

UC Irvine

UC Irvine Previously Published Works

Title

Enzyme variability in the *Drosophila willistoni* group. VI. Levels of polymorphism and the physiological function of enzymes

Permalink

<https://escholarship.org/uc/item/3wb6w1hv>

Journal

Biochemical Genetics, 7(3-4)

ISSN

0006-2928

Authors

Ayala, Francisco J
Powell, Jeffrey R

Publication Date

1972-12-01

DOI

10.1007/bf00484832

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at

<https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Enzyme Variability in the *Drosophila willistoni* Group. VI. Levels of Polymorphism and the Physiological Function of Enzymes

Francisco J. Ayala¹ and Jeffrey R. Powell¹

Received 8 Dec. 1971—Final 4 April 1972

We have studied allelic variation at 28 loci coding for soluble enzymes in three sibling species of Drosophila. The average proportion of polymorphic loci per population is 58% in D. willistoni, 71% in D. equinoxialis, and 60% in D. tropicalis. An individual of any one of these three species is heterozygous at 18-22% of its genes. The correlation between the amount of polymorphism at a given locus and the equilibrium constant of the reaction catalyzed by the corresponding enzyme is not significantly different from zero. We have separated enzymes into those involved in glucose metabolism (group I) and all other enzymes (group II). The genes coding for group II enzymes are about twice as polymorphic as those coding for group I enzymes.

INTRODUCTION

Recent studies have confirmed that natural populations of animals and plants have large amounts of genetic variation (Harris, 1970; Selander *et al.*, 1970; Marshall and Allard, 1970; Kojima *et al.*, 1970; Ayala *et al.*, 1972; and references therein). These studies are for the most part concerned with the genetic variation underlying variable enzyme patterns exhibited by gel electrophoresis. It has now become possible, at least to a first approximation, to quantify the amount of genetic variation in a population and to measure the amount of genetic differentiation between different populations and between different species.

This investigation was supported by NSF grants GB 20694 and GB 30895.

¹ Department of Genetics, University of California, Davis, California.

Table I. Genetic Variation in *Drosophila willistoni*

Gene	No. wild genomes sampled	No. populations sampled	Proportion of polymorphic populations	Mean frequency of heterozygous individuals	Frequency of most common allele	No. alleles with frequency ≥ 0.01
Group I (glucose-metabolizing enzymes)						
<i>αGpdh</i>	7032	57	0.044	0.021 ± 0.005	0.988	1
<i>Ald</i>	404	41	0.571	0.135 ± 0.048	0.866	4
<i>G3pdh</i>	190	27	0.500	0.148 ± 0.053	0.911	3
<i>Hk-1</i>	1620	44	0.333	0.099 ± 0.016	0.931	3
<i>Hk-2</i>	2228	49	0.556	0.138 ± 0.032	0.940	3
<i>Hk-3</i>	2060	47	0.071	0.039 ± 0.009	0.977	2
<i>Idh</i>	3168	63	0.370	0.086 ± 0.023	0.938	3
<i>Mdh-2</i>	6680	58	0.175	0.040 ± 0.008	0.979	2
<i>Me-1</i>	2882	61	0.275	0.079 ± 0.013	0.959	2
<i>Me-2</i>	1368	15	0.733	0.235 ± 0.044	0.849	3
<i>Pgm-1</i>	2636	28	0.815	0.186 ± 0.025	0.874	3
<i>Tpi-2</i>	2478	59	0.105	0.041 ± 0.010	0.980	2
Group II (other enzymes)						
<i>AcpH-1</i>	3066	24	0.588	0.102 ± 0.018	0.956	3
<i>AcpH-2</i>	874	10	0.900	0.151 ± 0.023	0.915	3
<i>Adh</i>	5916	56	0.515	0.103 ± 0.015	0.954	2
<i>Adk-1</i>	2150	40	1.000	0.527 ± 0.019	0.577	3
<i>Adk-2</i>	2580	60	0.722	0.171 ± 0.023	0.928	3
<i>Aph-1</i>	6272	65	0.941	0.225 ± 0.017	0.866	4
<i>Est-2</i>	7012	61	0.668	0.117 ± 0.008	0.943	3
<i>Est-3</i>	3914	42	0.727	0.108 ± 0.009	0.946	3
<i>Est-4</i>	9692	78	0.885	0.270 ± 0.012	0.847	3
<i>Est-5</i>	10432	81	0.333	0.089 ± 0.006	0.954	3
<i>Est-6</i>	2418	48	1.000	0.285 ± 0.030	0.783	3
<i>Est-7</i>	6819	67	1.000	0.601 ± 0.009	0.566	5
<i>Lap-5</i>	10348	80	1.000	0.629 ± 0.007	0.487	5
<i>Odh-1</i>	1088	30	0.765	0.180 ± 0.019	0.878	3
<i>Odh-2</i>	40	4	0.333	0.060 ± 0.024	0.975	2
<i>To</i>	4949	22	0.333	0.131 ± 0.029	0.905	2
Mean	3940	47	0.581	0.178 ± 0.030	0.881	2.89

