Nearly Identical Allelic Distributions of Xanthine Dehydrogenase in Two Populations of *Drosophila pseudoobscura*[1]

Tim P. Keith, Lisa D. Brooks, Richard C. Lewontin, Juan C. Martinez-Cruzado, and Donna Larson Rigby
Museum of Comparative Zoology, Harvard University

In a previous study, Keith (1983) showed by sequential gel electrophoresis of the esterase-5 protein in *Drosophila pseudoobscura* that a highly polymorphic locus with many alleles can have very similar frequency distributions in populations separated by 500 km. The present work studies another highly polymorphic locus, xanthine dehydrogenase, in the same California population samples, using the same technique to distinguish allelic classes. Twelve electromorphs were found in one population and 15 in the other. Both populations shared a single very frequent (~60%) allele, as well as five other alleles in low but similar frequencies. In addition, each population had an array of unique alleles present only once in one population sample but absent in the other. A statistical test against the stationary distribution for neutral alleles shows that, if the populations are at equilibrium, then purifying selection is operating on xanthine dehydrogenase. The extremely close similarity in frequency distributions of the alleles between populations for both the xanthine dehydrogenase and esterase-5 loci, despite differences in allele frequency distribution between loci, strongly emphasizes the importance of migration in influencing genic diversity in these populations.

Introduction

A major problem in understanding the allelic polymorphism of so many structural gene loci in natural populations has been the genetic ambiguity of the phenotypes seen in electrophoretic surveys. While different electromorphs can be shown unambiguously to be coded by different alleles, the inverse is not true. Proteins indistinguishable by electrophoresis may nevertheless differ by one or more amino acid substitutions, so that the genetic similarity between individuals and populations may be greatly overestimated. In order to resolve this ambiguity, methods have been developed to increase the discriminatory power of standard electrophoretic surveys, notably the technique of sequential gel electrophoresis (Coyne 1976; Singh et al. 1976). This method uses different pH's, buffers, and gel concentrations to break up electromorphic classes defined by a single condition of electrophoresis, doing so either by detecting small differences in the *pK* of ionizable side groups in different ionic environments or by detecting size and conformational differences. When the process has been applied to structural genes coding for

1. Key words: xanthine dehydrogenase, *Drosophila pseudoobscura,* similar electrophoretic distributions, sequential electrophoresis, polymorphism.

Address for correspondence and reprints: Dr. Tim P. Keith, Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138.

Mol. Biol. Evol. 2(3):206–216. 1985.

0737-4038/85/0203-4017$02.00

enzymes in *Drosophila,* it has been found that loci previously judged to be monomorphic or nearly so remain monomorphic (Coyne and Felton 1977; Kreitman 1980) but that loci previously judged to be highly polymorphic reveal an immense increase in genetic diversity under the more discriminating method. For example, the number of alleles at xanthine dehydrogenase (*Xdh*) in *D. pseudoobscura* increased from eight to 27 in 146 genomes examined (Singh et al. 1976), and the number of alleles at esterase-5 (*Est-5*) increased from 11 to 41 in the same species (Keith 1983). Moreover, a disjunct population of *D. pseudoobscura* from Bogotá, Colombia, previously thought to be monomorphic for the *Xdh* allele that is most common in the rest of the species, turned out to be highly polymorphic for geographically unique alleles (Singh et al. 1976).

It is precisely the highly polymorphic loci, with large numbers of alleles, that can be most revealing about the forces operating on genetic diversity when different populations are compared. Monomorphic loci, identical between populations, carry no statistical information at all about the relative roles of selection and drift. Loci with simple two- or three-allele polymorphisms are easily compatible with both a balancing selection and a drift theory, although evidence of consistent clines in allele frequencies such as those observed altitudinally (Grossman et al. 1969) and latitudinally (Oakeshott et al. 1982) in alcohol dehydrogenase of *D. melanogaster* argue strongly for selection. Loci with very large numbers of alleles, on the other hand, have a richness of statistical information that makes a number of tests possible both within and between populations. The tests of Ewens (1972) and Watterson (1977) to distinguish selection from neutrality are much more powerful when there are many alleles. Moreover, migration as an explanation of the similarity of certain allele frequencies could be eliminated if two populations had nearly identical frequencies of very frequent alleles but different arrays of less frequent variants. So, it is of considerable interest to use the technique of sequential gel electrophoresis to make a large-scale comparison of allelic frequency distributions of highly polymorphic loci. Such large-sample comparisons were not made in the original studies by Coyne (1976), Singh et al. (1976), and Singh (1979) since these studies were chiefly oriented toward demonstrating the high level of polymorphism rather than toward characterizing any particular population with great accuracy.

