# A COMPARATIVE STUDY OF THE ESTERASE-5 LOCUS IN *DROSOPHILA PSEUDOOBSCURA*, *D. PERSIMILIS* AND *D. MIRANDA*

GARY COBBS AND SATYA PRAKASH

*Department of Biology, University of Rochester, Rochester, New York 14627*

Manuscript received May 20, 1976
Revised copy received November 29, 1976

ABSTRACT

Electrophoretic phenotypes of the esterase-5 locus were examined in the sibling species *D. pseudoobscura*, *D. persimilis* and *D. miranda*. *D. persimilis* alleles were found to have uniformly higher charge on monomers than corresponding alleles of either *D. pseudoobscura* or *D. miranda*. Consequently, *D. persimilis* shares no alleles in common with either *D. pseudoobscura* or *D. miranda*, while the latter two species share a number of alleles. It was discovered that by increasing the concentration of acrylamide gel and increasing the length of migration, more allelic differences could be distinguished. Also more alleles were discovered by examining monomer mobility in addition to dimer mobility. In *D. persimilis* and *D. miranda* it was found that the previously known high frequency allelic classes broke down into several allelic classes. A test of goodness-of-fit to the infinite alleles model was done and a rough agreement with the model was found.

AN important quantity in molecular population genetics is the number of allelic forms of genetic loci coding for proteins in natural populations. The most common method used to detect genetic polymorphism of enzyme loci is gel electrophoresis. It has been proposed that routine methods of gel electrophoresis allow the investigator to detect roughly 27% of new mutations involving a change of a single amino acid (LEWONTIN 1974). Consequently, it is suspected that there are more alleles present in samples than have been reported in past studies.

In addition to the question of how many alleles are present in populations, there is much interest in the number and frequency of alleles common to related species. It is possible that some electromorphic alleles which are common to related species are not the same but rather have differences that are not revealed by routine methods of electrophoresis. This latter possibility is of particular interest, as the observation of the same electromorphic alleles in related species has been used in arguments attempting to distinguish between alternative models explaining the existence of enzyme polymorphisms (LEWONTIN 1974).

The esterase-5 locus in the sibling species *Drosophila pseudoobscura*, *D. persimilis* and *D. miranda* is of interest, as it has been found to be highly polymorphic

in all three species.There are 15 alleles known in *D. pseudoobscura* (PRAKASH, LEWONTIN and HUBBY 1969; COBBS 1976), 6 alleles in *D. persimilis* (PRAKASH 1969) and 5 alleles in *D. miranda* (PRAKASH, 1977). The *Est-5*$^{1.07}$ and *Est-5*$^{1.12}$ mobilities exist in all three species (PRAKASH, 1977 and present data).

This paper will report that with relatively minor changes in the methods of electrophoresis we have been able to find many more alleles in both *D. persimilis* and *D. miranda*. We have found that *D. persimilis* has no allele in common with either *D. pseudoobscura* or *D. miranda*. *D. pseudoobscura* and *D. miranda* have some alleles which, so far, appear to be common to both species.

## MATERIALS AND METHODS

*Collections*: Forty-three isofemale lines of *D. persimilis* collected from Mather, California in 1968 were sib mated until electrophoretically homozygous lines were obtained. These lines had been kept in vials for a number of years before sib matings and very few of them were segregating for electrophoretic alleles. Thirty-three isofemale lines of *D. miranda* collected from Sisters, Oregon in 1969 were also treated in the above manner. Again very few of these lines were segregating for electrophoretic alleles. Thirty *D. pseudoobscura* lines were derived as above from a sample taken at Mesa Verde, Colorado in 1971.

*Electrophoresis*: Vertical acrylamide gel electrophoresis was performed according to the methods of HUBBY and LEWONTIN (1966). The steps of centrifuging and immersion of the gel in boric acid solution for the staining of esterase-5 were omitted. The concentration of acrylamide varied from 4% to 12% and the gels were run for longer periods of time. The time of electro-

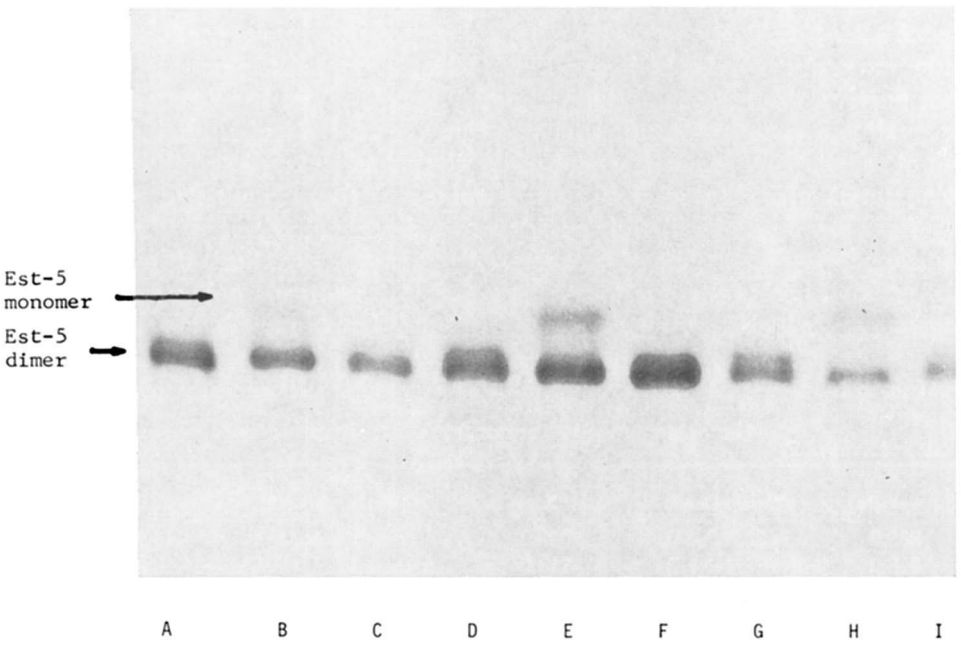


FIGURE 1.—*Est-5*$^{1.07}$ electrophoretic phenotypes in 5% acrylamide gel. Pockets A, D, and G are *D. pseudoobscura*, pockets B, E and H are *D. persimilis* and pockets C, F and I are *D. miranda*. Origin is toward bottom of picture.

phoresis varied with the percent gel. Higher concentration gels were run much longer than lower concentration gels. Gel concentrations of 5% and 8% were run for 2½ and 4 hours, respectively, at 16.1 volts/cm and 2 and 3 hours respectively for 11.6 volts/cm.