Only a relatively limited number of species and only a few dozen loci in any one of them have been studied so far. Yet certain patterns in the amount and kind of variation have started to emerge (Ayala, 1972). One pattern is that certain enzymes and the genes which code for them are generally more variable than others. Are there any general physiological reasons why some loci, in a given group of organisms, are more variable than others? Some authors have suggested that such reasons exist.

We have conducted an extensive survey of enzyme variation in several neotropical species of *Drosophila*. We summarize here our data about the genetic variation found in three sibling species, *Drosophila willistoni*, *D. equinoxialis*, and *D. tropicalis*. From 25 to 28 loci have been studied in each of these species. From several hundred to several thousand wild genomes have been studied at each locus. We ask the question whether the amount of genetic variation depends on the chemical properties and physiological role of the enzymes, as has been recently suggested (Gillespie and Kojima, 1968 *et al.*, 1970; Johnson, 1971).



Fig. 1. Localities sampled. ●, *Drosophila willistoni*; ▲, *Drosophila equinoxialis*; ▼, *Drosophila tropicalis*.

MATERIALS AND METHODS

The geographic distribution of *D. willistoni*, *D. equinoxialis*, and *D. tropicalis* extends through Central America, the Caribbean, and tropical and subtropical South America (Spassky *et al.*, 1971). We have obtained samples from more than 40 natural populations of each species. The localities sampled are spread throughout most of the distribution of the species (Fig. 1). From a few localities, only one or a few laboratory strains were available for study. From most localities, we have secured freshly collected samples and studied either the flies collected in nature or one F_1 progeny from each wild female.

Our techniques for starch gel electrophoresis and assay of enzymes have been described elsewhere (Ayala *et al.*, 1972). They are standard techniques with minor modifications to suit our materials.

The enzymes studied, with the symbols used to represent the genes coding for them are as follows: Glucose-metabolizing enzymes: α -glycerophosphate dehydrogenase, *α Gpdh* (1.1.99.5); aldolase, *Ald* (4.1.2.13); glyceraldehyde 3-phosphate dehydrogenase, *G3pdh* (1.1.1.8); hexokinase, three loci, *Hk-1*, *Hk-2*, *Hk-3* (2.7.1.1); isocitrate dehydrogenase, *Idh* (1.1.1.41); malate dehydrogenase, *Mdh-2* (1.1.1.37); malic enzyme, two loci, *Me-1* and *Me-2* (1.1.1.40); phosphoglucomutase, *Pgm-1* (2.7.5.1); triosephosphate isomerase, *Tpi-2* (5.3.1.1). Other enzymes: acid phosphatase, two loci, *Acph-1* and *Acph-2* (3.1.3.2); alcohol dehydrogenase, *Adh* (1.1.1.1); adenylate kinase, two loci, *Adk-1* and *Adk-2* (2.7.4.3); alkaline phosphatase, *Aph-1* (3.1.3.1); esterase, six loci, *Est-2*, *Est-3*, *Est-4*, *Est-5*, *Est-6*, and *Est-7* (3.1.1.2); leucine amino peptidase, *Lap-5* (3.4.1.1); octanol dehydrogenase, two loci, *Odh-1* and *Odh-2* (1.1.1.1); tetrazolium oxidase, *To*.

RESULTS

The geographic distribution of *D. willistoni* extends from Mexico and southern Florida through Central America, the Caribbean, and most of South America down to La Plata in Argentina. We have studied 81 samples from natural populations spread throughout that immense territory (Fig. 1). The results are summarized in Table I. For each locus, the table shows the number of genomes sampled, two for each wild individual, except for sex-linked loci, which are carried by males in single dose; the number of populations sampled; the proportion of polymorphic populations, that is, populations in which the most common allele has a frequency no greater than 0.95; the mean frequency of heterozygous individuals estimated from the allelic frequencies assuming Hardy-Weinberg equilibrium; the actual frequency of the most common allele in the total sample for the species; and the number of alleles with a frequency not smaller than 0.01 over the whole species.

