THE MODE OF SELECTION AT THE POLYMORPHIC ESTERASE 6 LOCUS IN CAGE POPULATIONS OF *DROSOPHILA MELANOGASTER*¹

KAREN YARBROUGH² AND KEN-ICHI KOJIMA

Department of Zoology, University of Texas, Austin, Texas 78712

Received June 22, 1967

IT is generally accepted that some selective mechanism is responsible for the maintenance of genetic polymorphisms in natural populations. In the past several years there have been mounting evidences indicating that the extent of polymorphisms in outbreeding populations, ranging from insects to humans, is far greater than previously considered. Electrophoretically separable protein and enzyme systems in these organisms are among the major sources of such evidences (e.g. LEWONTIN and HUBBY 1966; JOHNSON *et al.* 1966; HARRIS 1966).

Heterosis in fitness at the locus level may be a part of the answer to the above problem. As a matter of fact, several investigators pointed out theoretical possibilities under which a number of heterotic loci might be maintained at their equilibria (e.g. SVED, REED and BODMER 1967; KING 1967; MILKMAN 1967). However, there are experimental evidences indicating that another mechanism may be responsible for some cases of polymorphisms (e.g. KOJIMA and YARBROUGH 1967; TOBARI and KOJIMA 1967).

The objective of this study was to investigate the mode of selection associated with the polymorphism at the Esterase 6 locus in *Drosophila melanogaster*.

MATERIALS AND METHODS

In the fall of 1964, two wild-type inbred lines were crossed to form a large random-mating population which was maintained for 30 generations prior to the initiation of this experiment. It was found that this population was segregating for Esterase 6^F and 6^S alleles (hereafter denoted as the *F* and *S* alleles) at a frequency of approximately 30% *F*. This population will be referred to as the base population.

Throughout the period of time prior to the beginning of this experiment, the base population was maintained in four food cartons containing a cornmeal-molasses medium at approximately 25°C. Each carton was capable of supporting a population of several hundred flies. In each generation samples of 50 adult *Drosophila* from individual cartons, 200 flies in all, were mixed to form each of the four cartons for the next generation. This procedure was used to maintain a relatively large breeding population and to bring the genetic constitution to approximate gene frequency and linkage equilibria.

During generation 31, a number of single-pair matings from the base population were made, and each line was divided into several full-sib mating sublines for the next one to five generations. During this period of time, 14 lines homozygous for the *S* allele and six lines homozygous for the *F* allele were isolated by electrophoretic analyses of progeny genotypes in each subline.

¹ The first part of this work was carried out at North Carolina State University with the financial support of Public Health Service Grant GM 11546, and the final part was completed at the University of Texas with the aid of Atomic Energy Commission Contract AT-(40-1)-3681.

² Department of Biology, University of Southern Mississippi, Hattiesburg, Mississippi 39401.

Preparatory to starting the cage populations, each of the 20 lines, homozygous either for the *F* or *S* allele, was transferred to replicated cartons in order to increase the size of each line. From the 20 lines eight cage populations were started. Two initial frequencies of the *F* allele and two types of culture media were used with two replications of each experimental condition. Cornmeal-molasses medium was used in four of the cages while the other four contained banana medium.

Each of the eight population cages was started with a total of 506 adult males and 506 adult virgin females. The total number of flies of each genotype in a cage represented samples of equal size from either the *F* or *S* homozygous lines. Table 1 provides an identification and description of each of the eight cages.

By recombining the 20 Esterase 6 homozygous lines, it was felt that the original genetic constitution of the base population would be approximately regained with the exception of the Esterase 6 locus. It was further assumed that this procedure reduced the possibility of creating major linkage disequilibria between the Esterase 6 locus and all other loci.

In order to have a standard of reference for allele frequency changes at this locus under the cage population conditions over time, the original base population was transferred into a cage containing cornmeal-molasses medium. This cage was maintained for a period of ten months (approximately 20 generations), and the frequency of *F* remained stable between 0.27 and 0.32 for a sample size of approximately 150 individuals per observation.

During this study, the experimental cage populations were maintained on the appropriate medium under a constant temperature. Several days after the adults were placed in a cage, the food cups containing the eggs and larvae of a new generation were retained and their parent flies discarded. The adult *Drosophila* which emerged in the new cage 11 days later were anesthetized with CO₂ and removed. Several hundred *Drosophila* were retained to provide a sample of flies for electrophoretic determination of their genotypes. The more than 1,000 remaining adults were transferred to a new cage to begin the next generation.

