3-1969

An Analysis of Gene-Enzyme Variability in Natural Populations of Drosophila melanogaster and D. simulans

Stephen J. O'Brien
Cornell University, sobrien1@nova.edu

Ross J. MacIntyre
Cornell University

AN ANALYSIS OF GENE-ENZYME VARIABILITY IN NATURAL POPULATIONS OF *DROSOPHILA MELANOGASTER* AND *D. SIMULANS*

STEPHEN J. O'BRIEN AND ROSS J. MACINTYRE

Section of Genetics, Development, and Physiology, Cornell University, Ithaca, New York

INTRODUCTION

Two fundamental objects of research in experimental population genetics are the measurement of gene pool variability and the elucidation of both selective and nonselective mechanisms maintaining such variability. Relevant measurements include the number of breeding individuals in a population, the number of genetic loci in the organism under investigation, the potential number of alleles which can be maintained in a population, and the number of genes whose products are polymorphic. Estimates of the latter allow one to calculate the average number of heterozygous loci per individual in a population (Lewontin and Hubby, 1966).

Unfortunately, accurate techniques for estimating these parameters have not been developed. Consequently, geneticists have often relied upon mathematical models which are not immediately conducive to experimental analysis. Recently, however, several laboratories have applied the technique of gel electrophoresis of soluble proteins to directly estimate the average heterozygosity in natural populations. One of the purposes of this article is to expand these estimates to include a sample of 10 gene-enzyme systems from eight natural populations of *Drosophila melanogaster* and from two populations of a sibling species, *D. simulans*. In addition, an analysis has been made of a *D. melanogaster* population maintained in a population cage for 20 years.

We will show that natural populations of *D. melanogaster* are quite polymorphic—apparently much more so than comparable populations of *D. simulans*. The relevance and the evolutionary consequences of both the findings presented here and those of previous investigations will be critically reviewed.

MATERIALS AND METHODS

The Populations

Inseminated females from eight populations of *D. melanogaster* and two populations of *D. simulans* (which are sympatric with two of the *D.*

97

melanogaster populations) were collected from state parks in the eastern United States in the fall of 1966 by Dr. A. Chabora. A subculture of flies from the laboratory population, designated here as "Standard Kaduna," was kindly provided by Dr. Alan Robertson. The Standard Kaduna population cage is derived from flies caught in Kaduna, Nigeria, in 1949, and maintains an average breeding size of approximately 5,000 individuals. In addition, after the first generation in the laboratory, 10–15 progeny from each line were combined with progeny from the other lines from the same geographic population. The 10 populations were maintained in mass culture by trans-ferring 40–50 flies at each generation. Table 1 gives the geographic origin of the populations and the number of genes at a given locus in the original sample. Within this sample, one isofemale line could contain four possibly different alleles of an autosomal gene and three of a sex-linked gene. This of course assumes that each female carried the effective sperm from just one male at the time of capture. Hence, a population derived from 12 isofemale lines represents a maximum of 48 different genes at an autosomal locus and 36 at a sex-linked locus. If the collected females contained the effective sperm of more than one male, the number of genes sampled would of course be larger.

Sampling Procedures

The number of flies analyzed from each population depended upon the number of isofemale lines collected at that locale. Greater numbers were analyzed from populations set up from relatively large numbers of isofemale lines. Once a population was found to be polymorphic for a particular enzyme, sampling was discontinued. A population which was found to be monomorphic for an enzyme was usually tested again to avoid missing rare alleles.

We realize that this procedure would tend to miss rare alleles in the polymorphic populations, but since rare alleles contribute only slightly to the total population heterozygosity, this would not drastically alter our

TABLE 1

CHARACTERIZATION OF NATURAL POPULATIONS EMPLOYED IN THIS INVESTIGATION

| SPECIES | GEOGRAPHIC ORIGIN | NUMBER OF ISOFEMALE LINES | MAXIMUM NUMBER OF DIFFERENT GENES PER LOCUS | |
			Autosomal	Sex-Linked
D. melanogaster ...	Ceres, N.Y.	2	8	6
	Painesville, Ohio	12	48	36
	Mt. Sterling, Ohio	12	48	36
	Mammoth Cave, Ky.	12	48	36
	Red Top Mt., Ga.	12	48	36
	Columbia, Ga.	1	4	3
	Manning, S.C.	7	28	21
	Oxford, N.C.	5	20	15
D. simulans	Columbia, Ga.	8	32	24
	Manning, S.C.	1	4	3

final estimates. In those enzyme systems controlled by sex-linked genes, only females were analyzed to increase the sample size. The stock used as a reference in the electrophoretic analyses was a highly inbred line from Riverside, California, and is monomorphic for each of the enzyme systems. The enzymatic phenotypes of this stock are listed in Table 3, which is presented below.