The mean frequency of heterozygous individuals is the unweighted average of their mean frequency in each sample. Only samples with 20 or more genomes have been used to estimate the average, except at the *Ald*, *G3pdh*, and *Odh-2* loci, where all samples with ten or more genomes have been used. We have used the *expected* frequencies of heterozygotes, assuming Hardy-Weinberg equilibrium. The agreement between the expected and the observed frequency of heterozygotes is generally good.

To estimate the frequency of populations in which a given locus is polymorphic, we have used also only those samples with at least 20 genomes, except for the three loci mentioned in the previous paragraph.

We have studied in *D. willistoni* 28 loci coding for a variety of enzymes. On the average, we have sampled 3940 genomes from 47 different populations per locus. There is a great wealth of genetic variation. More than half of the loci, 58.1%, are polymorphic in a given population. Using a less stringent criterion of polymorphism, namely, that a locus is polymorphic when the second most common allele has a frequency of 0.01 or higher, the proportion of polymorphic loci per population is 86%. The mean overall loci of the frequency of heterozygous individuals per locus, estimates the proportion of loci at which an individual is heterozygous. On the average, a *D. willistoni* fly is heterozygous at $17.8 \pm 3.0\%$ of its genes.

The distribution of *D. equinoxialis* extends from southern Mexico through Central America, the Caribbean, and northern South America to central Ecuador and central Brazil. Unpublished data obtained in our laboratory indicate that the species consists of two main subspecies, one in most of Central America and the Caribbean, the other in South America. Crosses between the two subspecies give sterile male hybrids. Table II summarizes the genetic variation found in populations of the South American subspecies. A total of 41 samples of natural populations have been studied (Fig. 1). The proportions of polymorphic populations and of heterozygous individuals are calculated using only samples of at least 20 genomes, except for the *Ald* and *AcpH-2* loci, where all samples of ten or more genomes were taken into account.

We have studied 27 loci in *D. equinoxialis*. On the average, 1039 genomes and 18 populations have been sampled at each locus. On the average, 71.1% loci are polymorphic in a given population, or 88% if the less stringent criterion of polymorphism is used. An individual is heterozygous at $21.8 \pm 3.0\%$ of its loci. This figure is not significantly different from that obtained for *D. willistoni*. As in *D. willistoni*, populations of *D. equinoxialis* contain large amounts of genetic variation.

The geographic distribution of *D. tropicalis* is coextensive with that of *D. equinoxialis* throughout Mexico, Central America, and the Caribbean. In South America, *D. tropicalis* extends farther to the south of Brazil than *D.*

Table II. Genetic variation in *Drosophila equinoxialis*

Gene	No. wild genomes sampled	No. populations sampled	Proportion of polymorphic populations	Mean frequency of heterozygous individuals	Frequency of most common allele	No. alleles with frequency ≥ 0.01
Group I (glucose-metabolizing enzymes)						
<i>αGpdh</i>	1888	24	0.000	0.027 ± 0.009	0.983	2
<i>Ald</i>	66	13	1.00	0.297 ± 0.007	0.833	4
<i>G3pdh</i>	28	3	1.00	0.196	0.906	3
<i>Hk-1</i>	428	14	0.667	0.151 ± 0.028	0.914	2
<i>Hk-2</i>	662	15	0.571	0.146 ± 0.041	0.920	4
<i>Hk-3</i>	676	15	0.571	0.122 ± 0.038	0.954	2
<i>Idh</i>	822	23	0.444	0.088 ± 0.023	0.960	2
<i>Mdh-2</i>	1900	25	0.000	0.010 ± 0.004	0.994	1
<i>Me-1</i>	734	21	0.000	0.025 ± 0.010	0.990	1
<i>Me-2</i>	292	15	1.00	0.308 ± 0.058	0.801	3
<i>Pgm-1</i>	524	12	1.00	0.444 ± 0.034	0.622	3
<i>Tpi-2</i>	550	22	0.000	0.037 ± 0.013	0.978	2
Group III (other enzymes)						
<i>AcpH-1</i>	702	5	1.00	0.308 ± 0.010	0.811	3
<i>AcpH-2</i>	56	2	1.00	0.296 ± 0.018	0.821	4
<i>Adh</i>	1860	22	0.909	0.232 ± 0.032	0.861	2
<i>Adk-1</i>	658	22	1.00	0.474 ± 0.034	0.547	3
<i>Adk-2</i>	576	22	0.250	0.096 ± 0.052	0.941	3
<i>Aph-1</i>	1913	29	0.813	0.133 ± 0.023	0.919	3
<i>Est-2</i>	834	14	0.833	0.126 ± 0.019	0.939	4
<i>Est-3</i>	110	9	1.00	0.566 ± 0.029	0.473	3
<i>Est-4</i>	2682	32	1.00	0.373 ± 0.045	0.769	3
<i>Est-5</i>	2696	34	0.571	0.090 ± 0.011	0.940	3
<i>Est-6</i>	2756	32	1.00	0.256 ± 0.027	0.843	4
<i>Lap-5</i>	2774	41	1.00	0.486 ± 0.023	0.712	3
<i>Odh-1</i>	508	4	1.00	0.293 ± 0.073	0.827	3
<i>Odh-2</i>	24	2	1.00	0.165	0.917	2
<i>To</i>	1347	14	0.571	0.131 ± 0.046	0.956	2
Mean	1039	18	0.711	0.218 ± 0.030	0.857	2.74