The electrophoretic procedures followed in this experiment utilized, in general, the techniques developed by JOHNSON (1964). All gels were stained in a phosphate buffered solution containing α -naphthyl acetate, β -naphthyl acetate, *n*-propanol, and Fast Blue RR (or Fast Garnet GBC) salt. This staining technique permitted an accurate scoring of genotypes since the Esterase 6 bands appeared pinkish-red while all the other esterase bands were blue-black. The original description of the banding patterns of the three genotypes at this locus and their Mendelian behavior may be found in a paper by WRIGHT (1963).

RESULTS

The observed frequencies of the *F* allele in the eight cages are presented in Figures 1 (cornmeal medium) and 2 (banana medium) for the entire span of

TABLE 1
Identification and description of cage populations

Case No.	No. flies of each genotype		<i>f(F)</i> *	Medium
	<i>FF</i>	<i>SS</i>		
1	900	112	0.889	Cornmeal
2	102	910	0.101	Cornmeal
3	900	112	0.889	Banana
4	102	910	0.101	Banana
5	900	112	0.889	Cornmeal
6	102	910	0.101	Cornmeal
7	900	112	0.889	Banana
8	102	910	0.101	Banana

* The initial frequency of the *F* allele.

generations. The observed points are based on samples of size 100 to 200. These two figures clearly show that under both medium conditions a decline in F frequency occurred in those cages initiated at the high F frequency level, while F frequency increased in the cages started from the low F frequency. It appears that the rate of decline in F frequency in the high F cages was somewhat greater on the banana medium than on the cornmeal medium. After 30 generations, the F frequencies in all the cornmeal cages were converging between 0.45 and 0.25, indicating that the original equilibrium frequency of about 0.3 would eventually be attained. It should be noted here that the rate of the approach to the point of equilibrium slowed down considerably after the first 15 generations or so.

On the other hand, only three cages under the banana medium condition converged to the same range of F frequency as above, while one cage (cage 8) behaved in a quite different manner. In this cage, an initial increase in the frequency of the F allele to 0.21 was observed, and thereafter the F frequency declined to a lower level and was stabilized at approximately 0.10.

The pattern of frequency changes shown in Figures 1 and 2 seemed to suggest the presence of heterozygote superiority in fitness at this locus. Consequently, the Wrightian fitness model of $W_{FF} = 1-s$, $W_{FS} = 1$ and $W_{SS} = 1-t$ was used to estimate the values of the constants s and t . The method employed was the two-stage computer program developed by SMITH (1965) which was based on a minimum chi-square fit of observed allele frequency changes to those values pre-

FIGURE 1.—Allele frequency changes at the Esterase 6 locus in two high and two low F frequency cages maintained on cornmeal medium. Generations where there were no data are skipped by dashed lines.

FIGURE 2.—Allele frequency changes at the Esterase 6 locus in two high and two low F frequency cages maintained on banana medium. Generations where there were no data are skipped by dashed lines.

dicted by the model. At the first stage of the program a wide range of s and t combinations (e.g., $s = 0.0, 0.1, 0.2, \dots, 1.0$ combined with the same set of t) was chosen, and the value of chi-square fit was computed for each combination of s and t . The combination(s) which gave minimum chi-square(s) in comparison to the surrounding ones were used as the initial s and t for the second stage computation which iteratively minimized chi-square value by incrementing and decrementing s and t . The iteration was carried out until the chi-square did not decrease any longer by a 2.5% change in each of s and t values from the last iteration. The cut-off point of 2.5% was chosen, because the rate of decrease in chi-square value due to this degree of incrementing and decrementing was trivially small in all cases examined in this study. The final estimates of s and t for a given case were chosen as those with the best fit to the observed allele frequencies. The adequacy of the above estimating procedure was tested using a set of computer-simulated data, before actual data with unknown s and t were used.

The estimation of constant s and t was carried out for the data in Figures 1 and 2 for varying numbers of generations, and also for the curves smoothed out by taking moving average. The total result of this estimation effort is too bulky to present here; however, some typical examples may help to understand the situation. The results from the first 15 generations of the original data of the four cornmeal cages were (0.20, 0.16, 45.0) for cage 1, (0.20, 0.10, 56.4) for cage 5,

(0.72, 0.50, 16.5) for cage 2, and (0.44, 0.20, 43.2) for cage 6, where the inside of each parenthesis is arranged as (s , t , chi-square).

