*AN ANALYSIS OF POLYMORPHISMS AMONG ISOZYME LOCI IN DARK AND LIGHT DROSOPHILA ANANASSAE STRAINS FROM AMERICAN AND WESTERN SAMOA**

By F. M. Johnson, Carmen G. Kanapi, R. H. Richardson, M. R. Wheeler, and W. S. Stone

GENETICS FOUNDATION, DEPARTMENT OF ZOOLOGY, UNIVERSITY OF TEXAS, AUSTIN

Communicated May 27, 1966

There is mounting evidence that *Drosophila* populations are more polymorphic at the level of genetic loci than was formerly thought by many geneticists. There have been many experimental demonstrations of the extensive chromosomal polymorphisms, of the presence of considerable quantities of hidden deleterious and lethal alleles in the populations, and of the extensive complexity of the coadapted genetic systems integrated through natural selection.[1] Nevertheless, the proportion of polymorphisms among the majority of all loci is largely unknown because of inadequate screening techniques. Furthermore, the mechanisms for the maintenance of polymorphisms remain elusive to experimental verification.

The ideal situation of screening populations for genetic comparisons at and between individual loci has been brought a step closer to reality by electrophoretic separation of enzyme molecules. Though enzyme differences are not necessarily related with genetic differences on a one-to-one basis, most may be. At any rate, zymogram analysis of individual flies for electrophoretic mobility or qualitative activity differences gives a good first approximation of the underlying genetic situation for a different class of variation than has previously been studied directly. However, as Lewontin and Hubby[3] have pointed out, the randomness of the samples of loci studied by this technique is critical for broader generalizations which might be made from the findings in this class of loci. Sample characteristics are yet to be determined.

Lewontin and Hubby[3] have found about 40 per cent of all electrophoretically separable enzyme systems and protein bands of no known enzymatic activity of *D. pseudoobscura* to be polymorphic, having two or more electrophoretic forms of the same protein. One expects the difference in molecular forms to represent at least as much genetic difference, so that the variation of molecular types reflects similar variation of genotypes. If only a moderate bias is involved in estimating this proportion for the genome, then there exists a major discrepancy between the data and the expectation based upon the theory of segregational load. According to the theory, a population may bear, without being jeopardized by reduced average fitness, only a relatively small portion of loci in heterotic polymorphic conditions. In order to establish the meaning of the apparent inconsistency between the theory and observations, it must be established whether or not these isozyme polymorphisms are maintained more or less independently by some form of selection. It is very important to know if these variations are random representatives of a large number of such loci. It has been demonstrated in a laboratory population of *D. busckii*[4] that the polymorphism at the "leucine aminopeptidase" locus is maintained by heterotic selection, though this was not the case for an esterase locus in *D. melanogaster* and *D. simulans*.[5]

Previous works[6-8] indicate that a promising series of populations for analysis on the basis of genetic differences are the "light" and "dark" *D. ananassae* forms (species) of the South Pacific, particularly in American and Western Samoa. In this complex there are two sibling forms that phenotypically resemble *D. ananassae*, a cosmopolitan species. However, Futch[7] found chromosome inversions in one form which were totally lacking in the other, though the forms were sympatric. Tests by Stone *et al.*[6] indicated that the forms would hybridize in the laboratory if given no mating choice, producing fertile and vigorous F_2's and backcrosses. Spieth[8] found a strong positive assortive mating preference in the laboratory when males and females of both forms were together, which, together with the findings of Futch[7] and Stone *et al.*,[6] indicate essentially complete sexual isolation in nature. Thus, the two Samoan forms are two species of *D. ananassae*-like flies that are extremely closely related, living in overlapping habitats, and highly amenable to laboratory genetic tests.

In this communication, the number of polymorphic protein systems and the number of nonvariable protein forms are compared. Also a gene frequency comparison is made on the "Esterase-C" system for the natural populations of the Samoan Islands (this system is presumed but not yet shown to be homologous to the Esterase-C system of *D. melanogaster*).