equinoxialis. Table III summarizes our study of *D. tropicalis*. Forty-five samples of natural populations have been surveyed (Fig. 1). As before, the proportions of polymorphic populations and of heterozygous individuals are estimated from samples containing at least 20 genomes, except for *Ald* and *G3pdh*, where samples with ten or more genomes have been used.

An average of 1248 genomes and 20 populations have been sampled for each of 25 loci of *D. tropicalis*. The amount of genetic variation is approximately the same as in the other two species; 59.6% of the loci are polymorphic in a given population by the "0.95" criterion, or 88% by the "0.01" criterion. An individual is heterozygous at $19.2 \pm 3.1\%$ loci, which is not significantly different from the heterozygosity found either in *D. willistoni* or in *D. equinoxialis*.

Not only the total variation but also the amount of variation at any given locus is generally quite similar in the three sibling species. The correlation in the proportion of heterozygous individuals per locus is 0.65 between *D. willistoni* and *D. equinoxialis*, 0.84 between *willistoni* and *tropicalis*, 0.71 between *equinoxialis* and *tropicalis*. The three correlation coefficients are significantly greater than zero ($P < 0.001$). Similarly, positive correlations between species exist in the proportion of polymorphic populations per locus, in the frequency of the most common allele, and in the number of alleles.

What is the significance of the correlation between species in the amount of variation at any given locus? Gillespie and Kojima (1968) and Kojima *et al.* (1970) have suggested that enzymes involved in the metabolism of glucose are likely to be less variable than other enzymes. Their rationale is that energy metabolism is essential for the survival of the organism. The substrates of most enzymes involved in glucose metabolism are usually restricted to a single species of molecule; variation in the substrates of such enzymes is mainly quantitative. On the other hand, the majority of all other enzymes studied in *Drosophila* as well as in other animals have broad substrate specificities. These substrates often originate outside the organism and thus vary *qualitatively* as well as quantitatively. Variation in these enzymes may reflect the heterogeneity of the environment.

In Table IV, we compare the amount of genetic variation between glucose-metabolizing enzymes and all other enzymes. We compare the same three parameters of genetic variation used by Kojima *et al.* (1970): proportion of loci which are polymorphic in a given population, proportion of heterozygous loci per individual, and average number of alleles per locus. In all three measures, loci coding for glucose-metabolizing enzymes are less variable than loci coding for other enzymes. The most precise measure of genetic variation is the proportion of heterozygous loci. On the average, in the three sibling species, loci not involved in glucose metabolism are somewhat more than twice as variable as those coding for glucose-metabolizing enzymes.