Electrophoretic Procedures

Electrophoresis was conducted in horizontal starch gels using a continuous buffer system of 0.25 M tris-Cl, pH 8.6 in the bridges, and 0.05 M tris-Cl, pH 8.6 in the gel (Wright, 1963). This system was adequate for all enzymes except xanthine dehydrogenase. For this system the bridge and gel buffers described by Yen and Glassman (1965) were employed. Single flies were crushed on 5×7 mm strips of Whatman 3-MM filter paper previously soaked in gel buffer. The 10 enzymes examined are listed in Table 2, along with the gene symbols, loci, phenotypes of heterozygotes, and references. Modifications of referenced procedures include the following: with regard to NAD- and NADP-requiring enzymes (NAD: α-Gpdh-1, Adh, Xdh, Mdh-1; NADP: 6-Pgd and G-6-pd), 5 mg of coenzyme were added to 220 ml of gel buffer prior to gel preparation.

The voltage during the runs was kept below 8 v/centimeter. With Mdh, 6-Pgd, and G-6-pd, maximum of 6 v/centimeter was maintained. All electrophoresis was carried out at 4°C and all staining was done at 23°C. Stains for the various dehydrogenases were developed in the dark.

RESULTS

Analysis of Gene-Enzyme Systems

Both functional and genetic autonomy has been demonstrated for each gene-enzyme system listed in Table 2 either before or during the course of this investigation. Therefore, each enzyme is the product of a single and different structural gene. Homozygotes for each of these genes exhibit a single anodal zone of enzyme activity under appropriate conditions. Heterozygotes are characterized by either two or three such zones (see Table 2). The Mdh-1 is exceptional in that each major zone is invariably accompanied by a minor, more electropositive subband.

The raw data are presented in Table 3. Each phenotype is represented by one or two letters. Homozygous individuals are designated by a single letter, and heterozygotes are designated by two letters representing the alleles which they presumably carry. The allele producing the most electropositive (or "fastest") enzyme is designated by A. The allele designated as B produces the second "fastest" enzyme, and so on. Flies from the reference stock were always monomorphic for one band and thus were homozygous for one allele in each system. The characteristic allele of the control stock is indicated by a plus sign in Table 3. The body of the table

TABLE 2

GENE-ENZYME SYSTEMS OF *D. melanogaster* AND *D. simulans* ANALYZED IN THIS INVESTIGATION

Enzyme	Gene Symbol	Chromosome	Locus	Phenotype of Heterozygote*	Reference
6 phosphogluconate dehydrogenase......	*Pgd*	I	0.9	+	Young, 1966
Glucose-6-phosphate dehydrogenase......	*Zw* (*G-6pd*)	I	63	−	Young, Porter, and Childs, 1964; Young, 1966
α-glycerophosphate dehydrogenase-1......	*α Gpdh-1* (*Gdh*)	II	17.8	+	O'Brien and MacIntyre, 1968; Grell, 1967
Malic dehydrogenase......	*Mdh-1*	II	35.3	+	O'Brien, 1968, 1969)
Alcohol dehydrogenase......	*Adh*	II	50.1	+	Johnson and Denniston, 1964; Grell, Jacobson, and Murphy, 1965
Esterase-6......	*Est-6*	III	36	−	Wright, 1963
Larval alkaline phosphatase......	*Aph*	III	46.3	+	Beckman and Johnson, 1964*a*; MacIntyre, 1966*b*
Xanthine dehydrogenase......	*ry+* (*Xdh*)	III	52	+	Yen and Glassman, 1965
Leucine aminopeptidase-D......	*Lap-D*	III	98.3	−	Beckman and Johnson, 1964*b*; Falke and MacIntyre, 1966
Acid phosphatase-1......	*Acph-1*	III	101.3	+	MacIntyre, 1966*a*

* + or − refers to the presence or absence of an intermediate or "hybrid" enzyme in addition to both parental enzymes in the electropherogram of the heterozygote.