Materials and Methods.—The materials used in the present series of experiments were collections made by Profs. W. S. Stone and C. P. Oliver in the Samoan Islands in July and August 1965. Several locations on the islands of Tutuila, Upolu, and Savaii were sampled by sweeping with a net over fermenting papaya and other fruit. Both light and dark forms were obtained from the same collection site on Tutuila and Upolu, but on Savaii only light *ananassae* were collected. The shipment from Tutuila, made in July 1965, was received in Austin with only a few surviving flies, and the location was resampled in August. As the flies were received in Austin, females were placed individually in corn-meal-malt-agar food vials, and cultures of isofemale lines were initiated.

Beginning in mid-October, after no more than four generations of laboratory culture, electrophoretic analyses were initiated on the isofemale line cultures, with the primary intention of qualitatively examining the extent of isozyme polymorphisms in these populations. In the case of Esterase-C, all lines available were examined for allelic content. Since isofemale lines from wild-caught flies were available for analysis, gene frequency estimates in the Esterase-C system could be made for the Samoan populations. The estimates are of only moderate precision, since the samples of alleles from the natural populations are small. The number of isofemale lines in the original samples are 29 from Tutuila and 15 from Upolu for the dark forms, and 12 from Tutuila, 14 from Upolu, and 14 from Savaii for the light stocks. Usually about 12 flies were examined from each isofemale line (set of four alleles of the original sample), but as many as 30 flies per line were examined for the esterases. The average number examined was lower for the other enzymes, since some lines were not tested.

Most of the starch gel electrophoretic and isozyme assay methods were performed according to previously described procedures. These included alkaline phosphatase (APH),[9] esterase (EST),[10] leucine aminopeptidase (LAP),[11] and alcohol dehydrogenase (ADH).[12] The α-β specificity difference of esterases was determined with naphthyl acetates as described by Johnson *et al.*[2] In Tables 1, 2, and 3 of the present report, the α,β Esterase-A designations refer to α-naphthyl acetate and β-naphthyl acetate specificity, respectively. Acid phosphatase (ACPH) detection was similar to the APH method, except that a sodium acetate buffer, pH 5, was used for staining. Xanthine dehydrogenase (XDH) was assayed according to the ADH method, except that hypoxanthine was substituted for alcohol as substrate. One protein which showed no catalytic activity with the above methods was detected with a general protein stain, naphthol blue black. Single fly homogenization was done using the procedure outlined by Johnson,[13] and the electrophoretic separation was performed in the discontinuous system of buffers as described by Poulik.[14] The

isozymes were generally given letter designations reflecting their relative migration rate to the anode, with *A* indicating the faster migration rate.

Results and Discussion.—Polymorphic loci: Each gel stained for a particular class of enzymes had several regions of banding representing different enzyme systems. Where the regions for two systems overlapped, some other criteria, such as sensitivity to alcohol, were sometimes used to distinguish the different enzymes. In cases where the particular zone had only a constant pattern in all lines of a strain, the zone and the corresponding locus were considered to be monomorphic. However, if the bands from individuals varied, either as differences in electrophoretic mobility or as presence-absence (active-inactive) differences, they were considered to be polymorphic. An exception was an apparently sex-specific difference in alkaline phosphatase in adults.[15] In those cases where the variation was a difference in the intensity of stain, it was not considered as a polymorphism. Intensity variation could involve either variation in factors acting above the level of primary structure of the enzyme molecule or inadvertent concentration differences.

The procedure of relating banding patterns to genetic loci in some cases rests on the assumptions that proteins have their primary structure determined by one or a few genetic loci, and that changes in the primary structure of the protein will sometimes result in net charge changes of the molecule. Some changes in the genetic information at a locus may not be reflected in electrophoretic or gross activity variations, and multiple changes in a protein may be indistinguishable from single changes. Consequently, from the considerations of genic control of molecular structure affecting net molecular charge, the recognition of polymorphism by the method used is likely to give conservative estimates of the number of alleles in a population.

