LACK OF GENIC SIMILARITY BETWEEN TWO SIBLING SPECIES OF DROSOPHILA AS REVEALED BY VARIED TECHNIQUES[1]

JERRY A. COYNE

Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138

Manuscript received June 11, 1976

ABSTRACT

Acrylamide gel electrophoresis was performed on the enzyme xanthine dehydrogenase in sixty isochromosomal lines of *Drosophila persimilis* from three geographic populations. Sequential electrophoretic analysis using varied gel concentrations and buffers revealed twenty-three alleles in this species where only five had been described previously. These new electrophoretic techniques also detected a profound increase in divergence of gene frequencies at this locus between *D. persimilis* and its sibling species *D. pseudoobscura*. The implications of these results for questions of speciation and the maintenance of genetic variability are discussed.

ELECTROPHORETIC studies of the genetic changes accompanying species formation require a contemporary cross section of groups in different stages of speciation. The *willistoni* group of Drosophila (AYALA *et al.* 1974), the Bogotá population of *Drosophila pseudoobscura* (PRAKASH 1972), and certain Hawaiian drosophilids (CARSON *et al.* 1975) all provide good examples of incipient species. When compared with data on sibling species or distinct morphological species, these studies seem to show that a substantial amount of reproductive isolation can arise without much genetic differentiation (see LEWONTIN 1974, and AYALA 1975, for reviews of the problem). Two factors could account for this. First, genes detected by electrophoresis may have little to do with reproductive isolation, and electrophoretic differences between speciating groups may merely reflect random processes in the presence of such isolation. An alternative explanation is that conventional electrophoretic methods have been unable to detect more extensive genetic differences between nascent species. Since interspecific patterns of similarities and differences at isozyme loci have been used to support selective theories of polymorphism (e.g., AYALA and GILPIN 1974) and to make statements about the importance of regulatory genes in evolution (KING and WILSON 1975), it is additionally important to know whether these patterns are merely the result of the low resolving power of electrophoresis.

New techniques for detecting genetic variation have already been brought to bear on this question. BERNSTEIN, THROCKMORTON and HUBBY (1973) used heat denaturation in conjunction with electrophoresis at the xanthine dehydrogenase locus to investigate differences among eleven species in the *virilis* group of Drosophila. This two-method classification markedly reduced the genetic simi-

[1] Dedicated to the memory of PROF. TH. DOBZHANSKY.

larity between several species, and in one case it completely differentiated two species which had previously shared an electrophoretic allele. SINGH, HUBBY and THROCKMORTON (1975) studied species in the same group with electrophoresis and heat denaturation on the octanol dehydrogenase locus. This work also revealed a general decrease in similarity between species, and this was quite profound in some cases. In 43 out of 45 pairwise interspecific comparisons, genetic distance increased when heat denaturation was used in addition to electrophoresis.

The accompanying study of SINGH, LEWONTIN and FELTON (1976) on the xanthine dehydrogenase locus of *D. pseudoobscura* provides one clue that new methods of genetic classification might strongly reduce the allelic overlap of species. These workers used changes in gel concentration and pH to demonstrate a large decrease in similarity between the main body of *D. pseudoobscura* populations and the one from Bogotá, a population which is spatially and, to some extent, reproductively isolated from the others. Under standard conditions of electrophoresis (a 5% acrylamide gel of pH 8.9), all lines of the Bogotá population possessed the most common allele of the species. When several criteria were used, however, it was demonstrated that 9 of the 12 lines from Bogotá actually had alleles unique to that population.

The present work, undertaken with that of SINGH, LEWONTIN and FELTON (1976), compared the sibling species *D. pseudoobscura* and *D. persimilis* at the xanthine dehydrogenase locus using the new electrophoretic methods described in the accompanying paper.

Although these two species are almost morphologically identical, they differ in their ecology and exhibit almost complete reproductive isolation (DOBZHANSKY 1973). PRAKASH (1969) compared the two electrophoretically and found a strong

TABLE 1

Allele frequencies at the xanthine dehydrogenase locus in D. pseudoobscura *and* D. persimilis *at standard conditions of electrophoresis (5% acrylamide gel of pH 8.9)*

Present data			Previous data		
Allele	Pseudoobscura	Persimilis	Allele	Pseudoobscura	Persimilis
.904 (1—)	.007	—	*.90*	.013	—
.913 (2—)	.007	—	*.92*	.037	—
.924 (3—)	.034	—	*.97*	.031	.012
.939 (4—)	.027	—	*.99*	.209	.042
.957	—	.033	*1.00*	.688	.160
.972 (5—)	.130	—	*1.02*	.022	.715
1.000 (6—)	.705	.117	*1.04*	—	.071
1.014 (7—)	.041	.033			
1.030 (8—)	.048	.033			
1.040	—	.750			
1.077	—	.033			

The previous data is based on work by PRAKASH (*D. persimilis*) and PRAKASH, LEWONTIN and CRUMPACKER (*D. pseudoobscura*).

similarity of allele frequencies at most loci. They do differ to some extent at the xanthine dehydrogenase locus. The right half of Table 1 shows allelic frequencies at this locus in the two species from unpublished work of PRAKASH (*D. persimilis*) and PRAKASH, LEWONTIN and CRUMPACKER (*D. pseudoobscura*). Although the predominant allele differs for each species, they do share other alleles. I wished to ascertain whether the large increase in the number of alleles detected with additional techniques in *D. pseudoobscura* was also true in its sibling, but more particularly in how the use of the new methods would affect the distribution of the shared alleles. If such investigations qualitatively change the picture of isozyme frequencies, much of the previous data and conclusions about the genetic nature of speciation will need reexamination.