## RESULTS

### *Monomer mobility differences*

In 5% acrylamide gels run at 11.6 v/cm we found, as previously reported, *Est-5*$^{1.07}$ alleles in all three species. The *Est-5*$^{1.07}$ electrophoretic mobility of *D. pseudoobscura* and *D. miranda* appeared identical. The *D. persimilis Est-5*$^{1.07}$ phenotype differed from the above phenotypes by the presence of additional fast band. Examples of these phenotypes are shown in Figure 1. If the same strains are examined in 8% gels at 16.1 v/cm, then all three species give two banded phenotypes. Often the two bands have a continuum of light stain between the two bands. In 8% gels all three species retain the original *Est-5*$^{1.07}$ bands, which are still identical. The faster band of *D. pseudoobscura* and *D. miranda* appear identical and are slower than the faster band of *D. persimilis*. Examples of these phenotypes are given in Figure 2. The faster bands of *D. pseudoobscura* and *D. miranda* are not seen in 5% gels because both bands have the same mobility. The faster bands seen in 8% gels must have a lower molecular weight than the normal *Est-5*$^{1.07}$ enzyme because their relative mobilities increase with increasing concentration of gel. This follows from the well known relationship between log $R_m$

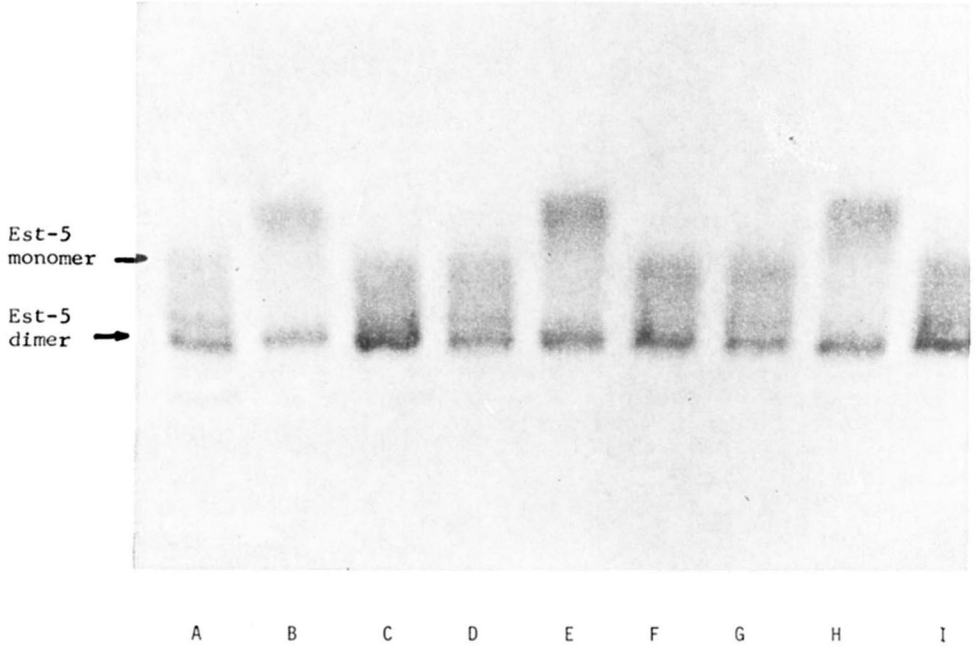


FIGURE 2.—*Est-5*$^{1.07}$ electrophoretic phenotypes in 8% acrylamide gel. Pockets A, D, and G are *D. pseudoobscura*, pockets B, E and H are *D. persimilis* and pockets C, F and I are *D miranda*. Origin is toward bottom of picture.

and molecular weight (see RODBARD and CHRAMBACH 1971). If the same strains are run in 3% gels, slow bands appear in *D. pseudoobscura* and *D. miranda* that are not seen in *D. persimilis*. These are the same bands that migrate faster than the normal *Est-5*$^{1.07}$ bands in 8% gels.

The *Est-5*$^{1.12}$ allele was also found to be common to all three species. Here again it was found that additional faster bands were present in 8% gels and the additional band in *D. persimilis* was faster than those of either *D. pseudoobscura* or *D. miranda*. This same result was also found when comparing *Est-5*$^{1.16}$ mobilities of *D. pseudoobscura* and *D. persimilis*.

Examination of the three species in 8% gels showed that all strains produce two bands connected by a continuum of stain similar to those shown in Figure 2. The mobilities of the slower and faster bands were found to be highly correlated. Strains containing the alleles *Est-5*$^{.85}$, *Est-5*$^{.95}$, *Est-5*$^{1.00}$, *Est-5*$^{1.04}$, *Est-5*$^{1.07}$, *Est-5*$^{1.12}$ and *Est-5*$^{1.16}$ in *D. pseudoobscura*, *Est-5*$^{1.07}$, *Est-5*$^{1.12}$, *Est-5*$^{1.16}$, *Est-5*$^{1.20}$ and *Est-5*$^{1.29}$ in *D. persimilis* and *Est-5*$^{1.00}$, *Est-5*$^{1.07}$ and *Est-5*$^{1.12}$ in *D. miranda* were studied in 8% acrylamide gels. The faster band and slower band mobilities were found to be correlated in all of these strains. Figure 3 also shows this correlation in esterase-5 mobility of faster band and slower band in both *D. pseudoobscura* and *D. persimilis*.