The mechanisms supporting the polymorphisms remain to be determined, but some possibilities may be suggested. These might include genotypic variation which is essentially neutral with respect to adaptation under present conditions, heterotic effects on a per-locus basis, functional association with a heterotic linkage group such as polymorphic inversion system, or chance association with a heterotic block of loci but with functional neutrality of the enzyme variation.

Wright[16] has pointed out the large number of essentially neutral alleles differing by only one or a few nucleotide pairs that may be maintained primarily at low frequencies in a population by mutation pressure. It may be that some of these neutral alleles are among the controlling elements which produced the electrophoretic variants detected here. However, in most cases the enzyme alleles are present in high frequencies and appear repeatedly in small samples. Some of the intermediate frequencies may represent an intermediate stage of gene substitution discussed by Haldane[17] and may contribute temporarily to the total genetic load.[18] However, in at least the few cases we have examined,[2, 4] the intermediate frequencies appear to be stable. Furthermore, as discussed below, the frequencies for various populations within the same species are similar, yet the frequencies are characteristically different between the species, which suggests a heterotic mechanism of maintenance of intermediate gene frequencies.

If we assume that the enzyme variants usually represent loci with adaptively heterotic allelic combinations, then it must be shown that the sample of genetic loci is a reasonably unbiased representative of the genome before the relevance to

TABLE 1

RELATIVE STAINING INTENSITY FOR DEVELOPMENTAL STAGES

Enzyme	Category	Larvae	Pupae	Adult
EST	*A* (α)	+	++	+
	A (β)	+	++	+
	C	++	+	+++
	6	++	++	++
	E'	+	+	+
	E	+	++	++
APH	*A*	+	0	+
	B	Trace	0	+++
	*C**	0	0	++
	D	+++	+	++†
ACPH	*A*	+++	+++	+++
LAP	*A*	Trace	Trace	Trace
	B	0	0	Trace
	C	+	0	++
	D	0	+++	+
	E	+	0	+
	F	+	0	+
ADH	*A*	0	0	+
	B	++	++	++
XDH	*A*	Trace	Trace	+
Protein	*A*	+	+	Trace

* Appears in older adults.
† Adult females show no, or a very lightly staining, APH D band.

genetic load theory may be stated. Sources of bias are largely obscure at the moment; therefore, conclusions are drawn with caution. For instance, it is fairly certain that the enzymes presently assayed are those in high concentration in the body and lymph, and it remains for further refinements in sensitivity of the assay to check the variability of less concentrated enzymes.

The enzyme patterns were examined in larval, pupal, and adult stages of the life cycle. Some systems were detected only in certain stages of development. Table 1 lists the band scored and their relative intensities of staining in the three major stages of the life cycle (third instar larvae, mid-pupae and adults). There are a number of experimental and biological factors which control the intensity of staining, but the scoring of intensity can be made with little ambiguity. Furthermore, zymogram assay of hybrids between light and dark *ananassae* indicate genetic homology of the Esterase C and acid phosphatase A system in the two species. The two species apparently contain the same alleles, as determined by these procedures.

In Table 2 the composite variations found in the different zones are indicated for the stages of the life cycle where the variations were observed. A value of 2 or more in the "mobility" column or "yes" in the "activity missing" column indicates a presumed genetic variation. Obviously the genetic analysis for each system must be done before the full genetic significance of the variants is available. These analyses are most completed for Esterase C and acid phosphatase. Multiple allelic inheritance is indicated (unpublished data).

With some strains the number of isofemale lines was small, so that one would expect to find only part of the variations in the population. Nevertheless, if the alleles were maintained at intermediate frequencies in wild populations, the chances are good (> 0.90) that the presence of the polymorphism would be detected with 20 or more flies taken from each of several lines of a strain. Working on this basis,

TABLE 2

SUMMARY OF HERITABLE ENZYME VARIATIONS

Enzyme	Category	No. of mobility positions	Enzyme activity missing	Hybrid band in heterozygote
EST	*A* (α)	2	Yes	No
	A (β)	2	Yes	No
	C	5	No	No
	6	1	No	—
	E'	2	Yes	No
	E	1	No	—
APH	*A*	1	No	—
	B	3	No	—*
	C	1	No	—
	D	4	No	Yes
ACPH	*A*	3	No	Yes
LAP	*A*	1	—	—
	B	1	—	—
	C	1	—	—
	D	3	No	No
	E	1	No	—
	F	1	No	—
ADH	*A*	2	Yes	No
	B	1	No	—
XDH	*A*	1	Yes	—

* Insufficient electrophoretic resolution.

the presence or absence of a polymorphism in a strain is evaluated with reasonable accuracy, and this level of resolution will indicate something of the genetic heterogeneity of some natural populations.