MATERIALS AND METHODS

Sixty isofemale lines of *D. persimilis* were obtained from SATYA PRAKASH at the University of Rochester. Twenty-four of the lines were from Fish Creek, Miranda, California (40° 14′ N, 123° 49′ W; collected August, 1968), 21 originated at Mather, California (37° 51′ N, 119° 5′ W; collected August, 1968), and 15 were from Sisters, Oregon (44° 18′ N, 121° 33′ W; collected August, 1969). These lines were made isochromosomal for the second chromosome using a Delta-cardinal marker stock provided by ELIOT SPIESS. As the F_1 males from crosses between the marker stock and the isofemale lines were used in further crosses, the procedure guaranteed that the stocks were indeed *D. persimilis* (male hybrids between *D. pseudoobscura* and *D. persimilis* are sterile). Lines were initially maintained at 17° on agar-yeast-sucrose medium and were later switched to cornmeal-yeast-Karo medium. Tests showed that the electrophoretic mobilities of the xanthine dehydrogenase alleles were not affected by the change of food.

The procedure for classifying variation at this locus was essentially that of SINGH, LEWONTIN and FELTON (1976). To facilitate comparison between the two species, the different electrophoretic conditions were used in the same order as that of their paper: initial classification was made on 5% acrylamide gels of pH 8.9, and further classification was made on 7% acrylamide gels of pH 8.9, 5% acrylamide gels of pH 7.1, and 7% acrylamide gels of pH 7.1, in that order. At each step of the classification my lines were compared with standards of *D. pseudoobscura* obtained from RAMA SINGH. All gels were run at 100 milliamps, and the running times were 3.5 hr. for 5% gels of pH 8.9, 6–7 hr. for 7% gels of pH 8.9, 7–9 hr. for 5% gels of pH 7.1, and 9–11 hr. for 7% gels of pH 7.1. Xanthine dehydrogenase was detected by the method of PRAKASH, LEWONTIN and HUBBY (1969) with the omission of KCl from the stain.

After these four electrophoretic conditions had been used, I experimented with buffers of two additional pH's in an attempt to find further genetic variation at this locus. A tris-HCl buffer of pH 10.4 yielded excellent gels but no additional variations in mobility. Gels were also run cathodally with an acetic acid-sodium acetate buffer of pH 5.1, but no xanthine dehydrogenase bands were visible under these conditions. One final technique was used in search of additional variation. COBBS (1976) has shown that the position, presence, or absence of a heterodimeric band in a heterozygous fly may be a sensitive indicator of genetic differences in subunits of dimeric enzymes. I crossed 20 of the 32 lines remaining in the main allelic class of the locus after all electrophoretic criteria had been used with the unique slow allele of the enzyme in *D. pseudoobscura* (*1000* in the terminology of the accompanying paper). These interspecific heterozygotes were run in gels of all four sets of conditions, but I observed no obvious differences in position or intensity of the heterodimer bands.

The criterion for judging the electrophoretic mobility of alleles as identical or different was a conservative one. Alleles were called "different" if they could be distinguished on essentially every gel run at a given set of conditions. Occasionally differences were noted which appeared only on "good" gels (i.e., those with unusually sharp bands). These were not classified as different, but they invariably became different when conditions were changed further. Each line

was run at least twice for each set of electrophoretic conditions, and alleles which appeared identical to those present in *D. pseudoobscura* were run three or more times.

Unlike SINGH, LEWONTIN and FELTON, I used the measured relative electrophoretic mobilities rather than a numerical designation to classify alleles. Their nomenclature of sequential integers is more convenient, but it gives no idea of the magnitudes of detectable electrophoretic differences. In this study, the electrophoretic mobility of an allele was measured as the distance from the bottom of the injection pocket to the middle of the stained band, relative to a standard allele run on the same gel. Although differences are small between alleles, repeated runs showed the relative mobilities to be surprisingly consistent: between-gel relative mobilities of an allele were always accurate to the second decimal place, and the relative mobilities between gels of alleles placed in the same class never overlapping those of adjacent classes. I give the relative mobilities to the third decimal place.

All intraspecific differences detected by the above techniques were assumed to be genetic based on the accompanying work in *D. pseudoobscura*. Moreover, the interspecific mobility differences in 5% gels of pH 8.9 were shown to be genetic by PRAKASH (1969), who observed heterodimeric hybrid bands at this locus in interspecific hybrids. However, interspecific crosses and backcrosses were made in this study to obviate the possibility that differences detected by the new methods reflected differences between the species not due to variation at the locus itself. These crosses are described in the RESULTS section.

As no heat denaturation or other quantitative techniques were used to uncover genetic varability in *D. persimilis*, the comparison with *D. pseudoobscura* is based only on the first four electrophoretic criteria of SINGH, LEWONTIN and FELTON. The additional alleles detected with heat denaturation in the latter species provide evidence that other genetic variation remains to be found in *D. persimilis*.

In this paper "standard conditions" of electrophoresis will be taken to mean those of a 5% acrylamide gel with gel and electrode buffers of 0.1 M Tris-borate-EDTA, pH 8.9.