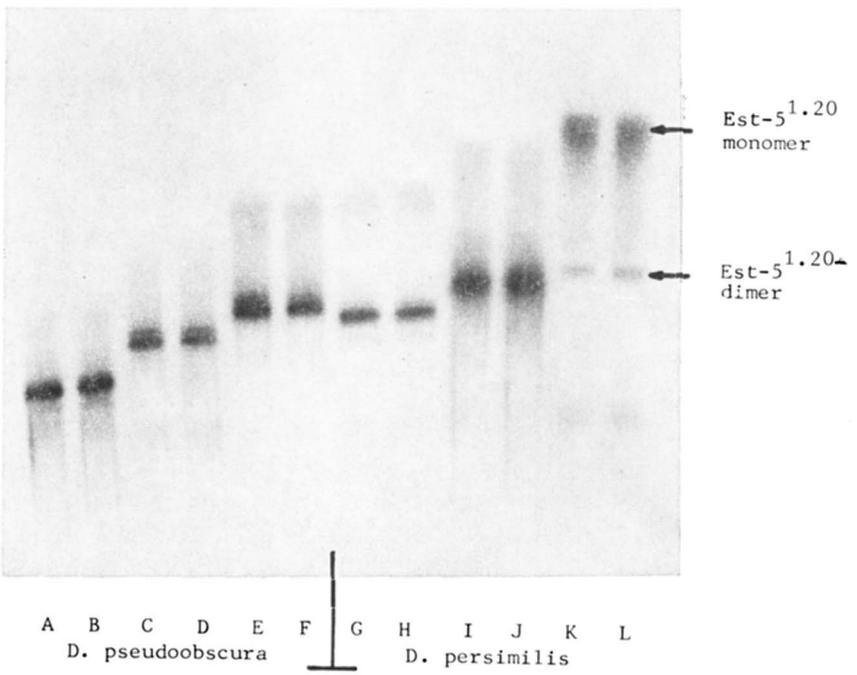


FIGURE 3.—Correlation between monomer (faster mobility) and dimer (slower mobility) forms of esterase-5 in *D. pseudoobscura* and *D. persimilis*. Pockets A and B contain *Est-5*$^{.85}$ C and D contain *Est-5*$^{1.00}$ and E and F contain *Est-5*$^{1.12}$ of *D. pseudoobscura*. Pockets G and H contain *Est-5*$^{1.07}$, I and J contain *Est-5*$^{1.14}$ and K and L contain *Est-5*$^{1.20}$ of *D. persimilis*. Origin is toward bottom of picture.

The high correlation between mobilities of the slow and fast bands suggests that they may have a common genetic basis. It is possible that the two bands are due to different degrees of polymerization of the *Est-5* gene product. COBBS (1976) has reported strains of *D. pseudoobscura* which apparently contain both dimer and monomer forms of *Est-5*. The phenotypes reported by COBBS (1976) showed a continuum of light stain between the two bands as do the phenotypes reported here. This continuum of stain between the two bands could result from the dissociation of subunits into lower molecular weight forms.

Evidence that the two bands are the same gene product is obtained by electrophoresing at high field strength. When 8% acrylamide gels are run at 32.3 volts/cm it is found that proportionally more of the esterase-5 activity is present in the faster band than in the slower band. At a field strength of 36 volts/cm nearly all the activity is in the fast band with a trailing smear. At 9 volts/cm nearly all the activity is present in the slow band. These results suggest that the slow band is dissociated by higher voltage to form the faster band.

It was suspected that these bands may be the dimer and monomer forms of the *Est-5* gene product, similar to the two-banded phenotypes reported by COBBS (1976). In order to test this hypothesis, molecular weights were compared by the method of Ferguson plots (see RODBARD and CHRAMBACH 1971; and COBBS 1976 for references). A graph of the relation between the log of relative mobility, $R_m$, and gel concentration is shown in Figure 4 for the fast and slow bands of a number of *Est-5* alleles in *D. pseudoobscura* and *D. persimilis*. Linear regression analysis indicates that the Ferguson plots of the suspected dimer bands shown in Figure 4b all have essentially the same slope of zero. The suspected monomer bands also all show essentially the same slope (Figure 4a). Consequently all these suspected monomer bands have the same molecular weight. In Figure 2 there are also bands slightly faster than the suspected dimers in pockets A, D, G and I. The molecular weight of these bands was found to be the same as the suspected dimers. These bands could be conformational forms of the dimer. The bands of $Est\text{-}5^{1.18MD}$ have been studied previously by COBBS (1976), and it was found that the slow band had twice the molecular weight of the fast band. In view of Figure 4, it appears that this finding is true for all the strains studied here.

This evidence is consistent with the hypothesis that the slower and faster bands in 8% gels are dimer and monomer forms of the *Est-5* gene product. Of course it is possible that the bands are some other level of polymerization such as tetramer and octomer. The fact that heterozygotes produce three bands with the molecular weight of the slow band suggests that the two bands present in homozygotes are dimer and monomer.