TABLE 3

NUMBERS OF INDIVIDUALS OF EACH PHENOTYPE FOR EACH ENZYME IN THE POPULATIONS SAMPLED IN THIS STUDY

Locus	Pheno-type	Con-trol	Melanogaster									Simulans	
			Ceres	Paines-ville	Mt. Sterling	Mammoth Cave	Manning	Ox-ford	Red Top Mt.	Colum-bia	Standard Kaduna	Colum-bia	Manning
Pgd	A	0	10	8	6	5	0	0	1	0	18	0	0
	AB	0	0	2	2	2	0	0	4	2	0	0	0
	B	+	0	2	5	5	9	12	8	7	0	15	15
G-6-pd	A	+	3	0	4	2	0	0	3	0	3	0	0
	AB	0	3	0	7	4	5	16	5	3	5	0	6
	B	0	4	11	3	7	8	6	1	6	5	12	8
αGpdh-1	A	+	3	9	16	18	8	0	10	9	18	12	0
	AB	0	0	6	0	0	0	3	0	0	0	0	0
	B	0	2	0	0	0	0	0	0	0	0	0	0
Mdh-1	A	0	0	0	1	0	0	0	0	0	0	0	0
	AB	+	0	0	5	0	0	3	0	0	0	0	0
	B	+	9	13	3	15	9	9	14	6	18	11	9
Adh	A	0	9	5	2	5	2	2	4	4	5	0	0
	AB	0	0	4	4	6	4	1	1	1	8	0	0
	B	0	0	3	6	1	5	6	3	0	3	6	6
	C	0	8	3	0	3	0	0	0	6	0	3	6
Est-6	A	+	0	3	5	3	7	6	1	6	9	6	0
	AB	0	0	6	9	4	0	1	3	0	5	1	0
	B	0	9	5	2	2	0	2	1	0	3	0	6
Aph	A	+	0	9	13	3	5	3	4	0	4	0	0
	AB	0	0	0	8	9	2	6	6	0	3	0	0
	B	0	0	0	1	6	2	0	4	0	8	11	6
Xdh	A	0	0	0	2	0	0	0	0	0	0	0	0
	AB	+	9	0	2	0	6	0	0	0	0	2	0
	B	0	0	17	18	19	0	6	13	6	18	7	6
	BC	0	4	0	0	4	4	5	0	0	0	0	0
Lap-D	A	+	3	10	13	19	4	12	14	5	5	0	6
	AB	0	1	14	4	2	2	9	5	2	10	0	0
	B	0	0	1	1	2	2	4	2	0	0	0	0
Acph-1	A	0	0	0	0	0	0	0	0	0	0	12	8
	B	+	10	15	18	15	9	19	15	6	15	12	9

NOTE.—See text for explanation.

TABLE 4

APPROXIMATE GENE FREQUENCIES CALCULATED FROM TABLE 3

Locus	Allele	Melanogaster									Simulans	
		Ceres	Paines-ville	Mt. Sterling	Mammoth Cave	Manning	Ox-ford	Red Top Mt.	Colum-bia	Standard Kaduna	Colum-bia	Manning
Pgd........	A	1.0	0.75	0.53	0.5	0	0	0.22	0.33	1.0	0	0
	B	0	0.25	0.47	0.5	1.0	1.0	0.78	0.67	0	1.0	1.0
G-6-pd.......	A	0.5	0	0.53	0.31	0.19	1.0	0.61	0.33	0.42	0	0
	B	0.5	1.0	0.47	0.69	0.81	0	0.39	0.67	0.58	1.0	1.0
αGpdh-1.......	A*	0.63	0.79	1.0	1.0	1.0	0.65	1.0	1.0	1.0	1.0	1.0
	B	0.37	0.21	0	0	0	0.35	0	0	0	0	0
Mdh-1........	A	0	0	0.40	0	0	0	0	0	0	0	0
	B*	1.0	1.0	0.60	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Adh........	A*	1.0	0.58	0.75	0	0	0	0	0	0	0	0
	B	0	0.42	0.25	0.67	0.36	0.30	0.56	0.75	0.56	1.0	1.0
	C	0	0	0	0.33	0.64	0.70	0.44	0.25	0.44	0	0
Est-6........	A*	1.0	0.44	0.59	0.56	1.0	0.70	0.5	1.0	0.68	0.59	1.0
	B	0	0.56	0.41	0.44	0	0.30	0.5	0	0.32	0.41	0
Aph........	A	1.0	1.0	0.75	0.42	0.68	0.65	0.5	0	0.37	0	0
	B	0	0	0.25	0.58	0.32	0.35	0.5	1.0	0.63	1.0	1.0
Xdh........	A	0	0	0.25	0	0	0	0	0	0	0.12	0
	B*	1.0	1.0	0.75	0.92	1.0	0.75	1.0	1.0	1.0	0.88	1.0
	C	0	0	0	0.08	0	0.25	0	0	0	0	0
Lap-D........	A*	0.75	0.69	0.81	0.88	0.64	0.65	0.79	0.75	0.67	1.0	1.0
	B	0.25	0.31	0.19	0.12	0.36	0.35	0.21	0.25	0.33	0	0
Acph-1........	A	0	0	0	0	0	0	0	0	0	1.0	1.0
	B*	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	0	0

* Most prevalent allele in the populations of D. melanogaster.

contains the number of flies of the different phenotypes from the samples of all the populations.

Table 4 gives estimates of gene frequencies which were calculated from the raw data of Table 3. Initially, each gene frequency was calculated in the conventional manner: two times the number of homozygotes plus the number of heterozygotes divided by twice the number of flies in the sample. For example, the frequency of the A allele of Lap-D in the Ceres population equals $[(2 \times 4) + 3]/2(4 + 3 + 1)$ or .69. The frequencies calculated in this way were then adjusted to more accurately reflect the frequencies in the original populations. Each frequency was multipled by the total number of genes represented at the locus in the original sample of inseminated females (see Table 1), and the product was rounded to the nearest whole number. This is an estimate of the actual number of these alleles in the original collection of isofemale lines. This value divided by the total number of possibly different genes at the locus gives the adjusted allele frequency. Continuing with the example of the A allele of Lap-D: .69 $\times$ 8 $\cong$ 6 A alleles, hence, $p\ (Lap$-$D^4) = 6/8$, or .75. Obviously, even the adjusted frequencies of Table 4 provide only a crude estimate of the gene frequencies in the original population.