RESULTS

The left half of Table 1 gives the allelic frequencies of xanthine dehydrogenase in the two sibling species under standard conditions of electrophoresis. Frequencies are given as frequencies of isochromosomal lines with a given allele, and the data for *D. pseudoobscura* is from SINGH, LEWONTIN and FELTON (1976). The most frequent allele in *D. pseudoobscura*, which was the standard allele for relative measurements under standard conditions, was arbitrarily designated *1.000*. In addition, the mobilities of *D. pseudoobscura* alleles in the left half of Table 1 are described in parentheses by the numerical designations of SINGH, LEWONTIN and FELTON. There is a total of 11 alleles detected under these conditions in both species, with 8 of them found in *D. pseudoobscura* and 6 in *D. persimilis*. Both species display the same general pattern of allele frequencies: one numerically dominant main class, which differs for the two species, flanked by several subsidiary classes with faster and slower mobilities.

The right half of Table 1 shows the allelic frequencies for the two species at this locus based on the previous work already cited. These investigators described a total of 7 alleles in the two species, with 6 of these present in *D. pseudoobscura* and 5 in *D. persimilis*. While the qualitative pattern of these results is similar to ours, the two sets of data differ in several respects. Our work shows that the predominant allele of *D. persimilis* under standard conditions (*1.040*/——) is not present at all in the sample of *D. pseudoobscura*, and that there are actually two previously undetected alleles intermediate in mobility to the main alleles of both

species. These two alleles, *1.014/---* and *1.030/---*, are found in both species. Moreover, this analysis also reveals that no allele in *D. pseudoobscura* with a mobility slower than *1.000* is shared by its sibling. Instead, two lines of *D. persimilis* have an allele, *.957/---*, which is not found in *D. pseudoobscura*.

The qualitative differences between previous data and those reported here are undoubtedly attributable to the improved resolution of alleles made possible by the longer running time of the gel (3.5 instead of 1.75 hr.) and to the fact that the isochromosomal lines precluded the presence of heterozygotes between alleles with nearly identical mobilities. These latter would probably have been classified as homozygotes in previous electrophoretic studies. The usual methods of interspecific surveys, in which individual flies are compared on gels with a few standards, would not be expected to yield results as accurate as those obtainable from repeated runs with homozygous lines.

This method of allelic classification has been used for ten years of work on Drosophila enzymes, and is the method on which almost all conclusions about intraspecific and interspecific gene patterns are based. Dramatically different results are obtained, however, from our sequential electrophoretic analysis. Table 2 gives the total allelic composition of all 60 populations of *D. persimilis* revealed by the four sets of electrophoretic conditions. We now find 23 alleles in the species instead of 6. The nomenclature in the table is based on the same principle as that given in the accompanying paper, but uses measured mobilities rather than integers designating ordered mobilities. Each allele is given a four-part classification corresponding to its mobility relative to standards under each of the four electrophoretic conditions. The first number gives the mobility under standard conditions relative to the most common allele of *D. pseudoobscura*, *1.000/---*. All of the lines possessing the same "allele" under standard conditions are then subjected to the second set of conditions, a 7% acrylamide gel of pH 8.9. The mobilities of the new alleles detected in this manner are then classified relative to the most common allele of all those which composed the same class under standard conditions. This process is repeated for 5% gels of pH 7.1 and 7% gels of pH 7.1, giving the third and fourth relative mobilities respectively. If an allele became completely unique (i.e., represented in only one line) before all classification was completed, the remainder of the numbers in its designation were assigned 1.000's. At the end of the classification there was one line of *D. persimilis* which was in the most common allelic class of *D. pseudoobscura* at the end of its electrophoretic classification. This allele was called *1.000/1.000/1.000/1.000*, as is shown in Table 2.

The striking fourfold increase in the number of alleles which results when several electrophoretic criteria are used instead of one corroborates in *D. persimilis* the findings documented in the accompanying paper for *D. pseudoobscura*. Of the 23 alleles present in our 60 lines, 6 are found under standard conditions, 7 more are observed when the gel concentration is raised to 7% at pH 8.9, 8 more are detected in a 5% gel of pH 7.1, and the remaining 2 are found with a 7% gel of pH 7.1. There are proportionately more alleles in our sample than in the sample of *D. pseudoobscura*, for I found nearly as many alleles (23 instead of 27)

TABLE 2

Distribution of alleles in 60 lines of D. persimilis *after all electrophoretic techniques are used*

Allele	Fish Creek	Mather	Sisters	Total
.957/1.000/1.000/1.000	.	.	1	1
.957/1.019/1.000/1 000	1			1
1.000/ .992/1.000/1.000		1		1
1.000/ .992/1.000/1.011		1		1
*1.000/1.000/ .988/1.000**	1			1
*1.000/1.000/1.000/1.000**			1	1
1.000/1.012/1.000/1.000	1		1	2
1.000/1.012/1.014/1.000			1	1
1.014/1.000/1.000/1.000		1	1	2
*1.030/1.000/1.000/1 000**			1	1
1.030/1.000/1.000/1.017	1			1
1.040/ .980/1.000/1.000		1		1
1.040/ .980/1.020/1.000			1	1
1.040/ .990/1.000/1.000		1		1
1.040/1.000/ .967/1.000	1			1
1.040/1.000/ .984/1.000		1	1	2
1.040/1.000/1.000/1.000	15	12	5	32
1.040/1.008/ .990/1.000	1			1
1.040/1.008/1.000/1.000	2	1	1	4
1.040/1.017/1.000/1.000			1	1
1.040/1.017/1.016/1.000		1		1
1.077/1.000/1.000/1 000	1			1
1.077/1.000/1.008/1.000		1		1
Total lines	24	21	15	60
Heterozygosity	.590	.653	.844	.703
Unique alleles	6	6	6	18

* Alleles also present in *D. pseudoobscura.*

Nomenclature for alleles is: relative mobility at 5% gel, pH 8.9/7% gel, pH 8.9/5% gel, pH 7.1/7% gel, pH 7.1 (see text for further explanation).

in a sample only 40% the size of SINGH, LEWONTIN and FELTON's, and from only 3 populations instead of their 12.