Further evidence that the slow and fast bands of these lines are products of the same gene comes from the examination of null alleles. Three strains homozygous for $Est\text{-}5^{null}$ alleles were examined. One of these was isolated from a Mesa Verde population by PRAKASH and MERRITT (1972) and two were isolated from a Riverside, California population by COBBS (1975). Under the electrophoretic conditions used here, none of the three null strains showed activity for either the low or high molecular weight bands. This suggests a common genetic basis for these

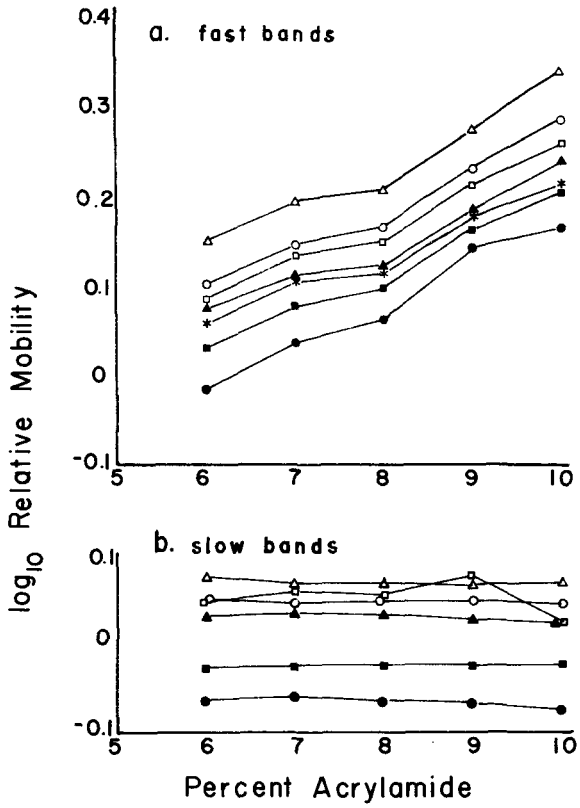


FIGURE 4.—Relation between $\log_{10}$ of relative mobility to slow band of *Est-5*$^{1.00}$ of *D. pseudoobscura*. Figure 4a is fast bands and figure 4b is slow bands. In both figures the symbols represent the following *Est-5* strains:

*D. pseudoobscura strains* *D. persimilis strains*

● — *Est-5*$^{.85}$ △ — *Est-5*$^{1.17}$
■ — *Est-5*$^{.95}$
★ — *Est-5*$^{1.00}$
▲ — *Est-5*$^{1.07}$
□ — *Est-5*$^{1.12}$
○ — *Est-5*$^{1.18MD}$

two bands. Whatever the reason for the lack of esterase-5 activity in the null lines, it is unlikely that a mutant would affect two loci simultaneously. Therefore, a single gene coding for these two bands seems a reasonable explanation.

There is the possibility that the difference in monomer mobilities results from an epigenetic modification of the charge on the binding region in one or the other species. In order to test this hypothesis, genetic studies were done. First, *D. persimilis Est-5*$^{1.07}$/*Est-5*$^{1.07}$ females were crossed to *D. pseudoobscura Est-5*$^{0.95}$ males. Second, *D. pseudoobscura Est-5*$^{1.07}$/*Est-5*$^{1.07}$ females were crossed to *D. persimilis Est-5*$^{1.20}$ males. For each of these two crosses at least 10 male $F_1$ progeny were examined in 8% gels at 16.1 v/cm. All $F_1$ males examined exhibited matriclinal inheritance of both monomer and dimer mobilities. This information alone

makes an epigenetic model seem highly unlikely. For each of the above crosses, $F_1$ females were backcrossed to both of the parental strains. Then for each back cross at least 20 $F_2$ males were examined. There was complete linkage of the *Est-5*$^{1.07}$ monomer and dimer mobilities in all crosses. These experiments suggest that the difference in monomer mobilities between *D. pseudoobscura* and *D. persimilis* is caused by a difference at the *Est-5* locus and is not the result of an epigenetic factor.

The high and low molecular weight bands in these gels are most easily interpreted as dimer and monomer forms of the *Est-5* gene product. The fact that *D. persimilis* alleles *Est-5*$^{1.07}$, *Est-5*$^{1.12}$ and *Est-5*$^{1.16}$ all have monomers with faster mobility than the corresponding alleles in either *D. pseudoobscura* or *D. miranda* suggests that there may be a fixed difference between these species at this locus. Specifically, an amino acid difference that changes monomer mobility but does not change dimer mobility is indicated.

There are at least two mechanisms by which an amino acid difference could cause the above effect. First, the change could occur in the region of the monomer which is involved in dimerization. The charge change in the monomer may not be apparent in the dimer because the site of change is no longer accessible to the ionic environment or the dissociation constants are altered by the dimerization or both of these. Second, the change may occur at a site, not necessarily in the dimerizing region, which causes the monomer but not the dimer to assume a different conformational structure with a higher charge. These two mechanisms are not mutually exclusive and some mixture of these two models may be operating. The two models will be referred to as the binding site model and the conformational change model.

*Measurement of charge ratios*

In order to obtain evidence that there is a fixed difference between *D. persimilis* and the other two species at the *Est-5* locus, a study of charge ratios was done. We will first present expressions for the interpretation of charge ratios in terms of the binding site and conformational change models and then present observations relating to the charge ratio. Methods for estimating net charge on the surface of proteins have been given by RODBARD and CHRAMBACH (1971). Their methods are based on several assumptions about the structure of proteins which are discussed in their paper. Consequently, estimates of absolute net charge have dubious accuracy. Fortunately, we may measure the ratio of net charge of two proteins with greater accuracy than we can measure absolute charge. Formulae for estimating the ratio of charges of two proteins are easily obtained from the formulae of RODBARD and CHRAMBACH (1971) and are given in the appendix to this paper. The values of charge ratios were estimated using equation (4) of the appendix and a single pocket of an electrophoretic gel. For each allele, several such estimates were made in gels of different concentration and the average of these was taken as the estimated value of the charge ratio.