Table 5 summarizes the data of Tables 3 and 4 for each population. The overall averages of the number of polymorphisms and heterozygosity per individual are also embodied in this table. The average heterozygosity per individual was calculated following Lewontin and Hubby (1966).

Several important results emerge from the data in Tables 3–5. First, there are high levels of enzyme polymorphism in natural populations of $D.$ $melanogaster$. On the basis of our estimates, these populations range from 30%–80% with an average of 54%. The frequency of polymorphisms in the Standard Kaduna laboratory population is virtually the same as the average

TABLE 5

PROPORTION OF LOCI, OUT OF 10, POLYMORPHIC AND PROPORTION OF THE GENOME
ESTIMATED TO BE HETEROZYGOUS IN AN AVERAGE INDIVIDUAL
FOR EACH POPULATION STUDIED

Species	Population	No. of Loci Polymorphic	Proportion of Loci Polymorphic	Proportion of Genome Heterozygous/ Individual
D. melanogaster.....	Ceres	3	0.30	0.134
	Painesville	5	0.50	0.212
	Mt. Sterling	8	0.80	0.339
	Mammoth Cave	7	0.70	0.271
	Red Top Mt.	6	0.60	0.269
	Columbia	4	0.40	0.163
	Manning	4	0.40	0.166
	Oxford	6	0.60	0.258
	Average (excluding Kaduna)	5.4	0.54	0.227
	Standard Kaduna	5	0.50	0.230
D. simulans	Columbia	2	0.20	0.070
	Manning	0	0.00	0.000

from the natural populations. However, *D. simulans* is markedly less poly-morphic for these gene-enzyme systems than *D. melanogaster*. The Manning population of *D. simulans* is monomorphic for all of the 10 enzymes, whereas four of the 10 were polymorphic in the sample of *D. melanogaster* from the same locale. This population of *D. simulans*, however, was derived from a single inseminated female. On the other hand, in the sample from the Columbia population of *D. simulans*, which was started with eight isofemale lines, only two of the 10 enzymes were polymorphic. Thus, even though four of the eight initial samples of *D. melanogaster* consisted of fewer inseminated females than the sample of *D. simulans* from Columbia, all showed higher levels of polymorphism.

Second, only certain enzymes are generally polymorphic (e.g., Lap-D and Adh). Others, such as Acph-1 or Mdh-1 tend to be completely or largely monomorphic in the sampled locales. Thus, the gene-enzyme systems do not seem to be equivalent in their contribution to the overall poly-morphism of the populations. It is interesting that the Standard Kaduna population, which presumably has had a very different recent history, exhibits the same polymorphisms as the natural populations. Furthermore, it is monomorphic for the much less variable Mdh-1 and Acph-1 enzymes. The data from the *D. simulans* populations are too scanty to make a mean-ingful comparison with the distribution of polymorphisms in the populations of *D. melanogaster*.

A third evident aspect of the data from Tables 3 and 4 is that in every system except three, there is a common allele, not only with respect to how many populations contain it, but also with respect to its frequencies in the polymorphic populations. The most common allele in such systems is marked with an asterisk in Table 4. Another somewhat surprising finding is that, despite the high levels of polymorphism found in the populations of *D. melanogaster*, in only one gene-enzyme system, Xdh, were more than two different alleles detected in the eight populations. In *D. pseudoobscura*, Hubby and Lewontin (1966) found one system (Esterase-5 with six alleles and several with four or three alleles.

It seems reasonable to assume, considering the studies of Wright and MacIntyre (1963), MacIntyre (1966a), and Courtright (1967) on gene homologies between *D. melanogaster* and *D. simulans*, that the gene-enzyme systems of the two sibling species in the populations studied here have had a common evolutionary origin. With the exception of Adh and Acph-1, the zone of activity seen in *D. simulans* flies is identical to a zone found in *D. melanogaster* populations. Thus, on the basis of electrophoretic migra-tion, a considerable (80%) conservatism of enzyme structure between the sibling species is indicated.

The data in Table 4, which are arranged with the more northern popula-tions on the left and the southern populations on the right, suggest that there is some geographic variation in two of the enzyme systems. In the 6-Pgd system, a north-south cline can be seen: The *A* allele, which is common in New York and Ohio, is relatively rare in the South. The transition appears

to be rather gradual. The same is true of the Aph system, again with the A allele common in the North and becoming less frequent in the middle and southern Appalachian regions. Variation in the other systems does not appear to be related to the geographic origin of the populations.