There were a few unsatisfactory aspects in the general results of the above analysis. First of all, the value of minimum chi-square was too large in the majority of cases. A close examination sometimes revealed that a large portion of unsatisfactorily large chi-square was attributable to the failure to fit the observed frequency changes during the early generations with the estimated fitness values from a large number of generations.

Another unsatisfactory aspect of importance was apparent differences in the magnitude of s and t estimates obtained from the high F versus low F frequency cages. In general, the s and t estimates from the high F cages were smaller in their magnitude than those from the low F cages. This last objection might be due to the differences of genes which were initially linked with the F and S alleles in the 20 lines extracted. However, taking all indications together, it was decided to abandon the model of constant s and t values and instead to look towards another model proposed earlier (KOJIMA and YARBROUGH 1967).

Approaching the problem of egg-to-adult viability of the genotypes in question, the authors reported evidence of frequency-dependent selection at this locus in the same base population. Using the same set of FF and SS homozygous lines, they synthesized populations of fertilized eggs which were laid in approximate Hardy-Weinberg ratios at the F frequencies of 0.70, 0.50, 0.30 and 0.15. When adults emerged, their genotypes were determined by the same techniques as those in the present study. The relative fitness values were computed by comparing the genotypic ratios in eggs with those in adults. Large selective differences were found to exist among the genotypes when the allele frequency was deviating from the original equilibrium value, ($F = 0.3$), and, further, the selective advantages were conferred upon the genotypes which were deficient in frequency in comparison to their equilibrium frequencies.

There were weaknesses and unavoidable assumptions involved in the analyses of KOJIMA and YARBROUGH's (1967) experiment and the present experiment, which were intrinsic to the designs of the two experiments. However, most of these weaknesses were different between the two experiments, and the results of the two should complement each other to bring out the total picture.

Based upon the above frequency-dependent selection model, the following analysis of fitness values was carried out with the cage population data. It is important to note that generations in the present study were always discrete. Let p_t be the frequency of the F allele among the adults at generation t . The fertilized eggs for the next generation are then expected to be laid in proportion to $p_t^2 : 2p_t(1-p_t) : (1-p_t)^2$ for FF , FS , and SS genotypes, respectively. Thus under the assumption that there is no selective difference among the three genotypes, the expected numbers of the three genotypes in N_{t+1} individuals sampled from generation $(t+1)$ are

$$\begin{array}{ccc} FF & FS & SS \\ p_t^2 N_{t+1} & 2p_t(1-p_t)N_{t+1} & (1-p_t)^2 N_{t+1} \end{array} \quad (1)$$

The ratio of the observed numbers of the three genotypes in generation $(t+1)$

to the above expected numbers will then provide a set of three relative fitness values for the genotypes involved in the transition from generation t to generation $(t + 1)$. Letting the three fitness values for genotypes FF , FS and SS be U_t , V_t , and W_t , respectively, for the transition of generation t to generation $(t + 1)$, the above method will estimate U_t/K_t , V_t/K_t , and W_t/K_t as the three relative fitness values, where K_t is the "so called" mean fitness. The proposed analysis is very sensitive to the sampling errors in the observed value of p_t , and, as Figures 1 and 2 indicated, there were considerable generation-to-generation fluctuations associated with the individual estimates of p_t . For this reason, the fluctuations in p_t were reduced by substituting a five-point moving average of p_t for the actual observed value.

The relative fitness values estimated according to the above procedure are listed in detail in Appendix Table 1 for cages 1 and 2 only, and the pooled values for early, middle and late stages (or generations) are presented graphically for each of cages 1 to 7 in Figure 3. The numbers of generations pooled in these figures are not consistent owing to missing data (skipped by dotted lines in Figures 1 and 2) and different speeds of frequency changes among the cages. The data from cage 8 were omitted from this analysis since there seemed to be an additional factor(s) involved in its evolutionary trends.

Considering cages 1, 3, 5 and 7 (high initial F frequency), it was evident that during the early generations there was a considerable amount of divergence between the selective values, and, in general, SS was superior to FS which was, in turn, superior to FF . However, during the middle to late states the degree of divergence decreased considerably. In the case of the low F frequency cages

FIGURE 3.—Relative fitness for cages with cornmeal (1, 2, 5, 6) and with banana (3, 4, 7) medium. The fitness value estimates of the three genotypes FF , FS and SS are pooled over various intervals of generations specified at the bottom of each graph. The estimates at individual generations are found in Appendix Table 1 for cages 1 and 2.