Table III. Genetic variation in *Drosophila tropicalis*

Gene	No. wild genomes sampled	No. populations sampled	Proportion of polymorphic populations	Mean frequency of heterozygous individuals	Frequency of most common allele	No. alleles with frequency ≥ 0.01
Group I (glucose-metabolizing enzymes)						
α Gpdh	2534	23	0.000	0.016 \pm 0.005	0.993	1
Ald	74	10	1.00	0.330 \pm 0.078	0.770	3
G3pdh	46	6	0.333	0.099 \pm 0.053	0.957	2
Hk-1	686	14	0.200	0.094 \pm 0.038	0.962	3
Hk-2	670	15	0.750	0.112 \pm 0.031	0.940	2
Hk-3	756	18	0.167	0.049 \pm 0.027	0.975	2
Idh	914	24	0.222	0.078 \pm 0.032	0.966	2
Mdh-2	2452	23	0.000	0.009 \pm 0.004	0.995	1
Me-1	862	23	0.625	0.151 \pm 0.042	0.906	3
Me-2	332	5	0.750	0.197 \pm 0.058	0.883	3
Pgm-1	632	15	0.000	0.038 \pm 0.012	0.973	2
Tpi-2	674	20	0.000	0.041 \pm 0.013	0.973	3
Group II (other enzymes)						
AcpH-1	788	5	0.400	0.124 \pm 0.040	0.953	2
Adh	2072	23	0.889	0.335 \pm 0.034	0.738	3
Adk-1	620	22	1.00	0.525 \pm 0.022	0.500	3
Adk-2	566	22	0.333	0.137 \pm 0.048	0.924	3
Aph-1	2236	45	0.947	0.183 \pm 0.023	0.904	3
Est-2	808	10	1.00	0.137 \pm 0.009	0.896	3
Est-3	2518	34	1.00	0.321 \pm 0.022	0.784	3
Est-4	2658	45	0.857	0.321 \pm 0.011	0.813	2
Est-6	134	3	1.00	0.247 \pm 0.020	0.858	4
Est-7	2344	35	1.00	0.528 \pm 0.019	0.599	5
Lap-5	2802	43	1.00	0.505 \pm 0.029	0.638	5
Odh-1	1072	5	1.00	0.125 \pm 0.006	0.929	3
To	1938	12	0.417	0.101 \pm 0.025	0.961	2
Mean	1248	20	0.596	0.192 \pm 0.031	0.872	2.72

Table IV. Level of Polymorphism in Glucose-Metabolizing Enzymes (Group I) and in the Other Enzymes (Group II)

Enzymes	No. loci	Proportion of polymorphic loci per population	Proportion of heterozygous loci per individual	Average No. of alleles with frequency ≥ 0.01
<i>Drosophila willistoni</i>				
Group I	12	0.379 ± 0.075	0.104 ± 0.019	2.58 ± 0.23
Group II	16	0.732 ± 0.062	0.234 ± 0.047	3.13 ± 0.22
Group II/group I		1.93	2.25	1.21
<i>Drosophila equinoxialis</i>				
Group I	12	0.521 ± 0.124	0.154 ± 0.039	2.42 ± 0.29
Group II	15	0.863 ± 0.059	0.268 ± 0.039	3.00 ± 0.17
Group II/group I		1.66	1.74	1.24
<i>Drosophila tropicalis</i>				
Group I	12	0.337 ± 0.102	0.101 ± 0.026	2.25 ± 0.22
Group II	13	0.834 ± 0.073	0.276 ± 0.045	3.15 ± 0.27
Group II/group I		2.47	2.73	1.40
Total				
Group I	36	0.412 ± 0.059	0.120 ± 0.017	2.42 ± 0.14
Group II	44	0.807 ± 0.037	0.258 ± 0.025	3.09 ± 0.13
Group II/group I		1.96	2.15	1.28

Table V. Regression Coefficient of the Level of Polymorphism on the Log Equilibrium Constant (K_{eq}) of the Enzyme^a

Measure of polymorphism	Regression coefficient	Standard error	Degrees of freedom	t	P
<i>Drosophila willistoni</i>					
Frequency of heterozygotes	-0.0001	0.0144	13	-0.07	> 0.5
Frequency of 2nd most common allele	-0.0008	0.0062	13	-0.13	> 0.5
Index of polymorphism (1 - P)	+0.0035	0.1085	13	0.32	> 0.5
Het./Het. max.	+0.0003	0.0352	12	0.01	> 0.5
<i>Drosophila equinoxialis</i>					
Frequency of heterozygotes	+0.0043	0.0086	13	0.50	> 0.5
Frequency of 2nd most common allele	+0.0031	0.0069	13	0.45	> 0.5
Index of polymorphism (1 - P)	+0.0024	0.0078	13	0.31	> 0.5
Het./Het. max.	-0.0253	0.0194	11	-1.30	> 0.1
<i>Drosophila tropicalis</i>					
Frequency of heterozygotes	+0.0051	0.0087	13	0.59	> 0.5
Frequency of 2nd most common allele	+0.0052	0.0071	13	0.73	> 0.4
Index of polymorphism (1 - P)	+0.0045	0.0081	13	0.56	> 0.5
Het./Het. max.	-0.0008	0.0205	11	-0.04	> 0.5
Total					
Frequency of heterozygotes	+0.0093	0.0593	43	0.16	> 0.5
Frequency of 2nd most common allele	+0.0075	0.0346	43	0.17	> 0.5
Index of polymorphism (1 - P)	+0.0104	0.0465	43	0.22	> 0.5
Het./Het. max.	-0.0258	0.1173	38	-0.22	> 0.5

^a Four measures of polymorphism are used, t and P for the significance of the regression.