DISCUSSION

Recently, several estimates of the extent of gene pool variability in *Drosophila* and human populations have been made. These were based upon electrophoretic or serological analyses. Table 6 summarizes the results of both the previous studies and the present investigation. The estimate of the average frequency of polymorphic loci in the populations of *D. melanogaster* analyzed by us is somewhat higher than most previously reported. There may be several reasons for this. First of all, we mass cultured our samples before analysis, instead of maintaining them in separate isofemale lines. In addition, we tried to analyze our population samples as soon as possible after collection. The elapsed time between collection and analysis was never greater than 6 months. Presumably, mass culturing and rapid analysis would reduce the chances of random fixation of segregating alleles. A second reason for our relatively high estimate of polymorphism may be that we did not examine larval extracts for non-enzymatic proteins. It is known that the product of a single gene can migrate as two or more species during electrophoresis. This can result from different levels of autopolymerization (Smithies, Connell, and Dixon, 1962; Isemura and Kakiuchi, 1962), conformational states (Kitto, Wasserman, and Kaplan, 1966), differential saturation with cofactors (Ursprung, 1966; Jacobson, 1967), or alteration by gel components (Brewer, 1967). Unless genetic and/or functional autonomy can be demontrated, two apparently mono-morphic proteins might be erroneously scored as the products of two separate genes. An error such as this would cause an underestimate of the relative frequency of polymorphic loci. We have, in this study, included only enzyme systems whose functional autonomy is obvious and whose genetic autonomy has been demonstrated (see Table 2). The importance of the latter can be seen in the $Mdh-1$ system, in which the homozygotes show two bands after electrophoresis and staining. If conditions such as gel com-ponents, cofactors, starch concentration, voltage, etc., are not carefully controlled, nongenetic minor bands can be seen in α-$Gpdh-1$, $Acph-1$ and Adh as well. If, in each system, electrophoretic variants in which both bands vary in a parallel fashion were not found and subjected to genetic analysis, more than four structural genes might have been assigned to these systems.

On the other hand, the use of only genetically autonomous systems may introduce a third bias which could be as serious or more so than the second. Our choice of enzymes for this investigation was influenced to some extent by previously published information on *Drosophila* gene-enzyme systems. These publications, in general, describe two or more alleles in the particular systems (see references in Table 2). It might be argued that these systems

TABLE 6

SUMMARY OF PREVIOUS ESTIMATES OF POLYMORPHISM IN NATURAL POPULATIONS

Species	No. of Loci Sampled	No. of Populations Sampled	Freq. Polymorphic Loci	Average Heterozygosity	Reference
D. pseudoobscura	18	5	0.39	0.12	Lewontin and Hubby, 1966
D. ananassae	20	2	0.4–0.5	*	Johnson et al., 1966
Homo sapiens	10	1	0.3	*	Harris, 1966
H. sapiens	33	1	0.36	0.16	Lewontin, 1967
D. melanogaster	10	8	0.54	0.228	This report
D. nasuta	3	6	1.0	0.33†	Stone et al., 1968
D. ananassae	4	4	0.5	0.133†	Stone et al., 1968

* Data sufficient to calculate average heterozygosity were not presented in these publications.
† Average heterozygosity calculated by present authors from referenced data.

were reported because they are the most polymorphic of the *Drosophila* gene-enzyme systems detected by electrophoresis. Our choice of these would, then, lead to inflated estimates of gene pool variability. It may be, however, that these systems were investigated and reported not primarily because electrophoretic variants of the enzymes were readily found, but because the histochemical methods of enzyme detection were well known at the time and/or because the banding patterns were sharp and relatively easy to analyze. They would then not necessarily be anymore polymorphic than gene-enzyme systems which are more difficult to detect and/or analyze. At any rate, in our decision to use this or that gene-enzyme system, genetic autonomy initially was not a *sine qua non* condition. We chose them because the enzymes give clean patterns in our hands and the staining techniques had been well worked out. In fact, for two gene-enzyme systems, genetic uatonomy was an *a posteriori* observation. There were no known variants of *Mdh-1* or *α-Gpdh-1* when this investigation began.

Finally, our estimate may be high not because of methodological reasons, but because populations of *D. melanogaster* are more variable than those of other species. Our own data indicate that *D. melanogaster* is more variable than its sibling species, *D. simulans*. Until we have more comparative information about the size and structure of Drosophila populations, little can be said about this possibility that is not entirely conjectural.

Despite the fact that there are quantitative differences between the various estimates reported in Table 6, all the data indicate that the gene pools of Drosophila and human populations as measured by the techniques of gel electrophoresis or serology are extremely variable. The importance of this conclusion again underscores the need to reexamine the technique itself, especially to ascertain the relative importance of the limitations and biases associated with it.

One obvious limitation of electrophoresis is that it detects only the mutational differences between alleles which result in proteins whose migration in an electric field is affected. Many amino acid substitutions probably would not affect this property of the protein. Can we decide what proportion of possible mutational events are detectable by electrophoresis? One of us (O'Brien, unpublished) has calculated that 37.8% of all missense base substitutions insert an amino acid which would alter the net charge of the protein product at the pH employed in this investigation. The assumptions underlying such estimates are many, and it is only with knowledge of these assumptions that the calculations become useful. Thus, extrapolations to the extent of gene pool variability from electrophoretic analyses presently suffer from the ignorance we have of the resolving powers of the technique.

