.

I am indebted to John H. Gillespie for reviewing an earlier version of this manuscript and for advice. This work was supported by Grant GM 22221 from the National Institutes of Health.

1. Ayala, F. J. & Kiger, J. A. (1980) *Modern Genetics* (Benjamin/Cummings, Menlo Park, CA), p. 662.
2. Marshall, D. R. & Brown, A. H. D. (1975) *J. Mol. Evol.* **6**, 149–163.
3. Kimura, M. & Ohta, T. (1971) *Nature (London)* **229**, 467–469.
4. Maynard Smith, J. (1972) *Nature (London) New Biol.* **237**, 31.
5. King, J. L. (1973) *J. Mol. Evol.* **2**, 317–322.
6. Ohta, T. & Kimura, M. (1973) *Genet. Res.* **22**, 201–204.
7. Wehrhahn, C. F. (1975) *Genetics* **80**, 375–394.
8. Brown, A. H. D., Marshall, D. R. & Weir, B. S. (1981) in *Genetics Studies of Drosophila Populations*, eds. Gibson, J. B. & Oakeshott, J. G. (Australian National Univ. Press, Canberra), pp. 15–43.
9. Fletcher, T. S., Ayala, F. J., Thatcher, D. R. & Chambers, G. K. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 5609–5612.
10. Fletcher, T. S. (1978) Dissertation (Univ. of California, Davis, CA).
11. Thatcher, D. R. (1980) *Biochem. J.* **187**, 875–886.
12. Schwartz, M. F. & Jörnvall, H. (1976) *Eur. J. Biochem.* **68**, 159–168.
13. Coyne, J. A., Felton, A. A. & Lewontin, R. C. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 5090–5093.
14. Singh, R. S. (1979) *Genetics* **93**, 997–1018.
15. Coyne, J. A. & Felton, A. A. (1977) *Genetics* **87**, 285–304.
16. Beckenbach, A. T. & Prakash, S. (1977) *Genetics* **87**, 743–761.
17. Kreitman, M. (1980) *Genetics* **95**, 467–475.
18. Cochrane, B. J. & Richmond, R. C. (1979) *Genetics* **93**, 461–478.
19. Sampsell, B. (1977) *Biochem. Genet.* **15**, 971–988.
20. Trippa, G., Loverre, A. & Catamo, A. (1976) *Nature (London)* **260**, 42–43.
21. Loukas, M., Vergini, Y. & Krimbas, C. B. (1981) *Genetics* **97**, 429–441.
22. Satoh, C. & Mohrenweiser, H. W. (1979) *Ann. Hum. Genet.* **42**, 283–292.
23. Ramshaw, J. A. M., Coyne, J. A. & Lewontin, R. C. (1979) *Genetics* **93**, 1019–1037.
24. Slightom, J. L., Blechi, A. E. & Smithies, O. (1980) *Cell* **21**, 627–638.
25. Britten, R. J., Cetta, A. & Davidson, E. H. (1978) *Cell* **15**, 1175–1186.
26. Grula, J. W., Hall, T. J., Hunt, J. A., Guigni, T. D., Davidson, E. H. & Britten, R. J. (1982) *Evolution*, in press.
27. King, J. L. (1974) *Genetics* **76**, 607–613.
28. Fitch, W. M. & Margoliash, E. (1970) *Evol. Biol.* **4**, 67–109.