The general pattern of the allelic distribution is similar for the two siblings. At each stage and in the final analysis, there is one numerically predominant allele (*1.040/1.000/1.000/1.000* in *D. persimilis*) and many others with rather low frequencies. At the end of this survey, 32 of the original 60 lines of this species are in the main allelic class. The other 27 lines are composed of 22 alleles, 18 of which are unique to one line. Each population has 6 of these unique alleles.

The heterozygosities for each population and for the total species assuming panmixia are given at the bottom of Table 2. *D. persimilis* has a species heterozygosity of .703, slightly lower than that of *D. pseudoobscura.* The Sisters population of *D. persimilis* has the highest heterozygosity (.845), for only one third of its lines have the common allele.

The genetic identities (the *I* value of NEI 1972) between each pair of populations are given in Table 3 for both standard conditions and for all conditions.

TABLE 3

Genetic similarities expressed as I (NEI *1962*) *between populations of* D. persimilis *at the xanthine dehydrogenase locus*

	Fish Creek	Mather	Sisters
Fish Creek	—	.995	.967
Mather	.958	—	.966
Sisters	.858	.861	—

Values above the diagonal are for classification under standard conditions only, and values below the diagonal are for the total classification under all four electrophoretic conditions.

Interpopulational similarities decrease considerably when all criteria are used: *I* values of .97 or greater are quite common for populations of a species, but those lower than .90 are rare (AYALA 1975). Again, the Sisters population exhibits a smaller identity with the other two because of the paucity of its lines with the common allele.

Our new methods of analysis at this locus reveal a substantial increase in allelic divergence between the species as well as in the number of alleles. Table 4 shows how the alleles held in common by the two species change when our sequential criteria are used for the comparison. Under each successive criterion in this table is given the number of lines of each allele in *D. persimilis* which is also found in *D. pseudoobscura* under those conditions. Next to each such allele in *D. persimilis* is given the number of lines of *D. pseudoobscura* sharing this allele at that stage of the analysis, plus the numerical designation of that allele according to the nomenclature of SINGH, LEWONTIN and FELTON. At standard conditions there are 11 lines of *D. persimilis* sharing a total of 3 alleles with

TABLE 4

Lines of Drosophila persimilis *and* D. pseudoobscura *having alleles in common at each stage of the electrophoretic analysis*

5% gel, pH 8.9				7% gel, pH 8.9				5% gel, pH 7.1				7% gel, pH 7.1			
persimilis		*pseudoobscura*		*persimilis*		*pseudoobscura*		*persimilis*		*pseudoobscura*		*persimilis*		*pseudoobscura*	
No.	Allele	No.	Allele	No.	Allele	No.	Allele	No.	Allele	No.	Allele	No.	Allele	No.	Allele
7	1.000/--	103	6---	2	1.000/.992/--	4	60---	2	1.000/.992/1.000/-	1	600-	1	1.000/.992/1.000/1.000[b]	0	Absent
												1	1.000/.992/1.000/1.011[b]	0	Absent
				2	1.000/1.000/--	88	63---	1	1.000/1.000/.988/-	6	630-				
								1	1.000/1.000/1.000/-	82	631-	1	1.000/1.000/.988/1.000	6	6300
				3	1.000/1.012/--[a]	0	Absent								
												1	1.000/1.000/1.000/1.000	68	6311
2	1.014/--	7	7---	2	1.014/1.000/--	1	71---	2	1.014/1.000/1.000[c]	0	Absent				
												1	1.030/1.000/1.000/1.000	3	8000
2	1.030/--	6	8---	2	1.030/1.000/--	6	80---	2	1.030/1.000/1.000	3	800-	1	1.030/1.000/1.000/1.017[d]	0	Absent
Alleles in common:		3/11 = .273			4/22 = .182				4/42 = .095				3/47 = .064		
Genetic similarity (I):		.156			.062				.035				.033		
Probable misdiagnosis:		.017			.006				.003				.001		

a = Faster than 63--; b = Faster than 6000; c = Faster than 7100 ; d = Faster than 8000.

Under each criterion is given the number of lines of each species having an allele in common, together with the designation of that allele in each species.

D. pseudoobscura, and 7 of these lines were in the most common allelic class of the latter species. After all criteria are employed, only 3 of these 11 lines still possess alleles present in *D. pseudoobscura*, the other 8 having alleles unique to *D. persimilis*. Each change of electrophoretic conditions further serves to differentiate the species. Of the 3 lines of *D. persimilis* still sharing alleles with its sibling at the end of the analysis, only one has the predominant allele of *D. pseudoobscura* (*1.000/1.000/1.000/1.000*, or *6311* in the terminology of SINGH, LEWONTIN and FELTON). The other two shared alleles are rare ones in *D. pseudoobscura*, with one (*6300*) having six representatives in the 146 lines of that species and the other (*8000*) only three. Moreover, SINGH, LEWONTIN and FELTON have subdivided the shared allelic classes *6300* and *6311* into still smaller ones with heat denaturation; and the genetic similarity between the species must be even smaller than we have shown here.

It is obvious, then, that these methods have tremendously increased the apparent differentiation between the species and have altered both the qualitative and quantitative description of the allelic distribution at this locus.