Another problem is the extent to which our sample of gene products is biased. Our extrapolated estimate assumes that all loci are equivalent with respect to tolerance of polymorphism. That this is probably not true can be seen from our own data in which some gene products are almost always polymorphic (e.g., Lap-D and Adh) while others are relatively mono-

morphic (e.g., Acph-1 and Mdh-1). Also, we have examined only soluble or readily solubilized enzymes in this study. Indeed, at the present time, only hydrolytic and oxidative enzymes (or those enzymes that can be readily linked to one of these in a reaction system) can be histochemically demonstrated (Shaw and Koen, 1968). Gene products whose variability we cannot detect at the present time include other classes of soluble enzymes, enzymes bound to membranes, complex enzyme aggregates, structural proteins, polypeptide hormones, and species of RNA which do not serve as templates for protein synthesis. In addition, there may well be gene products involved in the control of development and physiology of which we are not yet aware. Is there any reason to suspect that these other gene products are any more or less variable than our sample of 10 enzymes? One might expect other soluble enzymes to vary in a parallel fashion. Hubby and Lewontin (1966) found that soluble larval proteins varied to about the same degree as eight other soluble enzymes. Structural proteins, membrane bound enzymes, or enzymes which are parts of large aggregates (which must have a proper "fit" as well as be catalytically active) may not exhibit as much variability as the proteins we have examined. Similarly, we might expect less variability in allosterically controlled enzymes which require a functional allosteric site as well as a functional active site. We do not know if any of the 10 enzymes listed in Table 2 have allosteric sites. Nothing can be said about variability in nonmessenger RNA or in the products of the suspected regulatory genes. Nor do we know what proportion of the genome codes for soluble enzymes or proteins, what proportion for structural proteins, and so on. Thus, in view of our present ignorance concerning the above questions, the extrapolation to the genome from 10 genes which code for 10 soluble oxidative and hydrolytic enzymes seems presumptious. Obviously, we need estimates of gene pool variability from samples of the different classes of gene products. Lewontin's recent analysis (1967) is a step in the right direction.

Nevertheless, it is possible, if not probable, that the present estimates of gene pool variability (see Table 6) are reliable. What are the mechanisms responsible for maintaining the possibly high levels of variability in natural populations? These have been listed by Lewontin and Hubby (1966). Our data are relevant to certain of these mechanisms which will be discussed below.

Selective Neutrality of Isoalleles

The occurrence of nucleotide substitutions which are adaptively neutral would probably not be effective by themselves in maintaining the variability we observed. Kimura and Crow (1964) have shown that drift tends to randomly fix alternate alleles in all populations except those of inordinately large breeding size.[1] On the other hand, the models of Wright (1966) and

---

[1] The probability that an individual is homozygous at a locus is $1/(4\,N_e\,\mu + 1)$, where N_e is effective breeding size and μ is the mutation rate (Kimura and Crow, 1964).

Kimura (1968) suggest that isoallelic variation could also be maintained by high mutation pressures for these isoalleles. Whatever the mode of maintenance, the Kimura-Crow-Wright hypothesis of selective neutrality carries with it certain predictions which can be tested by examination of our data. Selective neutrality of isoalleles means the alleles are interchangeable from population to population in that one allele may be predominant or fixed in one population, another allele is fixed in a neighboring population, and a third is predominant or fixed in the next population. In our analysis, the only system which exhibits fixation of alternate alleles is *Pgd* (see Table 4), and that looks suspiciously like a north-south cline. Of the others there is not a single case where contrasting alleles have become fixed in different populations. The only apparent exception is the *Aph* locus in the Columbia population, which cannot be taken too seriously since this population descended from a single inseminated female. In fact, in most cases there is a predominant allele for each locus (see asterisk, Table 4) in all the populations, an observation that is in direct opposition to the predictions of selective neutrality. The limited variation of allele frequencies is reminiscent of a similar pattern of allele frequencies for the ABO blood antigens (Brues, 1954).

Migration and Selective Neutrality

Kimura (1968), in recognizing the enigma of having drift and mutation responsible for the large amount of variability, has further suggested that migration between *Drosophila* populations would ease the demands for both high mutation pressure and large breeding size. Presumably, *D. melanogaster,* which is commensal with man and his garbage, would be exposed to extensive migration from other populations. The analysis of enzyme variability, however, in a Standard Kaduna population of *D. melanogaster* which has been in a laboratory cage for 20 years, shows that as much variability exists in that cage population as in natural populations. At least in this laboratory population of about 5,000 individuals, migration cannot be in any way responsible for maintenance of the variability in enzyme systems.

Selection

Overdominance has been considered a primary mechanism for the maintenance of polymorphism (e.g., Crow, 1952; Dobzhansky, 1954). Still, the very high levels of polymorphism suggested by the data in Table 6 have indicated that the segregational component of the genetic load would be intolerable. Since Lewontin and Hubby (1966) raised the apparent dilemma, several theoretical models have been proposed to help solve the problem (Sved, Reed, and Bodmer, 1967; King, 1967; Milkman, 1967). Nevertheless, inbreeding depression data have suggested that the number of balanced polymorphisms which can be maintained with a 1% heterozygote advantage is 1,000 (Sved et al., 1967). This number is smaller than our data have predicted. The solution may be that overdominance occurs only in a fraction of the polymorphic loci and the others are maintained by different factors.