In the first application of sequential gel electrophoresis to large samples of genomes within populations, Keith (1983) studied *Est-5* polymorphism in about 120 genomes from each of two populations of *D. pseudoobscura* separated by 500 km. Thirty-three alleles were present in one population, and 22 were present in the other, for a total of 41 alleles overall, with a striking similarity in frequency distribution between the populations. The two major polymorphic alleles were present in 34% and 21% of one population and in 36% and 21% of the other population, and there was a significant correlation between populations for the rarer alleles. In both populations, the Watterson test showed evidence of purifying selection. It is the purpose of the present paper to compare the *Est-5* results with the situation at a second highly polymorphic locus, xanthine dehydrogenase, in the same two populations. Were this second locus to show great differentiation between populations, for example, then population structure and migration could not be invoked to explain both loci, and selection would have to be operating differently at the two loci. If, on the contrary, *Xdh* turned out to resemble *Est-5* closely in its statistical properties, the simple classical theory of purifying selection with migration would be greatly strengthened.

Table 1
Sequential Electrophoretic Conditions

Condition	Acrylamide Concentration (%)	Buffer	pH	Field Strength (V/cm gel)	Running Time (h)
1	5	TBE	8.9	18	3.5
2	5	TBE	7.1	10	7.5
3	8	TBE	7.1	10	14.0
4	5	Glycine-NaOH	10.4	12	6.0
5	7	TBE	8.9	18	6.0

NOTE.—Buffers were made according to the method of Keith (1983).

Material and Methods

The genomes for this study were obtained from the same isofemale lines collected by Keith (1983) in May 1979 at the James Reserve in the San Jacinto Mountains in southern California (elevation 1,646 m) and in a woodland near the Gundlach-Bundschu winery in the Sonoma Valley of northern California (elevation 30 m). Second chromosomes were extracted from 95 lines from the James Reserve and from 89 lines from Gundlach-Bundschu using a Delta/Bare^{Inv} balanced lethal stock and a single F_1 male from each isofemale line.

The isochromosomal lines were then maintained at 17 C for the repeated retests required by the experiment. It must be emphasized that, as discussed by Singh et al. (1976) and Keith (1983), sequential gel electrophoresis depends on comparing the mobility of lines side by side on the same gels, thus eliminating gel-to-gel differences. *Xdh* standards from the study of Singh et al. (1976) were run on all gels. Lines were first classified relative to these standards under condition 1 of electrophoresis (5% gel, Tris-borate EDTA [TBE], pH 8.9). Lines were classified as identical only after being run side by side on the same gel or when run adjacent to another line of the same mobility. After the first classification, each electromorphic class was reexamined under four other electrophoretic conditions to further subdivide the groups of lines. Table 1 shows the five electrophoretic conditions. Lines were compared over and over again in different side-by-side comparisons and, in most cases, were run and scored by one person as well as scored blindly by another to ensure that differences could be established unambiguously. Lines from both populations were included on the same gels so that electromorphs would be comparable across population samples. Staining of the gels followed the procedure of Prakash, Lewontin, and Hubby (1969).

Genetic Tests

To test whether the mobility differences observed for the different electromorphs were due to the structural locus itself (or to *cis*-acting modifier loci), each variant was crossed to two reference alleles of divergent mobilities (Coyne et al. 1979; Keith 1983). In this study the 1.04/1.00/1.00/1.00/1.00 and 0.90/1.00/1.00/1.00/1.00 alleles were used. In *all* cases, the variants segregated in a codominant manner.