At the bottom of Table 4, I give three measures of interspecific similarity for this locus and how they change as the criteria are used successively. The percentage of total alleles shared by both species drops from 27% to 6% when all electrophoretic criteria are used. Also provided in the table is NEI's (1972) index of genetic similarity, *I*. This measures the normalized probability of picking identical alleles from two species (or populations) compared to the probability of picking them from within one species (or population). The *I* value falls from .156 to .033 when all electrophoretic conditions are used for the comparison. Finally, we give the probability of misdiagnosis according to the definition of AYALA and POWELL (1972). This gives the probability of making a taxonomic error when assigning a fly of known genotype to the species in which that genotype is more frequent (panmixia and Hardy-Weinberg equilibria are assumed). This probability of misdiagnosis decreases from .017 under standard conditions to .001 at the end of the analysis. If the 99% probability of correct assignment is used as the definition of a diagnostic locus, the use of our additional methods have made this locus diagnostic when it was not so under standard conditions.

For the genetic analysis I selected two lines of *D. persimilis* which had two alleles "identical" with two in *D. pseudoobscura* under standard conditions, but which became unique to *D. persimilis* under further analysis. One line, with allele *1.000/1.012/1.000/1.000*, was identical to the allele *6*—— of *D. pseudoobscura* under standard conditions, but in a 7% acrylamide gel of pH 8.9 the former allele moved faster than the fastest subclass of the class in *D. pseudoobscura* (i.e., faster than *63*—). The other line, with allele *1.014/1.000/1.000/1.000*, appeared identical to the unique allele *71*— in *D. pseudoobscura* in 5% and 7% gels of pH 8.9, but became faster than it in a 5% gel of pH 7.1. Interspecific crosses were made between each pair of lines, followed by reciprocal backcrosses of the F_1 females to parental males. Figure 1 shows the electrophoretic results of the second cross described above. These gels were run in 5% acrylamide of pH 7.1, the conditions under which the alleles could be distinguished from

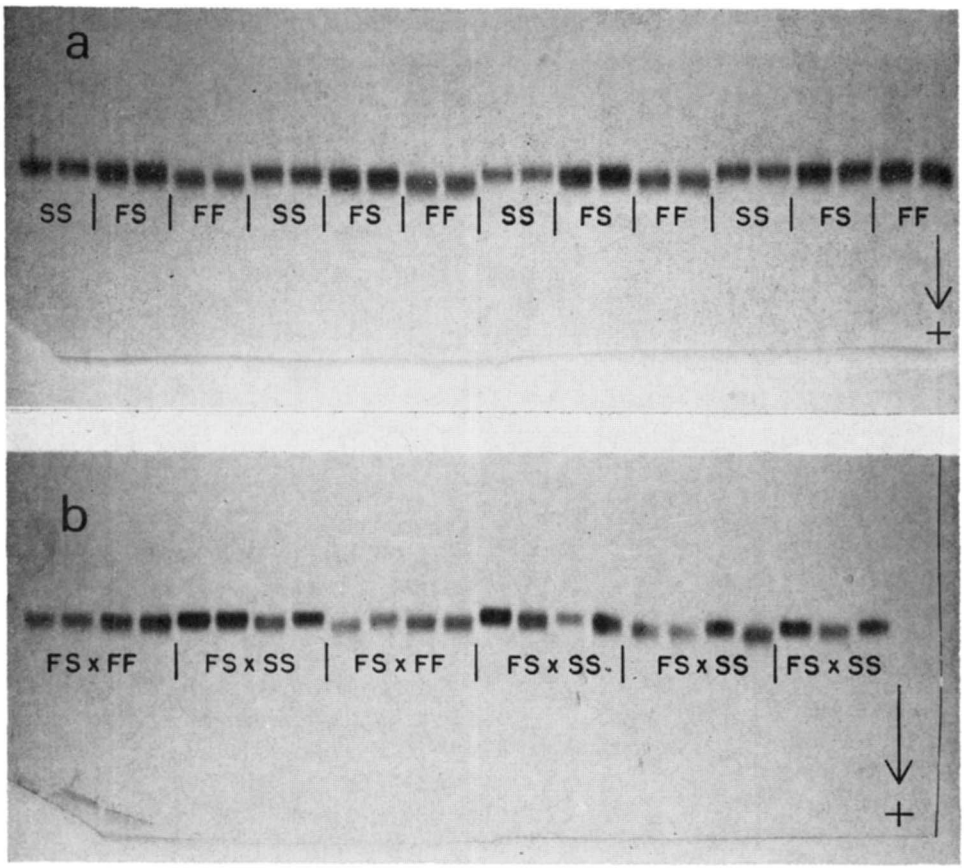

FIGURE 1.—Genetic analysis of interspecific mobility differences in a 5% acrylamide gel of pH 7.1. (a) Parental bands and interspecific hybrids. (b) Segregation in the reciprocal backcrosses. (See text for further explanation).

each other. In Figure 1a the F_1 hybrids are shown flanked by xanthine dehydrogenase bands of each of the parental species (each pair of adjacent pockets in this gel was injected with the homogenate of three individuals from the appropriate line ground in 30 microliters of buffered sucrose). "*FF*" is the allele *1.014/1.000/1.000/1.000* of *D. persimilis*, "*SS*" is allele *7100* of *D. pseudoobscura*, and "*FS*" is the interspecific hybrid between the two lines. The parental bands are clearly different from each other. Although three bands cannot be seen in the hybrids because of the proximity of parental band mobilities, the bands in the F_1 flies are clearly intermediate between the two and are also wider, encompassing the mobility ranges of the parental bands. There is also a hint of increased enzyme activity in the hybrids (all bands are somewhat fuzzy because these gels were run for much longer than the usual time for gels at these conditions). Figure 1b shows bands in individual progeny (1 fly per pocket in 15 microliters of sucrose solution) from the reciprocal backcrosses. These backcrosses are

alternated on the gel as indicated to emphasize mobility differences. Although the bands are again rather fuzzy, the segregation of fast, slow, and intermediate alleles appears as expected in the appropriate backcrosses, and heterozygotes and parental homozygotes can be tentatively identified. The results for the other cross and blackcrosses were similar, except that the F_1 hybrid bands, though intermediate, were not always wider than the parental bands. Taken together, these results indicate that the interspecific differences in mobility are genetic.