Denote the ratio of the charge of monomer to dimer for allele $i$ by $CR_i$. Now denote the contribution of each monomer to the charge of the dimer by $b_i$. It is

reasonable to assume the two monomer subunits contribute equally to the charge of the dimer because heterozygotes produce three dimer mobilities with the intermediate mobility equidistant from the other two mobilities. This was found to be true by measuring the distances in a number of *Est-5* heterozygous females. Now let the total charge on the monomer be $a_i + b_i$ where $a_i$ represents the charge change due to dissociation and conformational change. Here $a_i$ may be positive, negative, or zero. With these definitions we may then write

$$CR_i = \frac{a_i + b_i}{2b_i} \tag{1}$$

The above expression for $CR_i$ is valid for both the binding site and conformational change models. The two models differ by attributing the value of $a_i$ to different causes. Under either model, if the values of $a_i$ for one species were different in a consistent way from another species, then a fixed difference between the species is suggested. Consequently, we are most directly interested in the comparison of $a_i$ values between species. We therefore introduced the variable $z_i$ which is defined as

$$z_i = v_i \{ 2CR_i - 1 \}$$

where $v_i$ is the charge of the dimer of allele $i$ divided by the charge of the dimer of *Est-5*$^{1.00}$ of *D. pseudoobscura*. As seen in Figure 4b and to be reported elsewhere (COBBS and PRAKASH, in preparation), the value of $v$ is very close to or indistinguishable from $R_m$ at any concentration of gel for esterase-5 dimers. We will employ $R_m$ as an estimate of $v$. Since $z_i = a_i/b_{1.00}$, where $b_{1.00}$ is the $b$ value for *Est-5*$^{1.00}$ of *D. pseudoobscura*, comparison of $z_i$ values is essentially the comparison of $a_i$ values.

Values of $z_i$ for a number of alleles in *D. pseudoobscura* and *D. persimilis* are given in Table 1. The $z_i$ values of *D. persimilis* are uniformly higher than those of *D. pseudoobscura*. A Mann-Whitney test shows the two species to be significantly different with respect to this attribute. The data supports the hypothesis that there is a fixed difference between these species at this locus. Values of $z$ for *D. miranda* were found to be essentially the same as those for *D. pseudoobscura*.

TABLE 1

*Values of* z *for strains of* D. pseudoobscura *and* D. persimilis

| *D. pseudoobscura* Allele | $z$ | *D. persimilis* Allele | $z$ |
|---|---|---|---|
| *0.85* | 0.436 | *1.07* | 0.614 |
| *0.95* | 0.472 | *1.12* | 0.630 |
| *1.00* | 0.452 | *1.17* | 0.656 |
| *1.04* | 0.448 | *1.18* | 0.690 |
| *1.07* | 0.452 | *1.20* | 0.630 |
| *1.12* | 0.450 | *1.24* | 0.656 |

*Levels of intraspecies variation*

The ability to examine monomer mobilities may allow us to distinguish more alleles in a sample than can be distinguished on the basis of dimer mobility alone. If this is true, it should be possible to find alleles having the same dimer mobility but different monomer mobilities. Such cases were found in both *D. persimilis* and *D. miranda*. An example of two *Est-5*$^{1.17}$ dimer alleles with different monomer mobilities is shown in Figure 5. Similarly an example of *D. miranda Est-5*$^{1.12}$ dimer mobilities with different monomer mobilities is shown in Figure 6. The strains shown in Figure 6 give two monomer bands under these conditions. These patterns were found to be repeatable differences between strains and will be interpreted as allelic differences of the esterase-5 locus. The two monomer bands present in Figure 6 were found to have the same molecular weight by using Ferguson plots similar to those shown in Figure 4a. The two monomer bands may represent two conformational forms of the monomer.

Upon examination of all the strains listed in the MATERIALS AND METHODS there were 3 alleles in *D. persimilis* and 2 alleles in *D. miranda* which could not be detected by dimer mobility, but were detected by monomer mobility. The *Est-5*$^{1.33}$ allele of *D. persimilis* produced only monomers under the electrophoretic conditions used here and will therefore be referred to as *Est-5*$^{1.33M}$. In addition, there were new alleles found which were previously undetected. Most of these were

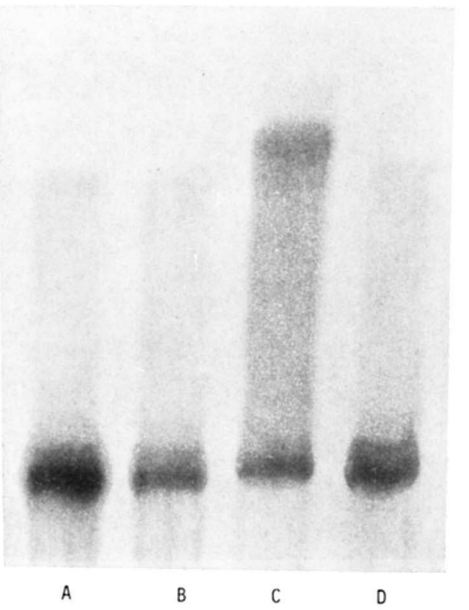


FIGURE 5.—*Est-5*$^{1.17}$ electrophoretic phenotype of *D. persimilis* in 8% acrylamide gel. Pockets B and D are both from one strain with a slow monomer mobility, Pocket A is another strain with slow monomer mobility, Pocket C is a strain with fast monomer mobility. Origin is toward bottom of picture.