Results

The classification of the lines by electromorphic phenotype is shown in table 2. The names of the allelic classes, given numerical codes, follow the method used

Table 2
Number of Isogenic Lines within Each Population Sample Falling in Different Sequential Electrophoretic Classes

Electromorph	No. of Lines		
	James Reserve	Gundlach-Bundschu	Total
0.90/1.00/1.00/1.00/1.00	0	1	1
0.92/0.98/0.98/1.00/1.00	1	0	1
0.92/1.00/1.00/0.95/1.00	0	1	1
0.94/1.00/1.00/1.00/1.00	4	2	6
0.94/1.00/1.00/1.02/1.02	1	0	1
0.94/1.03/1.00/1.02/1.01	0	1	1
0.94/1.03/1.00/1.03/1.03	1	0	1
0.98/0.98/1.00/1.02/1.02	0	1	1
0.98/1.00/1.00/1.00/1.00	10	9	19
0.98/1.00/1.00/1.02/1.00	1	2	3
0.98/1.02/1.01/1.02/1.02	6	8	14
0.98/1.03/1.01/1.03/1.03	1	0	1
1.00/0.98/0.98/0.98/0.98	0	1	1
1.00/1.00/1.00/1.00/1.00	58	52	110
1.00/1.02/1.01/1.01/1.02	9	4	13
1.00/1.02/1.01/1.02/1.03	2	4	6
1.00/1.03/1.01/1.02/1.03	0	1	1
1.00/1.04/1.01/1.02/1.03	0	1	1
1.03/1.00/1.00/1.00/1.00	1	0	1
1.04/1.00/1.00/1.00/1.00	0	1	1
	12 Electromorphs (5 unique); 95 genes; H = 0.600	15 Electromorphs (8 unique); 89 genes; H = 0.634	20 Electromorphs; 184 genes
	$X^2_{19} = 17.08$; $P = 0.6$		

by Keith (1983). The relative mobilities for sequential gel conditions are separated by a slash mark. The order of the mobilities corresponds to the order of electrophoretic conditions shown in table 1. The most common allele was arbitrarily given the standard mobility 1.00/1.00/1.00/1.00/1.00. Thus, allele 0.98/1.02/1.01/1.02/1.02 had mobility 0.98 relative to the standard under condition 1. Following the first condition, lines having the same mobility under condition 1 were tested as a group under each of the subsequent criteria. Thus, allele 0.98/1.02/1.01/1.02/1.02 migrates farther than allele 0.98/1.00/1.00/1.02/1.00 under conditions 2, 3, and 5 but has identical mobility under conditions 1 and 4. The relative mobilities are always redefined *within a gel condition,* so no significance is to be attached to similarities or differences in mobility constants across conditions. Moreover, the relative mobilities should not be taken too seriously as quantitative measurements. The method of side-by-side comparisons establishes unambiguous *orders* of mobilities rather than their absolute movement in centimeters.

As shown in table 2, both populations are highly polymorphic. There are 12 electromorphs in the James Reserve population, five of which are absent from the

other population, while the Gundlach-Bundschu population has 15 electromorphs, eight of which are unique to that population sample, giving a total of 20 electromorphic classes out of 184 genomes in the total sample. The average heterozygosities are 0.60 and 0.63, respectively, for the two populations.

Table 3 shows the distribution of electromorph frequencies both in each population and in the pooled sample. The distributions are strongly J-shaped. Each population has a single very frequent allele—the same allele in both populations—representing about 60% of all genomes in the sample. Each population then has several electromorphs at frequencies of 0.10 or less and a relatively large number of singleton electromorphs, each represented only once in the sample. This frequency distribution is in contrast with the case of esterase-5 (Keith 1983), in which there are two common alleles making up 35% and 21%, respectively, of the sampled genomes in addition to the long list of low-frequency and singleton electromorphs.

The most striking feature of the data is the close similarity between populations in the electromorph distribution. Table 3 shows that the general shapes of the distribution are very similar. The similarity is even more apparent when the detailed electromorph frequencies in table 2 are compared. The same allele is present in very high frequency in both populations. With the single exception of class 0.98/1.00/1.00/1.02/1.00, which is present twice in Gundlach-Bundschu and only once in James Reserve, every class represented more than once in one population sample is represented more than once in the other. On the other hand, morphs present only once in a population are, with the same exception of allele 0.98/1.00/1.00/1.02/1.00, always absent from the other population sample, and alleles present in one population and not the other are always found only once in that population. Thus, there appears to be a group of electromorphs that are present in both populations in low but similar frequencies and a group of unique alleles in each population represented only once in the sample. Singletons aside, there are no polymorphisms present in one population sample and absent in the other.