DISCUSSION

The most important result of this investigation is the discovery that the interspecific electrophoretic distribution of alleles at the xanthine dehydrogenase locus becomes almost completely disjunct under our new methods of analysis. Eight of the 11 lines of *D. persimilis* with alleles thought homologous to those in *D. pseudoobscura* actually possessed alleles unique to *D. persimilis*. Especially striking is the almost complete differentiation of *D. persimilis* lines which originally appeared to have alleles in main class of *D. pseudoobscura* (*1.000/*——, or *6*—— in the terminology of SINGH, LEWONTIN and FELTON). There were 7 lines with this allele under standard conditions, but after all electrophoretic techniques had been used only one line remained which was still in the most common class of *D. pseudoobscura* (*1.000/1.000/1.000/1.000*). Obviously, the distribution of subclass alleles within this originally common class is truly different for the two species. It may be argued that these results are expected because our species were differentiated under standard conditions at the locus, but the divergence of the Bogotá population of *D. pseudoobscura* from the rest of the species populations, which is documented in the accompanying paper, shows that profound divergence can be obtained even with originally similar allele frequencies. These results indicate that previous interspecific comparisons of gene frequencies would bear repeating with methods similar to ours.

The other noteworthy outcome of these investigations is the discovery of the large number of alleles previously undetected in this species, substantiating the results of the accompanying paper. We have detected 23 alleles at this locus in *D. persimilis*. As this species shares only 3 of these with its sibling, there are at least 47 electrophoretically detectable alleles of xanthine dehydrogenase present in both species. The discovery of at least 10 additional alleles in *D. pseudoobscura* by their differential responses to heat leads one to believe that there are well over 60 alleles in the two species. Thus the total detectable number of alleles has increased by more than a factor of five with the use of the new methods. Before generalizing this multiplicity of new alleles to other loci, however, one should remember that xanthine dehydrogenase is an abnormally large enzyme in Drosophila. According to SEYBOLD (1974), the molecular weight of the monomer exceeds 135,000.

Two previous studies have compared genetic variation detected by electrophoresis with that found with other biochemical techniques. BOYER *et al.* (1972) concluded from an analysis of primate hemoglobins that the total genetic vari-

ation is at least five times that which is found by electrophoresis in starch gels of pH 8.5. KING and WILSON (1975) compared genetic variation of proteins in chimpanzees and men using electrophoresis, immunology, and protein sequencing. They concluded that only about 27% of the total variants in this panoply of proteins would be detected electrophoretically. These figures are comparable to ours, and may mean that our methods uncover most of the genetic variation.

The form of the allelic distributions of the two species is remarkably similar. One main allelic class in *D. persimilis* contains over half the lines, and there are 22 other classes, none of which comprises more than four lines. Eighteen of these lines have a "unique" allele. If one can generalize from the work on *D. pseudoobscura*, in which over half of the rare alleles are represented more than once, we can conclude that most of these "unique" alleles are probably due to the small sample size and are not truly unique mutations. Finally, as in *D. pseudoobscura*, I found interpopulational differences in the alleles detected by the new methods. BERNSTEIN, THROCKMORTON and HUBBY (1973) also noticed that while alleles detected under standard conditions have fairly similar frequencies in all populations of a species, alleles detected by heat denaturation differ from population to population. Whether this is a real phenomenon awaits much more extensive surveys using many lines from each of many populations.

The implications of our results regarding the maintenance of genetic variability within a species has been thoroughly discussed in the accompanying paper and will not be treated here *in extenso*. Obviously, one cannot make firm generalizations about such mechanisms from the analysis of a single, abnormally large protein.

The similarity of allelic distributions between the two species and the form of the distributions themselves are intriguing. In contrast to the plethora of studies examining alleles in one species and comparing such distributions interspecifically, relatively little attention has been paid to the specific form of allelic distributions for isozyme loci. Many species of Drosophila exhibit rather similar forms of the distributions. Under standard conditions of electrophoresis one typically finds one common allele in the middle of the distribution flanked by faster and slower alleles of much lower frequencies (see AYALA *et al.* 1974, for representative data). Our distributions are also of this form, although we cannot say that the predominant allele is in the "middle" of the multidimensional distribution because not all electrophoretic criteria were used independently on all alleles.

Selective hypotheses concerning the maintenance of genetic variability have commonly failed to compare the distributions of alleles predicted by theory with those actually found. Frequently, for example, the presence of multiple alleles is explained by "balancing selection", usually implying single-locus overdominance for fitness. However, no models of this type have been made to account for the common, uneven gene-frequency distributions. Other forms of balancing selection, such as multiple-niche polymorphism, frequency-dependent selection, and fluctuating selection have also been held responsible for polymorphism, but these theories, too, have not predicted particular distributions. In the absence of appropriate models one can only guess that rather specific assumptions might be

needed to explain strongly peaked allelic distributions with a selective theory.