But what are these other factors? Kojima and Yarbrough (1967) and Tobari and Kojima (1967) have suggested that frequency-dependent selection is responsible for maintaining widespread polymorphism. This type of selection is suggested by experiments with inversions (Tobari and Kojima, 1967), with mating preference (Ehrman et al., 1965; Ehrman, 1966), and with the *Est-6* locus in *D. melanogaster* (Kojima and Yarbrough, 1967; Yarbrough and Kojima, 1967). The difficulty with these analyses is that the effects of blocks of genes rather than single genes were measured (see MacIntyre and Wright, 1966). Moreover, the method of measuring fitness in the experiments on frequency-dependent selection might not be accurate since it seems to detect only a fraction of the total selective pressure (see Prout, 1965).

Sved et al. (1967) and King (1967) have demonstrated that fitness is not a function of the genotype in the absolute physiological sense; rather, it is modified by the environment, both ecological and genetic. This is certainly apparent when one compares the homologous loci of *D. melanogaster* and *D. simulans* in our sample. The genetic environment of the different species seems to have imposed a certain level of tolerability of polymorphism in each species. An ecological component as reflected by effective population size of the two species might also be critical.

It may be that linkage disequilibria are important in the maintenance of gene-enzyme polymorphisms (Sved, 1968). These could maintain several polymorphic loci simultaneously, even though only a fraction of the loci were overdominant. This may be the case in a linkage disequilibrium established in a population cage between alleles at the *Lap-D* and *Acph-1* loci (O'Brien, MacIntyre, and Fine, 1968). Much more research must be done before the role of gene interaction in the maintenance of polymorphisms can be evaluated.

SUMMARY

Nine populations of *D. melanogaster* and two populations of *D. simulans* were analyzed for polymorphism in 10 gene-enzyme systems by the technique of gel electrophoresis. In the eight natural populations of *D. melanogaster,* an average of 54% of the enzymes were polymorphic, and the average heterozygosity was 22.7%. An experimental population of *D. melanogaster,* which has been maintained in a laboratory cage for 20 years, showed levels of polymorphism equivalent to those of natural populations. The *D. simulans* populations had much less variability. The possible factors involved in maintaining these polymorphisms are discussed.

also indebted to Dr. Richard Lewontin for several helpful suggestions in the analysis of the results.

LITERATURE CITED

Beckman, L., and F. M. Johnson. 1964a. Genetic control of aminopeptidases in *Drosophila melanogaster*. Hereditas 51:221–230.

———. 1964b. Variation in larval alkaline phosphatase controlled by *Aph* alleles in *Drosophila melanogaster*. Genetics 49:829–835.

Brewer, John. 1967. Artifact produced in disc electrophoresis by ammonium persulfate. Science 156:256–257.

Brues, Alice M. 1954. Selection and polymorphism in the A-B-O blood groups. Amer. J. Phys. Anthropol. 12:559–597.

Courtright, J. B. 1967. Polygenic control of aldehyde oxidase in *Drosophila*. Genetics 57:25–39.

Crow, J. F. 1952. Dominance and overdominance. *In* J. W. Gowen [ed.] Heterosis. Iowa State Press, Ames, Iowa.

Dobzhansky, Th. 1954. Evolution as a creative process. 9th Int. Congr. Genet., Proc. Caryologia 6 (Suppl.)435–449.

Ehrman, L. 1966. Mating success and genotype frequency in *Drosophila*. Anim. Behav. 14:332–339.

Ehrman, L., B. Spassky, O. Pavlovsky, and Th. Dobzhansky. 1965. Sexual selection, geotaxis and chromosomal polymorphism in experimental populations of *Drosophila pseudoobscura*. Evolution 19:337–346.

Falke E., and R. J. MacIntyre. 1966. The genetic localization of non-specific leucine aminopeptidase in *Drosophila melanogaster*. Drosophila Information Service 41:165–166.

Grell, E. H. 1967. Electrophoretic variants of α-glycerophosphate dehydrogenase in *Drosophila melanogaster*. Science 158:1319–1320.

Grell, E. H., K. B. Jacobsen, and J. B. Murphy. 1965. Alcohol dehydrogenase in *Drosophila melanogaster*. Science 149:80–82.

Harris, H. 1966. Enzyme polymorphism in man. Roy. Soc. (London), Proc., B. 164:298–310.

Hubby, J. L., 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.

Isemura, Toshizo, and K. Kakuichi. 1962. Association and dissociation of bacterial α-amylase molecule. J. Bochem. 51:385–392.

Jacobson, K. B. 1967. Alcohol dehydrogenase in Drosophila: interconversion of isoenzymes. Science 159:324–325.

Johnson, F. M., and C. Denniston. 1964. Genetic variation of alcohol dehydrogenase in *Drosophila melanogaster*. Nature 204:906–907.

Johnson, F. M., C. G. Kanapi, R. H. Richardson, M. R. Wheeler, and W. S. Stone. 1966. An analysis of polymorphism among isozyme loci in dark and light *Drosophila ananassae* strains from American and Western Samao. Nat. Acad. Sci., Proc. 56:119–125.