The similarity of the two electromorph distributions in table 2 was tested by a 2×20 χ^2 test of homogeneity, giving a value of $X^2 = 17.08$ with 19 df. This corresponds to a probability of $P = 0.6$. Although a large number of the expectations are very small, the χ^2 test is in fact valid, as shown by Lewontin and Felsenstein (1965), who demonstrated that small expectations do not bias the χ^2 test in $2 \times n$

Table 3
Frequency Distribution of Electromorph Classes within Each Population and in the Pooled Sample

James Reserve		Gundlach-Bundschu		Pooled	
Class Frequency	No. of Times Represented	Class Frequency	No. of Times Represented	Class Frequency	No. of Times Represented
0.610	1	0.584	1	0.598	1
0.105	1	0.101	1	0.103	1
0.095	1	0.090	1	0.076	1
0.063	1	0.045	2	0.071	1
0.042	1	0.022	2	0.033	2
0.021	1	0.011	8	0.016	1
0.010	6			0.005	13

tables, especially with large numbers of classes. Indeed, if the singletons are representative of a class of one-time-only mutations, in contrast to the repeated polymorphisms, then a χ^2 test excluding the singletons is a better test of the similarity of the populations. Such a test gives a $X^2 = 3.79$ with 6 df, corresponding to a probability of $P = 0.75$, confirming that the singletons make the largest contribution to the overall X^2.

Discussion

The two main features of the data that are relevant to questions of forces operating on the genetic diversity are (1) the shape of the allele frequency distribution and (2) the close similarity of distributions between the populations. Before attempting to relate the experimental results to questions about the forces operating, however, we must consider the relationship between the electromorphic classes and allelic genetic variation at the *Xdh* locus. The results of our genetic analysis of the present material as well as of the genetics carried out by Singh et al. (1976) on this locus leave no doubt that the electromorphic differences observed result from allelic differences. We must still consider the question, however, of whether there may be allelic variation *within* electromorphs. Any comparison of populations or any conclusions drawn from the shape of the frequency distribution are useful only if we can assume allelic identity of phenotypically identical morphs. Indeed, that is the entire purpose of carrying out sequential electrophoretic comparisons. The evidence that sequential gel electrophoresis distinguishes a high percentage of amino acid substitutions comes from two sources. First, the tests made by Ramshaw et al. (1979) and by McLellan (1984) of known amino acid substitutions in human hemoglobin and myoglobin respectively show that between 85% and 93% of amino acid substitutions are detectable by sequential techniques. However, additional proteins with known amino acid substitutions need to be tested to determine whether these results are general. Second, the DNA sequence studies of the alcohol dehydrogenase locus (Kreitman 1983) showed no hidden amino acid variation within electromorphs that had previously been studied by sequential gel electrophoresis (Kreitman 1980). On the other hand, the detection by peptide mapping of at least one amino acid substitution in *Adh* that leads to differential heat sensitivity with no accompanying electrophoretic difference (Chambers et al. 1981) and the detection of several heat-sensitive, nonelectrophoretic variants in *Xdh* (Singh et al. 1976) make it clear that some genetic heterogeneity within electromorphs is possible. The totality of the evidence, however, suggests that in the present study we have detected nearly all of the genetic variation present. For the remainder of this discussion, then, we will refer to electromorphic classes as representing single alleles, although, of course, there may be a small amount of hidden genetic variation still present.

The shape of the allelic frequency distribution makes it unlikely that the total genetic diversity at the *Xdh* locus can be the result of heterotic balancing selection. Lewontin et al. (1978) have shown that the maintenance of 12 or 15 alleles in stable, balanced polymorphism by heterosis implies that fitness values lie in a very small region of parameter space. Even if we assume that the singleton alleles are new mutations and not part of the stable heterotic distribution, the six remaining allelic classes are too many to be kept in stable equilibrium except by very special sets of selection coefficients. Moreover, their frequency distribution, with one class very frequent and the other classes having a frequency of 10% or less, does not

correspond to the stable frequency distributions found to be possible by Lewontin et al. (1978). An alternative would be to abandon the simple heterotic models and invoke an ad hoc frequency-dependent selection model to explain the observations. While there is no logical flaw in such a procedure, there is no independent evidence that an appropriate form of frequency-dependent selection actually operates. While the entire distribution of allelic frequencies cannot reasonably be ascribed to heterotic selection, it is always possible that the allelic classes observed belong to a small number of selective classes, within which the allelic variation is unrelated. Thus allele 1.00/1.00/1.00/1.00/1.00 might be one physiological class, while the remaining five low-frequency polymorphic alleles could belong to a second class, but the two classes could be selectively distinguishable from each other. Such a mixed selective-neutral model cannot be tested on the basis of the static data of allele frequency distribution but requires some kind of measurement of selective forces themselves.