Table 3 summarizes the occurrence of polymorphisms detected in the various *ananassae* forms from the Samoan Islands. In general, the polymorphic bands found in one strain are usually polymorphic in all other strains, and the proportion of polymorphic to monomorphic zones agrees with the findings of Lewontin and Hubby;[3] that is, somewhere around 40–50 per cent of all loci examined are polymorphic. The implications of this massive heterogeneity remain to be determined.

TABLE 3

DISTRIBUTION OF POLYMORPHISM IN LIGHT AND DARK STRAINS OF *D. ananassae*

		Light			Dark	
Enzyme	Category	Tutuila	Upolu	Savaii	Tutuila	Upolu
EST	*A* (α)	+	+	+	+	+
	C	+	+	+	+	+
	6	0	0	0	0	0
	E'	+	+	+	+	+
	E	0	0	0	0	0
APH	*A**	0	0	0	0	0
	B	0	+	0	+	+
	D	+	+	+	+	+
LAP	*C*	0	0	0	0	0
	D	0	+	+	+	0
ADH	*A*	+	+	+	+	0
	B	0	0	0	0	0
XDH	*A*	+	+	0	0	0
Protein	*A*	0	0	0	0	0

* Intensity of stain not considered.

TABLE 4

GENE FREQUENCY ESTIMATES FOR *D. ananassae*-LIKE SAMPLES FROM SAMOA: *Est C*

Species	Location	Allelic frequency F	M	S
Light	Tutuila	0.12	0.73	0.15
	Savaii	0.19	0.77	0.04
	Upolu	0.11	0.89	0.00
	Average	0.14	0.80	0.06
Dark	Tutuila	0.84	0.14	0.02
	Upolu	0.67	0.33	0.00
	Average	0.76	0.23	0.01

Most protein bands (without known catalytic activity) are polymorphic in *D. pseudoobscura*.[3] Consequently, further analysis of *ananassae* may reveal such proteins to be polymorphic in *ananassae* also. From a number of species comparisons, which in several cases involved different laboratory strains of the same species, the Esterase E is the only enzyme zone that appears to be universally monotypic, even between species of different subgenera.[2] Consequently, this enzyme may be in a special category in which constancy of molecular form is especially important for life in *Drosophila*.

Population patterns: The inheritance of Esterase C has been shown (unpublished results) to be a multiple allelic system (at least three alleles). There may be slight differences in adaptive values among the genotypes involving the three alleles (F, M, S) isolated in laboratory stocks. Two rare presumptive alleles have not yet been studied.

To gain information of adaptive importance in nature, it would be of interest to evaluate the gene frequencies of the two *ananassae* species in the Samoan Islands. Table 4 gives a summary of the estimates of frequencies of the three major alleles in the populations. Differences between allelic frequencies of the same species on different islands are not significant. On the other hand, the differences between the frequencies of the "fast" alleles and of the "medium" alleles for the two forms (species) are highly significant. Apparently, the two species have attained different polymorphic balances under natural conditions.

The obvious implication of this difference is that either the environments are different such that isozyme homozygotes have different selective disadvantages, or that there are nonallelic interactions between the isozyme alleles and other genes of the system. These possibilities are now being investigated.