Johnson (1971) has argued that enzymes which catalyze irreversible and rate-limiting reactions along a pathway affect production of the end result of the pathway more severely than enzymes with equilibrium constants not very different from 1. Changes in the activity of the former type of enzymes will have larger effects on the pathway than changes in the latter, since the effect of the latter will be largely diluted by mass action. From these considerations, Johnson concludes that natural selection is more likely to promote stable polymorphisms among enzymes catalyzing physiologically irreversible reactions than among enzymes whose reactions are freely reversible or among enzymes with equilibrium favoring the reverse of the physiologically significant direction. According to his hypothesis, the regression of the amount of polymorphism at a given locus on the equilibrium constant of the enzyme should be positive.

In Table V, we give the coefficients of the linear regression of various measures of polymorphism on the logarithm of the equilibrium constant, K_{eq} , of the enzymes. The following equilibrium constants are taken from Johnson (1971): alcohol dehydrogenase, 1×10^{11} ; α -glycerophosphate dehydrogenase, 1×10^4 ; hexokinase, 155; isocitrate dehydrogenase, 20; malic enzyme, 20; phosphoglucomutase, 17; aldolase, 2×10^{-4} ; malate dehydrogenase, 7×10^{-13} . The following enzymes were not considered by Johnson: the equilibrium constants used are triosephosphate isomerase, 20 (Oesper and Myerhof, 1950); adenylate kinase, 1 (Bowen and Kerwin, 1956); glyceraldehyde 3-phosphate dehydrogenase, 8×10^{-1} (Walsh and Sallach, 1965). Besides the proportion of heterozygous individuals, we have used two measures of polymorphism used by Johnson (1971) and a third one suggested by him (personal communication). The "index of polymorphism" is estimated by $1 - P$, where P is simply the average frequency of the most common allele, weighted for sample size (Johnson, 1971). Het./Het. max. is the ratio of the observed heterozygosity to the heterozygosity which would occur if all alleles occurring with a frequency no less than 0.01 would in fact occur all with equal frequencies (and the genotypic frequencies would satisfy Hardy-Weinberg expectation).

It is clear that our data do not support Johnson's hypothesis. Each of the four measures of genetic polymorphism for each of three species of *Drosophila* and for the three species combined gives coefficients of regression none of which is significantly different from zero.

DISCUSSION

Studies of enzyme variation in *Drosophila* flies, house and field mice, man, and other animals and plants have shown that natural populations of these organisms contain a great deal of genetic variation. Much of this variation

appears to be maintained by balancing natural selection (see, for instance, Prakash *et al.*, 1969; Ayala *et al.*, 1971, 1972; Allard and Kahler, 1972). What forms of balancing selection are involved in each particular case largely remains to be ascertained. The occurrence of heterosis has been demonstrated for two different loci in *Drosophila* (Richmond and Powell, 1970; Wills and Nichols, 1971). That diversifying selection plays a major role has recently been demonstrated by Powell (1972), who found that levels of polymorphism are directly related to environmental heterogeneity. Kojima and his collaborators have shown that genotypic fitnesses are sometimes frequency dependent—a mechanism which underlies adaptation to environmental heterogeneity (Kojima and Yarbrough, 1967; Kojima and Tobari, 1969; Huang *et al.*, 1971).

The number of loci studied in any particular organism and the number of organisms studied are still fairly limited. Yet several patterns are starting to emerge (Ayala, 1972): (1) the amount of polymorphism varies considerably from locus to locus; (2) at a given locus, the amount and pattern of the variation remain fairly constant from population to population throughout the whole distribution of a given species; (3) among closely related species, the amount of variation at a given locus is fairly constant from species to species. These generalizations suggest that the amount of variation at a given locus may be an attribute of the physiological role of the enzyme. Can we predict at least to a first approximation the degree of variation at a locus from knowledge of the physiological function or the biochemical properties of the enzyme?

Johnson (1971) has suggested that the amount of polymorphism may be a direct, positive function of the equilibrium constant, K_{eq} , of the enzyme in its physiological milieu. He analyzed data from several *Drosophila* species and found a positive linear regression of the frequency of the second most common allele on the logarithm of K_{eq} . Similarly, there seemed to be a positive linear regression of the logarithm of his "index of polymorphism" ($1 - P$, see *Results*) on the logarithm of K_{eq} . It should be noted, however, that Johnson uses K_{eq} values which do not obtain at the physiologically significant pH, which must be around 7. For alcohol dehydrogenase, he has taken $K_{eq} = 10^{11}$ and for malate dehydrogenase $K_{eq} = 7 \times 10^{-13}$. These values obtain when the molarity of hydrogen ions, H^+ , is 1 (pH 0). At pH 7, the equilibrium constants are, approximately, 10^4 and 10^{-5} for alcohol dehydrogenase and malate dehydrogenase, respectively (see Johnson, 1960).