Kimura, Motoo. 1968. Evolutionary rate at the molecular level. Nature 217:624–626.

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

King, J. L. 1967. Continuously distributed factors affecting fitness. Genetics 55:483–492.

Kitto, G. B., P. M. Wassarman, and N. O. Kaplan. 1966. Enzymatically active con-

formers of mitochondrial malate dehydrogenase. Nat. Acad. Sci., Proc. 56:578–585.

Kojima, K., and K. M. Yarbrough. 1967. Frequency dependent selection at the *Esterase-6* locus in *Drosophila melanogaster* Nat. Acad. Sci., Proc. 57:645–649.

Lewontin, R. C. 1967. An estimate of average heterozygosity in man. Amer. J. Hum. Genet. 19:681–685.

Lewontin, R. C., and J. L. Hubby. 1966. A molecular approach to the study of genic heterozygosity in natural populations. II. Amount of variation and degree of heterozygosity in natural populations of *Drosophila pseudoobscura*. Genetics 54:595–609.

MacIntyre, R. J. 1966*a*. The genetics of an acid phosphatase in *Drosophila melanogaster* and *Drosophila simulans*. Genetics 53:461–474.

———. 1966*b*. Locus of the structural gene for third larval instar alkaline phosphatase (*Aph.*) Drosophila Information Service 41:62.

MacIntyre, R. J., and T. R. F. Wright. 1966. Responses of *esterase-6* alleles of *Drosophila melanogaster* and *D. simulans* to selection in experimental populations. Genetics 53:371–387.

Milkman, R. 1967. Heterosis as a major cause of heterozygosity in nature. Genetics 55:493–495.

O'Brien, C. J. 1968. Electrophoretic variants of soluble malic dehydrogenose in Drosophila melanogaster. Isozyme Bull. 1:40 (Abstr.)

———. 1969. Genetics of *malic dehydrogenose-1 in Drosophila melanogaster*. Drosophila Information Service 44(Abstr.) (in press.).

O'Brien, S. J., and R. J. MacIntyre. 1968. α-GPDH-1^A: α-glycerophosphate dehydrogenase-1^A and α-GPDH-1^B in new mutants. Drosophila Information Service 43:60 (Abstr.)

O'Brien, S. J., R. J. MacIntyre, and W. Fine. 1968. A linkage disequilibrium between two gene enzyme systems in an experimental population of *Drosophila melanogaster*. Genetics 60:208–209. (Abstr.)

Prout, T. 1965. The estimation of fitnesses from gene frequencies. Evolution 19:546–551.

Shaw, C. R., and A. L. Koen. 1968. Starch gel electrophoresis of enzymes. *In* Ivor Smith [ed.] Chromatographic and electrophoretic techniques. 2d ed. Interscience New York.

Smithies, O., G. E. Connell, and G. H. Dixon. 1962. Chromosomal rearrangement and the evolution of haptoglobin genes. Nature 196:232–236.

Stone, W. S., M. W. Wheeler, F. M. Johnson, and K. Kojima. 1968. Genetic variation in natural island populations of *Drosophila nasuta* and *Drosophila ananassae* subgroups. Nat. Acad. Sci., Proc. 59:102–109.

Sved, J. 1968. The stability of linked systems of loci with a small population size. Genetics 59:543–563.

Sved, J. A., T. E. Reed, and W. F. Bodmer. 1967. The number of balanced polymorphisms that can be maintained in a natural population. Genetics 55:469–481.

Tobari, Y. N., and K. Kojima. 1967. Selection modes associated with inversion karyotypes in *Drosophila ananassae*. I. Frequency-dependent selection. Genetics 57:179–188.

Ursprung, H. 1966. Multiple molecular forms of alcohol dehydrogenase in *Drosophila melanogaster*. N.Y. Acad. Sci. Ann. (in press).

Wright, S. 1966. Polyallelic random drift in relation to evolution. Nat. Acad. Sci., Proc. 55:1074–1081.

Wright, T. R. F. 1963. The genetics of an esterase in *Drosophila melanogaster*. Genetics 48:787–801.

Wright, T. R. F., and R. J. MacIntyre. 1963. A homologous-gene enzyme system, *Esterase-6*, in *Drosophila melanogaster* and *Drosophila simulans*. Genetics 48:1717–1728.

Yarbrough, K., and K. Kojima. 1967. The mode of selection at the polymorphic *Esterase-6* locus in cage populations of *Drosophila melanogaster*. Genetics 57:677–686.

Yen, T. T. T., and E. Glassman. 1965. Electrophoretic variants of xanthine dehydrogenase in *Drosophila melanogaster*. Genetics 52:977–981.

Young, W. J. 1966. X-linked electrophoretic variation in 6-phosphogluconate dehydrogenase. J. Hered. 57:58–60.

Young, W. J., I. Porter, and B. Childs. 1964. Glucose-6-phosphate dehydrogenase in *Drosophila:* X-linked electrophoretic variants. Science 143:140.