Having rejected a simple multiple-allelic heterotic model, we can test its polar opposite, the totally neutral model. Ewens (1972) and Watterson (1977) have derived sampling theories of neutral alleles that predict a relation between the number of alleles present in a sample and a measure of allelic diversity in the sample provided that the population distribution of allele frequencies is at the stationary state under random genetic drift. The Watterson test, which uses the homozygosity, F, as a measure of diversity, is the more powerful of the two (Watterson 1977; Ewens 1979), and it is the one we use here. Using a computer simulation algorithm of Stewart (appendix to Fuerst et al. [1977]), we chose 1,000 samples having the same number of alleles as each of our population samples from the theoretical stationary, neutral distribution. From these simulated samples, the sampling distribution of F values was calculated and compared with our observed F values both in each of the two populations and in the pooled sample. The results are shown in table 4 together with the two-sided probabilities of observing the samples by chance. Both populations and the pooled sample deviate from the expectation in the direction of too little diversity (F being too high) for the observed number of alleles. This diminished diversity can be taken as evidence of purifying selection in favor of a wild-type allele, but there is no way to distinguish this explanation from the possibility that the alleles are selectively neutral but that the populations are not at the stationary distribution for purely historical reasons.

When the same test was performed on the data for *Est-5* from these populations (Keith 1983), the same result was obtained, as shown in the lower half of table 4. Either purifying selection is working at both loci in both populations or both deviate from the stationary distribution as a consequence of an historical event affecting both loci.

The second important feature of the data is the close similarity in allele frequency distributions between populations. In the Results section we showed that tests of homogeneity either between the entire distributions or only between the repeated polymorphic alleles revealed very close similarity between the populations. Again, the results very closely parallel what Keith (1983) found at the esterase-5 locus. In the latter case, the test of homogeneity for the entire distribution gave $P = 0.65$.

The similarity between the two California populations at two different loci may be a consequence of migration between them as well as of common selective pressures in the two populations, as indicated by the Watterson test. A moderate

Table 4
Results of Watterson Tests for Neutrality

Population	Observed (F)	Simulation Mean (F)	Simulation Variance (F)	Probability (two sided)
		Xdh (present study)		
James Reserve	0.3997	0.2204	0.0057	<0.07
Gundlach-Bundschu	0.3657	0.1677	0.0028	<0.01
Pooled	0.3816	0.1494	0.0021	<0.05
		Est-5 (Keith 1983)		
James Reserve	0.1712	0.0682	0.0002	<0.01
Gundlach-Bundschu	0.1924	0.1180	0.0011	<0.06
Pooled	0.1801	0.0685	0.0003	<0.01

gene flow across 500 km is not unlikely for *Drosophila pseudoobscura* populations in California, as shown by the long-distance migration studies of Jones et al. (1981) and Coyne et al. (1982). For the *Xdh* locus, it is possible to get some idea of the importance of migration by a comparison with the geographically much more extensive data given by Singh et al. (1976). They examined 11 population samples from the continuous North American distribution of *D. pseudoobscura* in small samples varying from four to 17 genomes. Their results cannot be compared allele for allele with the present results because all the electromorphs from their study would need to be run together on the same gels as those reported here. In the absence of such a comparison, it is not clear, except for the most common allele, what the correspondence is between experiments. Nevertheless, the general statistical properties of the allelic distribution can be matched. Table 5 shows the distribution of allele frequencies in the two studies. There is an overall similarity, with one allele

Table 5
Comparison of the *Xdh* Allele Frequency Distribution in the Present Study with That of Singh et al. (1976)

Present Study		Singh et al.	
Allele Frequency	No. of Times Represented	Allele Frequency	No. of Times Represented
0.598	1	0.507	1
0.103	1	0.083	1
0.076	1	0.060	1
0.071	1	0.037	4
0.033	2	0.022	6
0.016	1	0.015	2
0.005	13	0.007	11
Total alleles	20	Total alleles	25
Total genomes	184	Total genomes	134

representing 50%–60% of all genomes, a number of alleles at low frequencies, and a large number of singletons. There are, however, some clear differences. The present study found only 20 alleles in 184 genomes studied, while Singh et al. (1976) found 25 alleles in only 134 genomes. This larger number of alleles for the genomes studied cannot be explained by the wider geographical sample because it is largely accounted for by low-frequency alleles that are present more than once, at frequencies between 0.015 and 0.037, a class of frequencies represented only three times in the present study. That is, three-quarters of all alleles in the current study are either singletons or high-frequency alleles, while only half of the alleles in the Singh et al. (1976) study have such frequencies. If a wide geographical range simply covered a diversity of populations, the small samples should have given more singletons. While it cannot be seen in table 5, all but one of the alleles, represented three (0.022) and five (0.037) times in the pooled sample of Singh et al. (1976), were observed in several widely separated geographical localities. Thus, they are not locally specialized alleles.