Our data do not support Johnson's hypothesis. We have obtained extensive data on the natural polymorphisms of three closely related species of *Drosophila*. Using four different measures of genetic polymorphisms, we have attempted to correlate the amount of polymorphism with the logarithm of

the equilibrium constant of the enzyme. In every case the regression of amount of polymorphism on $\log K_{eq}$ was not significantly different from zero.

Lack of linear correlation between amount of polymorphism and equilibrium constant is not surprising. Johnson argues that "enzymes which exert acute control over flow through the pathway," that is, enzymes catalyzing physiologically irreversible reactions, should be more polymorphic. If this rationale is valid, the amount of polymorphism should be high in enzymes both with very high and with very low K_{eq} . Both catalyze irreversible reactions, although the direction of the reaction is opposite. Yet he attempted a straight linear regression for enzymes whose K_{eq} ranged from 7×10^{-13} to 1×10^{11} . A V-shaped function, with the vertex around $K_{eq} = 1$ would have been more appropriate. The validity of his rationale may be questioned on other grounds as well. Even if natural selection acts strongly on enzymes catalyzing irreversible reactions, because of their crucial role in a pathway, it is not clear why the mode of selection should be in such cases balancing selection. It might be argued, at least with the same *a priori* probability, that we should expect directional selection to act on those enzymes. Because of their crucial role, natural selection might be expected to favor that form of the enzyme which optimizes for the organism the amount of flow in the reaction. Furthermore, controlling enzymes in a pathway often have allosteric sites as well as active sites of catalysis. Thus to maintain a functioning enzyme, at least two sites must be protected from radical change.

Kojima *et al.* (1970) studied enzyme variation in two natural populations of each of two species, *D. melanogaster* and *D. ananassae*, and in one natural population of each of three species, *D. simulans*, *D. affinis*, and *D. athabasca*. Six to 11 enzymes involved in glucose metabolism (group I) and four to eight other enzymes (group II) were studied in each population. On the average, somewhat more than 100 flies were studied from each population at each locus. They found that enzymes from group II were on the average more polymorphic than those from group I. Three measures of polymorphism were used. The ratio of the amount of polymorphism in group II to that in group I was 3.70, 4.95, and 3.27 for the number of polymorphic populations, the frequency of heterozygotes, and the number of alleles per locus.

Our data confirm the observation of Kojima *et al.* (1970) that glucose-metabolizing enzymes are, on the average, less polymorphic than other enzymes. Richmond (1972) has made an extensive study of genetic variation in *D. paulistorum*, another sibling species of the *D. willistoni* group. He has also found less variation in glucose-metabolizing enzymes than in other enzymes. In our study, however, the amount of polymorphism is only about twice as large in enzymes from group II as in those from group I. (It should be pointed out that in the number of alleles per locus we have included all alleles with frequency ≥ 0.01 , while Kojima *et al.* included only those with

frequencies ≥ 0.05 .) In our data, differences between the two groups of enzymes in the amount of polymorphism are similar to those found in the house mouse by Selander and Yang (1969) (see also Kojima *et al.*, 1970).

Gillespie and Kojima (1968) were the first to propose the hypothesis that enzymes involved in glucose metabolism should be less variable than other enzymes. This interesting hypothesis needs further examination and testing. If the rationale is correct, one should be able to find that, in general, enzymes involved in biochemical pathways essential to the organism and which act on a single species of substrate should be less variable than those involved in peripheral pathways and having broad specificities. A possible test for the hypothesis could be made by examining enzyme variation in plants. In green plants, enzymes involved in photosynthesis play an essential role analogous to that of the glucose-metabolizing enzymes in animals.

ACKNOWLEDGMENTS

We are grateful to Professors Eric E. Conn, Department of Biochemistry, and Theodosius Dobzhansky, Department of Genetics, University of California, Davis, who read the manuscript and made many valuable suggestions.