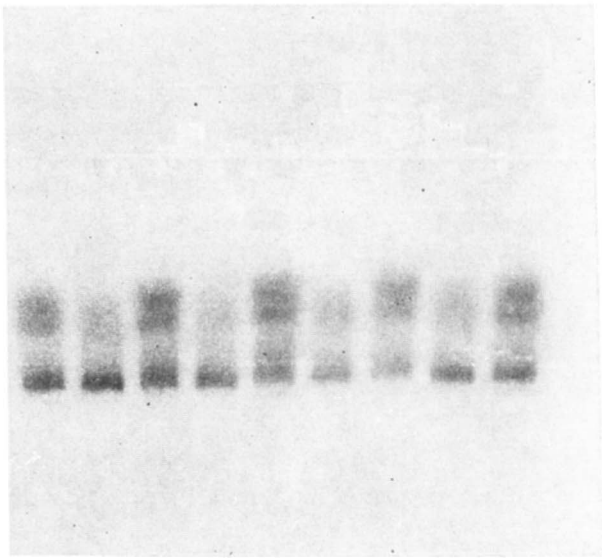


FIGURE 6.—*Est-5*$^{1.12}$ electrophoretic phenotypes of *D. miranda* in 8% acrylamide gel. Pockets A, C, E, G and I are all a single strain having normal monomer mobility. Pockets B, D, F, and H are all a single strain having slow monomer mobility. (Note the presence of two monomer bands in both strains.) Origin is toward bottom of picture.

very small differences in $R_m$ that were not resolved with shorter periods of electrophoresis. For example, a difference in $R_m$ of less than 1% is found in *D. persimilis*. Such small differences were often undetected in previous studies, as repeated alternating pocket tests and extended length of migration are necessary to be confident of such differences.

A summary of the alleles and their frequency in the samples is given in Table 2 for *D. persimilis* and Table 3 for *D. miranda*. The $R_m$ values were calculated using the *Est-5*$^{1.00}$ allele of *D. pseudoobscura* as the standard mobility. Homologies with the allelic classification of PRAKASH (1969) for *D. persimilis* are given in Table 2. If the data of Table 2 are compared with those of PRAKASH (1969) we find a factor of increase in number of alleles of two. If we compare the number of alleles detected in the two gel conditions we find that the 8% gels resolved 1.5 and 1.8 times more alleles than 5% gels in *D. persimilis* and *D. miranda*, respectively. The allele frequency distribution of data from 8% gels is different than that for 5% gels. In both species the more common classes were the ones which broke into more classes. For example, the 1.17 mobility class of *D. persimilis* separated into 5 allelic classes in 8% gels. The less frequent allelic classes rarely split into more classes. Consequently, the data obtained using 8% gels indicate that neither species has an extremely common allele at this locus. A similar result has also been found in *D. pseudoobscura* (MCDOWELL and PRAKASH 1976).

TABLE 2

*Allele frequency distributions in* D. persimilis

| Allele of PRAKASH 1969 | Relative mobility in 5% acrylamide | Relative mobility in 8% acrylamide | Frequency of alleles |
|---|---|---|---|
| *1.07* | *1.07* | *1.07* | 0.093 |
| *1.12* | *1.12* | *1.12* | 0.047 |
| | *1.14* | *1.14* | 0.047 |
| *1.16* | *1.16* | *1.16* | 0.023 |
| *1.20* | *1.17* | *1.17*+sm | 0.023 |
| | | *1.17* | 0.349 |
| | | *1.17*+fm | 0.023 |
| | | *1.18* | 0.209 |
| | | *1.18*+fm | 0.070 |
| | *1.20* | *1.20* | 0.047 |
| *1.29* | *1.24* | *1.24* | 0.023 |
| *1.33* | *1.33* | *1.33*m | 0.047 |

The symbols +fm and +sm stand for fast and slow monomer mobility, respectively.
Number of strains = 43.
Number of alleles = 12.
Effective number of alleles = 5.3.
Heterozygosity = 0.81.

## *Levels of interspecies variation*

A study was done to reveal differences between the *Est-5*$^{1.07}$ mobilities of *D. pseudoobscura* and *D. miranda*. Six *Est-5*$^{1.07}$ lines were selected from each of these two species and were compared under a large number of buffer conditions. The buffer conditions used included the following: (1) 0.01 M, 0.05 M and 0.1 M Tris-

TABLE 3

*Allele frequency distributions in* D. miranda

| Relative mobility in 5% acrylamide | Relative mobility in 8% acrylamide | Frequency of alleles |
|---|---|---|
| *1.00* | *1.00* | 0.182 |
| *1.03* | *1.03* | 0.030 |
| *1.04* | *1.04* | 0.030 |
| *1.07* | *1.06* | 0.091 |
| | *1.07*+sm | 0.091 |
| | *1.07* | 0.212 |
| | *1.08* | 0.061 |
| *1.12* | *1.11* | 0.091 |
| | *1.12*+sm | 0.061 |
| | *1.12* | 0.030 |
| *1.15* | *1.15* | 0.121 |

The symbol +sm stands for slow monomer mobility.
Number of strains = 33.
Number of alleles = 11.
Effective number of alleles = 7.8.
Heterozygosity = 0.872.

borate EDTA, pH = 9.1; (2) 0.05 M Tris-HCl, pH = 7.5, 8.5; (3) 0.005 M $NaPO_4$, pH = 7.0, 7.5, 8.0; (4) 0.05 M Tris-borate, pH = 9.1; (5) 9 mM Tris-3mM citrate pH 6.8, in 12% starch gel. In all the above conditions the *Est-5*$^{1.07}$ dimer mobilities of both *D. pseudoobscura* and *D. miranda* were not perceptively different. Also the other mobilities common to both *D. pseudoobscura* and *D. miranda* did not appear to be different in 0.1 M TBE and 8% gel concentration.

## DISCUSSION

A question which naturally arises from the data presented here is whether or not there is information available to distinguish between alternative models proposed to explain the existence of protein polymorphism. Much literature has been directed toward deciding if natural selection plays an important role in determining the allele frequencies at loci such as the one studied here. As pointed out by EWENS (1972), the hypothesis that selection exists is not sufficiently precise to formulate an appropriate test to use in making such decisions. Consequently, most tests pertaining to the above question have in one way or another been limited to a test of goodness-of-fit to a selectively neutral model of one type or another.