The theory of allelic neutrality can, of course, explain the abundance of alleles if one postulates populations of large effective size. In our case mutation rates and population sizes need not be drastically adjusted to explain the new variation, for although the actual number of alleles in *D. persimilis* has risen by a factor of four, the effective number has only doubled (from 1.7 under standard conditions to 3.4 after all methods are used). Our knowledge of population sizes and isozyme mutation rates is insufficient to preclude the existence of even hundreds of alleles at a locus in Drosophila. Neutral theory can also account for a peaked allelic distribution. KIMURA and OHTA (1975), for example, derive a J-shaped distribution of gene frequencies, with a few alleles common and many rather rare. Other modifications of this theory, such as that of a simple mutational balance between classes (MAYNARD-SMITH 1972) can explain the position of the most common allele in the center of the electrophoretic distribution.

One existing statistical test—that of EWENS (1972)—has been devised to compare actual frequency distributions of alleles with those predicted by neutral theory. EWENS's model is based on certain assumptions about the specific distribution of his statistics, and also assumes that all mutations are to alleles not previously existing. The test itself uses information-theoretic measures, and yields the comparison of data with theory as an F ratio. Significantly large values of F provide evidence for an allelic distribution too even to be accounted for by neutrality ("heterotic selection", according to EWENS), while significantly large values of 1/F indicate distributions too uneven for neutrality ("selection favoring one allele"). Applying the computer program given in Appendix 4 of EWEN's paper to the present data, I found that the Mather and Fish Creek populations as well as the summed data for all three populations had F values below 1, in the direction of unevenness. The Sisters population, on the other hand, had an F ratio in the direction of excess evenness because of its lower frequency of the predominant allele. Because of the small sample sizes, however, the degrees of freedom were too small for the values to achieve significance. Much larger sample sizes than those used here will be necessary to properly test the data for correspondence to neutrality.

A simple model has recently been proposed by KING and OHTA (1975) to account for a symmetrical, peaked allelic distribution under standard conditions of electrophoresis. They assume one selectively favored "type" allele with recurrent stepwise mutation, each step away from the type incurring additional selective disadvantage. Each selective value is on the order of the mutation rate. This mutation-selection balance is then combined with the probability that mutations will be electrophoretically detectable to yield a final model which predicts a symmetrical distribution of electrophoretic alleles about one common allele. By adjusting the magnitude of the selection, the common allele can be made as predominant as desired, and the model does take into account our observed heterogeneity of allelic classes. King and Ohta's model explains the features of our allelic distribution if one assumes that allele *1.040/1.000/1.000/1.000* is the type allele and all other alleles are slightly and almost equally deleterious in

comparison to it. Whether the distribution of alleles predicted by this model is statistically in consonance with our results awaits a suitable test. Too, the theory is an equilibrium one, applicable only to very large populations. It is not certain if populations of temperate zone Drosophila flies meet these assumptions.

It is obvious, then, that mere impressionistic inspection of electrophoretic data in one species will not settle the controversy over the maintenance of genetic variability.

Interspecific comparisons of gene frequencies have also been used to buttress one side or another of the controversy (e.g. AYALA *et al.* 1974). In the case of our data, both neutralist and selectionist arguments can explain interspecifically disjunct allelic distributions with a different mode for each species. For balancing selection, one need only suppose that the selective value of a given allele is different in different species. This would also explain why alleles appearing identical between species under standard conditions become so different when examined in other ways. A pure mutation-drift hypothesis could explain the pattern invoking founder effects, population bottlenecking, or long periods of separation between species. The model of KING and OHTA would account for the pattern by assuming a new "type" allele for a new species. Such an allele could arise by either ecological changes during speciation or sampling effects in a small, speciating population and subsequent modifying selection. The establishment of a new type allele could then result in changes in mutational and selective constraints, and this might lead one to expect that alleles appearing interspecifically identical under standard conditions would actually be quite different (i.e., subclasses of electrophoretic classes would have different frequencies in both species). Here again, simple observation of interspecific gene patterns is unlikely to lead to a resolution of the controversy.

While our methods may not lead to an immediate preference for one mechanism maintaining genetic diversity, they do have taxonomic value. The use of multiple conditions of electrophoresis has made the xanthine dehydrogenase locus diagnostic for these two species when it was not so under standard conditions. However, since these techniques require the use of multiple gels run sequentially for identification of an allele, they might prove unwieldly for population samples in which individual flies require classification.

Our work, then, is a second step in the description of the genetic changes which accompany speciation. Research until now has indicated that reproductive isolation is not concomitant with marked changes in gene frequencies and hence has fostered the belief that reproductive isolation may be due to changes at only a few loci. The present results suggest that this conclusion is premature. LEWONTIN (1974) has noted in interspecific studies the absence of "species-distinguishing genes"—those loci which are fixed for alternate alleles in closely related species. If isozymes contribute to reproductive isolation, one might expect to find this type of locus frequently. Our new methods may be applied to loci which appear identically monomorphic in two species, for such an identity may mask profound differences.

If our findings prove to be general, new methods of isozyme surveys will be

necessary to insure accurate identification of alleles. These should include longer running times, more liberal use of standards on gels, and, where possible, the use of isochromosomal or highly inbred lines.

Our methods can be expanded in several directions to uncover additional genetic variation in populations. A range of pH's in gel and buffer can be used, along with several gel concentrations both higher and lower than those normally used. Modifying agents such as succinate may react differently with different isozymes and perhaps alter their electrophoretic mobilities (D. WALLACE, personal communication). Heat denaturation may also be included in the repertoire of techniques, though I have not found data generated by this method to be consistently repeatable.