REFERENCES

- Allard, R. W., and Kahler, A. L. (1972). Patterns of molecular variation in plant populations. *Proc. Sixth Berkeley Symp. Math. Stat. Prob.*
- Ayala, F. J. (1972). Darwinian versus nondarwinian evolution in natural populations of *Drosophila*. *Proc. Sixth Berkeley Symp. Math. Stat. Prob.*
- Ayala, F. J., Powell, J. R., and Dobzhansky, T. (1971). Enzyme variability in the *Drosophila willistoni* group. II. Polymorphisms in continental and island populations of *Drosophila willistoni*. *Proc. Natl. Acad. Sci.* **68**:2480.
- Ayala, F. J., Powell, J. R., Tracey, M. L., Mourão, C. A., and Perez-Salas, S. (1972). Enzyme variability in the *Drosophila willistoni* group. IV. Genic variation in natural populations of *Drosophila willistoni*. *Genetics* **70**:113.
- Bowen, W. J., and Kerwin, T. D. (1956). The kinetics of myokinase. II. Studies of heat denaturation, the effects of salts and the state of equilibrium. *Arch. Biochem. Biophys.* **64**:278.
- Gillespie, J. H., and Kojima, K. (1968). The degree of polymorphism in enzymes involved in energy production compared to that in nonspecific enzymes in two *Drosophila ananassae* populations. *Proc. Natl. Acad. Sci.* **61**:582.
- Harris, H. (1970). *The Principles of Human Biochemical Genetics*, Elsevier, New York.
- Huang, S. L., Singh, M., and Kojima, K. (1971). A study of frequency-dependent selection observed in the esterase-6 locus of *Drosophila melanogaster* using a conditional media method. *Genetics* **68**:97.
- Johnson, G. (1971). Metabolic implications of polymorphisms as an adaptive strategy. *Nature* **232**:347.
- Johnson, M. J. (1960). Enzymic equilibria and thermodynamics. In Boyer, P. D., Lardy, H., and Myrbäck, K. (eds.), *The Enzymes*, Vol. 3, Part B, Academic Press, New York and London, pp. 407-441.
- Kojima, K., and Tobari, Y. N. (1969). The pattern of viability changes associated with genotype frequency at the alcohol dehydrogenase locus in a population of *Drosophila melanogaster*. *Genetics* **61**:201.

- Kojima, K., and Yarbrough, K. M. (1967) Frequency-dependent selection at the esterase-6 locus in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci.* **57**:645.
- Kojima, K., Gillespie, J. H., and Tobari, Y. N. (1970). A profile of *Drosophila* species' enzymes assayed by electrophoresis. I. Number of alleles, heterozygosities, and linkage disequilibrium in glucose-metabolizing systems and some other enzymes. *Biochem. Genet.* **4**:627.
- Marshall, D. R., and Allard, R. W. (1970). Isozyme polymorphisms in natural populations of *Avena fatua* and *A. barbata*. *Heredity* **25**:373.
- Oesper, P., and Myerhof, O. (1950). The determination of triose phosphate isomerase. *Arch. Biochem.* **27**:223.
- Powell, J. R. (1972). Genetic polymorphisms in varied environments. *Science* **74**:1035.
- Prakash, S., Lewontin, R. C., and Hubby, J. L. (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**:848.
- Richmond, R. C. (1972). Enzyme variability in the *Drosophila willistoni* group. III. Amounts of variability in the superspecies, *D. paulistorum*. *Genetics* **70**:87.
- Richmond, R. C., and Powell, J. R. (1970). Evidence of heterosis associated with an enzyme locus in a natural population of *Drosophila*. *Proc. Natl. Acad. Sci.* **67**:1264.
- Selander, R. K., and Yang, S. Y. (1969). Protein polymorphism and genic heterozygosity in a wild population of the house mouse (*Mus musculus*). *Genetics* **63**:653.
- Selander, R. K., Yang, S. Y., Lewontin, R. C., and Johnson, W. E. (1970). Genetic variation in the horseshoe crab (*Limulus polyphemus*) a phylogenetic relic. *Evolution* **24**:402.
- Spassky, B., Richmond, R. C., Perez-Salas, S., Pavlovsky, O., Mourão, C. A., Hunter, A. S., Hoenigsberg, H., Dobzhansky, T., and Ayala, F. J. (1971). Geography of the sibling species related to *Drosophila willistoni*, and of the semispecies of the *Drosophila paulistorum* complex. *Evolution* **25**:129.
- Walsh, D. A., and Sallach, H. J. (1965). Purification and properties of chicken liver D-3-phosphoglycerate dehydrogenase. *Biochemistry* **4**:1076.
- Wills, C., and Nichols, J. (1971). Single-gene heterosis in *Drosophila* revealed by inbreeding. *Nature* **233**:123.