Even a test of goodness-of-fit to a selectively neutral model may be impossible to interpret in terms of our original question of whether or not natural selection is involved. All selectively neutral models are based on assumptions which may not be satisfied in natural populations. For example, the now famous infinite alleles model of KIMURA and CROW (1964) assumes that each new mutant is unique, there is no subdivided structure in the population, and the population is in an equilibrium state. The first of these assumptions may be approximately correct for a locus such as esterase-5; however, the second and third assumptions may be considerably off. Consequently, a significantly poor fit to a neutral model may indicate either that some type of selection is occurring or that some other assumption of the model is not correct.

In spite of these and other problems in interpreting goodness-of-fit tests to neutral models, such tests may be instructive. The nature of deviation, if any, may indicate which assmptions are incorrect. For example, if allele frequencies were too uniform to fit the infinite alleles model, it would be difficult to attribute this to subdivided populations and possibly some form of balancing selection would be implied. Alternatively, if allele frequencies were too skewed, then both selection and subdivided populations could be responsible. Further investigation may be able to distinguish between these possibilities.

Another difficulty arises from the fact that most of the forces which may affect allele frequencies are not mutually exclusive. Some mixture of forces affecting allele frequency distributions may be impossible to detect from the study of allele frequencies alone.

With the above factors in mind, a test of goodness-of-fit of our data to the infinite alleles model was done. The test derived by EWENS (1972) was applied to the data of Tables 2 and 3. Neither species deviated significantly from the expectations of the infinite alleles model. The values of $L$, $F$ $(v_1v_2)$ were $-0.37$, 0.86

(64.4, 12.7) and 1.22, 1.99 (68.4, 11.8) corresponding to probabilities of approximately 0.79 and 0.08 for *D. persimilis* and *D. miranda*, respectively. Examination of the *L* values for these two species indicates that *D. persimilis* fits the model well and *D. miranda* fits the model only moderately well. A similar test on esterase-5 of *D. pseudoobscura* gives a moderately good fit ($p \sim 0.12$) to the infinite alleles model (McDowell and Prakash 1976). We may therefore assume that the allele frequencies are in rough agreement with the expectations of the infinite alleles model. These results suggest that many of the alleles at this locus may be effectively neutral with respect to fitness. This conclusion is not proven, as there are many selective models that could give the same conclusion. Also, the test applied here is not extremely powerful in terms of detecting slight deviations from expected distribution (Ewens 1972).

Other evidence also suggests that esterase-5 alleles may be neutral. The experiments of Yamazaki (1971) failed to find selective differences between two *Est-5* alleles in laboratory environments. Also, Cobbs (1975) failed to find any selective disadvantage even with null phenotypes of this locus in *D. pseusoobscura*. Cobbs (1976) has found alleles in natural populations which have unusual dimerizing properties. The existence of these alleles suggests that the enzyme may tolerate changes in quaternary structure. *All these reports in addition to the present study suggest that mutation to effectively neutral alleles and genetic drift may be a predominant force in maintaining the extensive polymorphism at the esterase-5 locus.*

A curious finding is that the alleles of *D. persimilis* have undergone at least one complete substitution while *D. pseudoobscura* and *D. miranda* still share a number of apparently common alleles. The *Est-5*$^{1.00}$, *Est-5*$^{1.03}$, *Est-5*$^{1.04}$, *Est-5*$^{1.07}$, *Est-5*$^{1.12}$ alleles all occur in both *D. pseudoobscura* and *D. miranda*. This observation is surprising, as *D. pseudoobscura* and *D. persimilis* are much more closely related to each other than to *D. miranda*. Observations of interspecies hybrids, karyotype differences, and morphology all show *D. pseudoobscura* and *D. persimilis* to be more closely related than *D. pseudoobscura* and *D. miranda* or *D. persimilis* and *D. miranda* (Dobzhansky and Tan 1936).

It seems unlikely that this pattern of interspecies difference could have resulted from mutation and genetic drift alone. It is possible that some of the differences between the *Est-5* of *D. persimilis* and the other two species is the result of natural selection. This could happen even if most of the extant variation of this locus is selectively neutral—or nearly neutral. Clearly, more work, both experimental and theoretical will be necessary to obtain a more complete understanding of the intra- and interspecies differences at this locus.

One must be extremely cautious in attempting to extend the results reported here to other enzyme loci. The esterase-5 locus was chosen for this study because it was previously known to be highly polymorphic. It certainly should not be assumed to be an average or typical enzyme locus.

We are grateful to Rod Norman for assistance with the computer program and to Nancy Istock for expert technical assistance. This work was supported by grant GM 19217 and by Genetics Training Grant GM-00658–15, both from the Public Health Service.

APPENDIX

A method for measuring the ratio of the charge of two proteins can be derived from the equations of RODBARD and CHRAMBACH (1971) and other well known equations of electrophoresis. The charge ratio ($CR$) will be estimated from the relative electrophoretic mobility of the two proteins, $Pi$ and $Pj$, under identical experimental conditions. In the application used here, the relative mobilities of the proteins were measured in the same pocket of an electrophoretic gel.

An expression for free electrophoretic mobility derived by GORIN (1939) and also given in TANFORD (1961, eq. 24–8) may be written as

$$U_i = \left\{\frac{Q_i}{6\pi\eta Ri}\right\} \left\{\frac{1+\kappa r)}{1+\kappa\,(R_i+r)}\right\} f(\kappa R_i) \tag{1}$$

where

$U_i$ = free mobility of $P_i$
$Q_i$ = charge of $P_i$
$R_i$ = radius of $P_i$
$f(\kappa R_i)$ = Henry's function
$r$ = average radius of counter ions
$\eta$ = viscosity.