Finally, one should bear in mind that these methods may not be applicable to other loci or other organisms. Several other second-chromosome loci have proved refractory to changes in electrophoretic conditions. The larval proteins Pt-7 and Pt-8, for example, form bands too diffuse to score at gel concentrations other than 5%, and they run so close together when the pH is altered that they cannot be scored. Aldehyde oxidase exhibits band splitting at acrylamide concentrations higher than 5%. These methods are being applied successfully, however, to two other second-chromosome loci, octanol dehydrogenase and alcohol dehydrogenase-6. In the latter case I have already found several additional alleles and these two species are becoming further differentiated.

I am most grateful to DR. RAMA SINGH for his help in these experiments. Without his advice, instruction, and his isochromosomal lines of *D. pseudoobscura*, this work would not have been possible. RICHARD LEWONTIN and DONALD WALLACE provided welcome suggestions in the course of the research. Thanks are also due DR. SATYA PRAKASH for the strains of *D. persimilis* and for permission to use his unpublished data. LYNDAL ANDREWS, A. H. COLEMAN, A. A. FELTON, KATHY HORTON, LASZLO MESZOLY and DR. ELIOT SPIESS provided additional assistance. I am indebted to RUSS LANDE, RICHARD LEWONTIN, PETER HAAS and DON HICKEY for critically reading the manuscript. This work was supported by a National Science Foundation predoctoral fellowship to the author and by NIH Grant GM-21179 to R. C. LEWONTIN.

LITERATURE CITED

AYALA, F. J., 1975 Genetic differentiation during the speciation process. Evolutionary Biology **8**: 1–78.

AYALA, F. J. and M. E. GILPIN, 1974 Gene frequency comparisons between taxa: support for the natural selection of protein polymorphisms. Proc. Natl. Acad. Sci. **71**: 4847–4849.

AYALA, F. J. and J. R. POWELL, 1972 Allozymes as diagnostic characters of sibling species of Drosophila. Proc. Natl. Acad. Sci. **69**: 1094–1096.

AYALA, F. J., M. L. TRACEY, L. G. BARR, J. F. MCDONALD and S. PEREZ-SALAS, 1974 Genetic variation in natural populations of five Drosophila species and the hypothesis of the selective neutrality of protein polymorphisms. Genetics **77**: 343–384.

BERNSTEIN, S. C., L. H. THROCKMORTON and J. L. HUBBY, 1973 Still more genetic variability in natural populations. Proc. Natl. Acad. Sci. **70**: 3928–3931.

BOYER, S. H., A. N. NOYES, C. F. TIMMONS and R. A. YOUNG, 1972 Primate hemoglobins: polymorphisms and evolutionary patterns. J. Human Evolution **1**: 515–543.

CARSON, H. L., W. E. JOHNSON, P. S. NAIR and F. M. SENE, 1975 Allozymic and chromosomal similarity in two Drosophila species. Proc. Natl. Acad. Sci. **72**: 4521–4525.

COBBS, G., 1976 Polymorphism for dimerizing ability at the esterase-5 locus in *Drosophila pseudoobscura*. Genetics **82**: 53–62.

DOBZHANSKY, TH., 1973 Is there gene exchange between *Drosophila pseudoobscura* and *Drosophila persimilis* in their natural habitats? Amer. Nat. **107**: 312–314.

DOBZHANSKY, TH. and C. EPLING, 1944 Contributions to the genetics, taxonomy, and ecology of *Drosophila pseudoobscura* and its relatives. Carnegie Inst. Wash. Pub. **554**: 1–183.

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

KIMURA, M. and T. OHTA, 1975 Distribution of allelic frequencies in a finite population under stepwise production of neutral alleles. Proc. Natl. Acad. Sci. **72**: 2761–2764.

KING, J. L. and T. OHTA, 1975 Polyallelic mutational equilibria. Genetics **79**: 681–691.

KING, M.-C. and A. C. WILSON, 1975 Evolution at two levels in humans and chimpanzees. Science **188**: 107–116.

LEWONTIN, R. C., 1974 *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York. 346 p.

MAYNARD-SMITH, J., 1972 Protein polymorphism. Nature New Biol. **237**: 31.

NEI, M., 1972 Genetic distance between populations. Amer. Nat. **106**: 283–292.

OHTA, T. and M. KIMURA, 1975 Theoretical analysis of electrophoretically detectable polymorphisms: models of very slightly deleterious mutations. Amer. Nat. **109**: 137–145.

PRAKASH, S., 1969 Genic variation in a natural population of *Drosophila persimilis*. Proc. Natl. Acad. Sci. **62**: 778–784. ——, 1972 Origin of reproductive isolation in the absence of apparent genic differentiation in a geographic isolate of *Drosophila pseudoobscura*. Genetics **72**: 143–155.

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

RICHMOND, R. C., 1972 Enzyme variability in the *Drosophila willistoni* group. III. Amount of variability in the superspecies *D. paulistorum*. Genetics **71**: 87–112.

SEYBOLD, W. D., 1974 Purification and partial characterization of xanthine dehydrogenase from *Drosophila melanogaster*. Biochim. Biophys. Acta **334**: 266–271.

SINGH, R. S., J. L. HUBBY and L. H. THROCKMORTON, 1975 The study of genic variation by electrophoretic and heat denaturation techniques at the octanol dehydrogenase locus in members of the *Drosophila virilis* group. Genetics **80**: 637–650.

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.

Corresponding editor: W. J. EWENS