It is possible to ask directly whether the James Reserve and Gundlach-Bundschu samples could have been drawn from an allele distribution like the one found by Singh et al. (1976). To do so, a Monte Carlo simulation of the sampling process was performed, drawing 1,000 samples of size 89 and 95 from a multinomial distribution with frequencies found in the Singh et al. pooled data (table 5, right hand column). The frequency distribution of three statistics were then constructed from the 1,000 samples: (1) the number of alleles present per sample, (2) the number of singletons per sample, and (3) the heterozygosity per sample. The observed values of these three statistics from the two California populations were then compared with the simulated sampling distributions. The results are given in table 6. With the exception of the small number of alleles observed in the James Reserve sample and the concomitantly low heterozygosity in that population, all statistics of the two present samples are in agreement with the expectation based on the broad geographical survey of *Xdh*.

Finally, we can compare the results of our *Xdh* study with Keith's (1983) sample of esterase-5. We have already stated that for both loci the populations are remarkably similar to each other and that both loci show evidence of purifying

Table 6
Comparison of Allelic Distributions Observed with Those Generated by Monte Carlo Sampling from the Allelic Distribution Found by Singh et al. (1976)

Statistic and Population	Observed	Mean from Simulation	Probability of Being ≤ Observed Value
No. of alleles:			
James Reserve	12	18.0	0.001
Gundlach-Bundschu	15	17.6	0.12
No. of singletons:			
James Reserve	6	6.4	0.53
Gundlach-Bundschu	8	6.3	0.84
Heterozygosity			
James Reserve	0.600	0.72	0.05
Gundlach-Bundschu	0.634	0.72	0.10

selection or a common historical deviation from the stationary distribution. The two loci, however, have certain definite differences. While *Xdh* has a simple major allele with a long list of less-frequent alleles and singletons, esterase-5 has a major two-allele polymorphism superimposed on the long list of minor alleles. Second, esterase-5 has more alleles detected per genome sampled than does *Xdh*, that is, 41 alleles in 237 genomes (0.173 alleles per genome sampled) compared with 20 alleles in 184 genomes (0.109 alleles per genome sampled). Sampling theory alone predicts the opposite, that the yield of new alleles will be less than proportional to increased sample size (Lewontin and Prout 1956). Third, alleles unique to one population are often (seven cases out of 27 unique alleles) found more than once in that population at the esterase-5 locus, while that was never the case for 13 unique alleles at the *Xdh* locus. For *Xdh*, unique alleles give the definite appearance of rare mutants, while at the *Est-5* locus they appear more often in multiple copies.

Despite differences between the loci, they both give very similar evidence about the forces operating. There may be sufficient migration between the populations to prevent any divergence between them except for unique mutations that arise in each population separately. In addition, purifying selection appears to be operating in these populations—and possibly in other populations—at the *Xdh* locus (Singh et al. 1976). Both loci are unusual in the very large number of rare alleles segregating, in contrast to the 70% of *Drosophila* enzyme-coding structural genes known to be monomorphic and in contrast to most polymorphic genes that have only two or three alleles segregating in a population. It may be that the larger size of *Xdh* (subunit molecular weight 140,000) and esterase-5 (subunit molecular weight 55,000) allow more amino acid substitutions that are quasi-neutral in their physiological effect than smaller proteins like alcohol dehydrogenase (molecular weight 27,000).

Acknowledgments

We would like to thank Doreen Lewis for her work in preparing *Drosophila* medium, cleaning glassware, and keeping things lively and Becky Jones for deciphering handwriting, keeping us from going bankrupt, and preparing the manuscript.

The experimental work was supported by National Institutes of Health Grant no. GM 21179. The computer simulations were supported by Department of Energy Contract no. DE-AC02-76EV02472.