Here $\kappa$ is the DEBYE-HUCKLE constant for the buffer and may be calculated from knowledge of the Dielectric constant, absolute temperature and ionic strength. The value of $\kappa$ was calculated as

$$\kappa = \left(\frac{8\,\pi N\,\varepsilon^2}{1000\,DkT}\right)^{1/2} I^{1/2}$$

where

$N$ = Avogadro's number
$\varepsilon = 4.80 \times 10^{-10}$ electrostatic units
$D$ = Dielectric constant
$k$ = Boltzman's constant
$T$ = absolute temperature
$I$ = ionic strength

Equation (i) of RODBARD and CHRAMBACH (1971) may be rewritten to obtain an expression for electrophoretic mobility, $M$, in a gel of concentration $C$ which may be written as

$$M_i = 10^{-K_i C} U_i \tag{2}$$

where

$M_i$ = electrophoretic mobility of $P_i$
$K_i$ = retardation coefficient of $P_i$
$U_i$ = free electrophoretic mobility as defined in equation (1)
$C$ = concentration of gel.

Using equations (1) and (2) we may obtain an expression for the ratio of the mobilities of proteins $i$ and $j$ in a gel of concentration $C$. This expression may be written as

$$\frac{M_i}{M_j} = 10^{(K_j - K_i)C} \left\{\frac{Q_i}{Q_j}\right\} \left\{\frac{R_j}{R_i}\right\} \left\{\frac{1+\kappa(R_j+r)}{1+\kappa(R_i+r)}\right\} \left\{\frac{f(\kappa R_i)}{f(\kappa R_j)}\right\} \tag{3}$$

Let $CR = Q_i/Q_j$, $R_m = M_i/M_j$ and $K = K_i - K_j$. A slight rearrangement gives

$$CR = Rm\,\{10^{KC}\} \left\{\frac{R_i}{R_j}\right\} \left\{\frac{1+\kappa(R_i+r)}{1+\kappa(R_j+r)}\right\} \left\{\frac{f(\kappa R_j)}{f(\kappa R_i)}\right\}. \tag{4}$$

Equation (4) provides the means to estimate the charge ratio of two proteins without knowing the absolute charge of either protein. The radius of a protein may be estimated from its molecular weight ($MW$) by equation (vii) of RODBARD and CHRAMBACH (1971) which is given as (5) below.

$$R = \{3\,(MW_i\bar{\nu})/4\pi N\}^{1/3} \tag{5}$$

where

$MW_i$ = molecular weight of $P_i$
$\bar{\nu} = 0.74$

Equations (4) and (5) now permit one to estimate $CR$ from knowledge of $Rm$, $K$, $MW_i$, $MW_j$, $C$ and $r$. The values of $C$ and $r$ are experimental constants. $Rm$ may be measured directly and $K_i$, $K_j$, $MW_i$, and $MW_j$ may be determined from FERGUSON plots as described by RODBARD and CHRAMBACH (1971). The values of $MW_i$ and $MW_j$ may also be estimated by the methods of COBBS (1976), which are only a slight modification of the methods of RODBARD and CHRAMBACH (1971). The value of $K$ may be estimated as the slope of the relationship between log $(Rm)$ and $C$. The value of Henry's function may be evaluated numerically or through an approximation given by equation (vii) in appendix V of the paper by RODBARD and CHRAMBACH (1971).

## LITERATURE CITED

COBBS, G., 1975 Population phenomena at the esterase-5 locus in *Drosophila pseudoobscura*. Ph.D. dissertation. University of California at Riverside. ——, 1976 Polymorphism for dimerizing ability at the esterase-5 locus in *Drosophila pseudoobscura*. Genetics **82**: 53–62.

DOBZHANSKY, TH. and C. C. TAN, 1936 Studies on hybrid sterility. III. A comparison of the gene arrangements in two species, *Drosophila pseudoobscura* and *Drosophila miranda*. Zeits Indukt: Abst. Verebong **72**: 88–114.

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

GORIN, M. H., 1939 An equilibrium theory of ionic conductance. J. Chem. Phys. **7**: 405.

HUBBY, J. and R. C. LEWONTIN, 1966 A molecular approach to the study of genic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*. Genetics **54**: 577–594.

KIMURA, M. and J. F. CROW, 1964 The number of alleles that can be maintained in a finite population. Genetics **49**: 725–738.

LEWONTIN, R. C., 1974 *The genetic basis of evolutionary change*. Columbia University Press, New York and London.

MCDOWELL, R. E. and S. PRAKASH, 1976 Allelic heterogeneity within allozymes separated by electrophoresis. Proc. Natl. Acad. Sci. U.S. **73**: 4150–4153.

PRAKASH, S., 1969 Genic variation in a natural population of *Drosophila persimilis*. Proc. Nat. Acad. Sci. U.S. **62**(3): 778–784. ——, 1977 Genetic divergence in closely related sibling species *Drosophila pseudoobscura, Drosophila persimilis* and *Drosophila miranda*. Evolution (In press).

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.

PRAKASH, S. and R. B. MERRITT, 1972 Direct evidence of genic differentiation between sex ratio and standard gene arrangements of $X$ chromosome in *Drosophila pseudoobscura*. Genetics **72**: 169–175.

RODBARD, D. and A. CHRAMBACH, 1971 Estimation of molecular radius, free mobility and valence using polyacrylamide gel electrophoresis. Anal. Biochem. **40**: 95–134.

TANFORD, C., 1961 *Physical Chemistry of Macromolecules*. Wiley, New York.

YAMAZAKI, T., 1971 Measurement of fitness at the esterase-5 locus in *Drosophila pseudoobscura*. Genetics **67**: 579–603.

Corresponding editor: R. C. LEWONTIN