Summary.—An analysis of electrophoretic variations in several enzymes and one protein (without known catalytic activity) of Samoan strains of "light" and "dark" *Drosophila ananassae* indicated about 40–50 per cent of the loci variable for two or more alleles. In the case of one esterase, *Est C*, gene frequencies from the populations on Tutuila, Upolu, and Savaii were estimated. The data indicated a strong similarity of "light" *ananassae* from different islands with allelic frequency estimates averaging 0.14 (F), 0.80 (M), and 0.06 (S), while a different pattern of similarity between islands was obtained for "dark" *ananassae* with frequency estimates averaging 0.76 (F), 0.23 (M), and 0.01 (S). The S-allele was not found in either "light" or"dark" forms from one island.

* This work is, in part, research by Frank M. Johnson to be included in a dissertation to be submitted to the University of Texas in partial fulfillment of the Ph.D. requirements in Zoology. Supported (in part) by a U.S. Public Health Service research grant (GM 11609 to W. S. Stone and

M. R. Wheeler) and training grant (2 T1-GM 337-06 to R. P. Wagner *et al.*) from the National Institutes of Health and by a contract from the U.S. Atomic Energy Commission [AT-(40-1)-2952 to W. S. Stone].

[1] For a review, see the introduction by E. B. Spiess in *Papers on Animal Population Genetics* (Boston: Little, Brown and Co., 1962).

[2] Johnson, F. M., C. G. Kanapi, R. H. Richardson, M. R. Wheeler, and W. S. Stone, "An operational classification of *Drosophila* esterases for species comparison," *Univ. of Texas Publication*, in press.

[3] Lewontin, R. C., and J. L. Hubby, "A molecular approach to the study of genic heterozygosity in natural populations, II. Analysis of geographical populations of *Drosophila pseudoobscura*," *Genetics*, in press.

[4] Richardson, R. H., and F. M. Johnson, "A tri-allelic isozyme system in *Drosophila busckii*, II. A stable polymorphic system," manuscript in preparation.

[5] MacIntyre, R. J., and T. R. F. Wright, "Response of esterase 6 alleles of *Drosophila melanogaster* and *D. simulans* to selection in experimental populations," *Genetics*, **53**, 371 (1966).

[6] Stone, W. S., M. R. Wheeler, F. D. Wilson, V. L. Gerstenberg, and H. Yang, "Genetic studies of natural populations of *Drosophila:* I. Pacific island populations," *Univ. of Texas Publication*, in press.

[7] Futch, D. G., "A study of speciation in South Pacific populations of *D. ananassae*," *Univ. of Texas Publication*, in press.

[8] Spieth, H. T., "Mating behavior of *D. ananassae* and *ananassae*-like flies from the Pacific," *Univ. of Texas Publication*, in press.

[9] Beckman, L., and F. M. Johnson, "Variations in larval alkaline phosphatase controlled by *Aph* alleles in *Drosophila melanogaster*," *Genetics*, **49**, 829 (1964).

[10] Beckman, L., and F. M. Johnson, "Esterase variation in *Drosophila melanogaster*," *Hereditas*, **51**, 212 (1964).

[11] Beckman, L., and F. M. Johnson, "Genetic control of aminopeptidases in *Drosophila melanogaster*," *Hereditas*, **51**, 221 (1964).

[12] Johnson, F. M., and C. Denniston, "Genetic variation of alcohol dehydrogenase in *Drosophila melanogaster*," *Nature*, **204**, 906 (1964).

[13] Johnson, F. M., "Rapid single fly homogenization for the investigation of *Drosophila* isozymes," *Drosophila Information Service*, **41**, 193 (1966).

[14] Poulik, M. D., "Starch gel electrophoresis in a discontinuous system of buffers," *Nature*, **180**, 1477 (1957).

[15] Johnson, F. M., "A comparison of developmental differences of alkaline phosphatase zymograms from *Drosophila melanogaster* and *D. ananassae*," *Nature*, in press.

[16] Wright, S., "Polyallelic random drift in relation to evolution," these PROCEEDINGS, **55**, 1074 (1966).

[17] Haldane, J. B. S., "The cost of natural selection," *J. Genet.*, **55**, 511 (1957).

[18] Kimura, M., "Optimum mutation rate and degree of dominance as determined by the principle of minimum genetic load," *J. Genet.*, **57**, 21 (1960).