LITERATURE CITED

CHAMBERS, G. K., W. G. LAVER, S. CAMPBELL, and J. B. GIBSON. 1981. Structural analysis of an electrophoretically cryptic alcohol dehydrogenase variant from an Australian population of *Drosophila melanogaster.* Proc. Natl. Acad. Sci. USA **78**:3101–3107.

COYNE, J. A. 1976. Lack of genic similarity between two sibling species of *Drosophila* as revealed by varied techniques. Genetics **84**:593–607.

COYNE, J. A., I. A. BOUSSEY, T. PROUT, S. H. BRYANT, J. S. JONES, and J. A. MOORE. 1982. Long distance migration of *Drosophila.* Am. Nat. **119**:589–595.

COYNE, J. A., W. F. EANES, and R. C. LEWONTIN. 1979. The genetics of electrophoretic variation. Genetics **92**:353–356.

COYNE, J. A., and A. A. FELTON. 1977. Genic heterogeneity at two alcohol dehydrogenase loci in *Drosophila pseudoobscura* and *Drosophila persimilis.* Genetics **87**:285–304.

EWENS, W. 1972. The sampling theory of selectively neutral alleles. Popul. Biol. **3**:87–112.

——. 1979. Mathematical population genetics. Vol. **9**. Biomathematics. Springer, New York.

FUERST, P. A., R. CHAKRABORTY, and M. NEI. 1977. Statistical studies on protein polymorphism in natural populations. I. Distribution of single locus heterozygosity. Genetics **86:** 455–483.

GROSSMAN, A. I., L. G. KORENEVA, and L. E. ULITSKAYA. 1969. (Variation of the alcohol dehydrogenase locus in natural populations of *Drosophila melanogaster.*) [In Russian] Genetika **6:**91–96.

JONES, J. S., S. H. BRYANT, R. C. LEWONTIN, J. A. MOORE, and T. PROUT. 1981. Gene flow and the geographical distribution of a molecular polymorphism in *Drosophila pseudoobscura.* Genetics **98:**157–178.

KEITH, T. P. 1983. Frequency distributions of esterase-5 alleles in two populations of *Drosophila pseudoobscura.* Genetics **105:**135–155.

KREITMAN, M. 1980. Assessment of variability within electromorphs of alcohol dehydrogenase in *Drosophila melanogaster.* Genetics **95:**467–475.

KREITMAN, M. 1983. Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster.* Nature **304:**412–417.

LEWONTIN, R. C., and J. FELSENSTEIN. 1965. The robustness of homogeneity tests in $2 \times n$ tables. Biometrics **21:**19–33.

LEWONTIN, R. C., L. R. GINZBURG, and S. D. TULJAPURKAR. 1978. Heterosis as an explanation for large amounts of genic polymorphism. Genetics **88:**149–170.

LEWONTIN, R. C., and T. PROUT. 1956. Estimation of the number of different classes in a population. Biometrics **12:**211–223.

MCLELLAN, T. 1984. Molecular charge and electrophoretic mobility in cetacean myoglobins of known sequence. Biochem. Genet. **22:**181–200.

OAKESHOTT, J. G., J. B. GIBSON, P. R. ANDERSON, W. R. KNIBB, D. G. ANDERSON, and G. K. CHAMBERS. 1982. Alcohol dehydrogenase and glycerol-3-phosphate dehydrogenase clines in *Drosophila melanogaster* on different continents. Evolution **36:**86–96.

PRAKASH, S., R. C. LEWONTIN, and J. L. HUBBY. 1969. A molecular approach to the study of genic heterozygosity. IV. Patterns of genic variations in central, marginal, and isolated populations of *Drosophila pseudoobscura.* Genetics **61:**841–858.

RAMSHAW, J. A. M., J. A. COYNE, and R. C. LEWONTIN. 1979. The sensitivity of gel electrophoresis as a detector of genetic variation. Genetics **93:**1019–1037.

SINGH, R. S. 1979. Genic heterogeneity within electrophoretic "alleles" and the pattern of variation among loci in *Drosophila pseudoobscura.* Genetics **93:**997–1018.

SINGH, R. S., R. C. LEWONTIN, and A. A. FELTON. 1976. Genetic heterogeneity within electrophoretic "alleles" of xanthine dehydrogenase in *Drosophila pseudoobscura.* Genetics **84:**609–629.

WATTERSON, G. A. 1977. Heterosis or neutrality? Genetics **85:**789–814.

WALTER M. FITCH, reviewing editor

Received September 7, 1984; revision received December 13, 1984.