(cages 2, 4 and 6), the same general trends toward a decreasing amount of divergence were observed, and the order of ranking seemed to reverse, i.e., $SS < FS < FF$, in general. This reversal of selective values of the homozygous genotypes was quite striking and consistent for the early generations of all seven cages. The selective values of the heterozygotes did not appear to be significantly superior to those of either homozygote except at the beginning stages of cage 4. Rather, the selective value of FS tended to be intermediate between the two homozygotes. Considering data from all cages together, it appeared that no single genotype was consistently superior to any other at or near the point of equilibrium. This finding is in complete agreement with the conclusions reached in the earlier study by KOJIMA and YARBROUGH (1967). However, this concept should be accepted with some reservation until the evolutionary behavior of cage 8 and the differences among some of replicate cages are adequately explained.

DISCUSSION

The mode of selection at the Esterase 6 locus found in these experiments presents a mechanism by which a random mating population can preserve genic polymorphisms as an alternative to the popular concept of heterozygote superiority. Frequency-dependent selection as a mechanism for maintenance of genetic variability was suggested earlier by TEISSIER (1954a, b) and more recently by KOJIMA and YARBROUGH (1967).

Under this type of selective system large deviations of allele frequencies from their equilibrium values are forced to return quickly to the neighborhood of the original frequency range. Within this range, differences in selective values of segregating genotypes are far smaller than those associated with genotypes in a population which exhibits a considerable departure from the state of equilibrium.

It is conceivable that the rate of return of perturbed populations to a state of stable equilibrium may differ considerably depending upon the particular genetic systems (e.g., locus, inversion, etc.; for a case of frequency-dependent selection with inversion systems, see TOBARI and KOJIMA 1967) and internal and external conditions of the populations. As a matter of fact, MACINTYRE and WRIGHT (1966) reported a variety of genetic situations involving the Esterase 6 locus in *Drosophila melanogaster* and *D. simulans*. The rate of return to a gene frequency equilibrium varied from situation to situation: in some instances cages started from 80% and 20% F converged to practically the same frequency within a period of several generations, while in other cages there remained a difference of approximately 27% after 21 generations. The rate of return to an equilibrium state must be affected by such factors as differential physiological actions of the alleles involved, linkage relationships, and all other genes which might exhibit epistasis with the locus in question. It is difficult to evaluate the relative contributions of these factors in various populations in their study.

In considering this mode of selection as a general mechanism for maintaining genetic variability, its prevalence in other systems in a variety of populations must be investigated. Some evidences in support of the prevalence of this mecha-

nism exist in reports concerning mating behavior of some *Drosophila* species (PETIT 1958; SPIESS and LANGER 1964; EHRMAN *et al.* 1965; EHRMAN 1966). Furthermore, PAVLOVSKY and DOBZHANSKY (1966) reported that selection coefficients associated with the third chromosome inversion polymorphism in *Drosophila pseudoobscura* were variable depending upon the frequency of the inversions. The concept proposed in the present report is similar to the selective mechanism associated with the inversion systems of the second and third chromosomes of *D. ananassae* (TOBARI and KOJIMA 1967).

However, some cautions must be taken with respect to the method of analysis using the ratio of observed adult frequency to the expected (fertilized egg) frequency. This method is adequate for analyzing the selection pattern associated with zygotic survival, but not for dealing with selection at the production of fertilized eggs. The authors are grateful to Drs. M. KIMURA, R. E. COMSTOCK and R. C. LEWONTIN who communicated directly to KOJIMA on this point. Our earlier data with the same material as the one used here (KOJIMA and YARBROUGH 1967) indicated that there was some heterosis in female fecundity when the three genotypes were tested separately. However, there were far stronger selective differences associated with zygotic viability, which seemed to control the point of gene frequency equilibrium. The pattern of change in viability selection was, in all practical purposes, identical to the one found in the present cage experiments.

Taking these evidences *in toto*, it is felt that the phenomenon of frequency-dependent selection with approximate selective neutrality at equilibrium is a fairly widespread mechanism for maintaining large amounts of genetic variability in natural populations.

The authors express their sincere appreciation to Dr. FRANK B. ARMSTRONG and Dr. FRANK M. JOHNSON for their technical advice. The second author is very grateful to Dr. MICHAEL LERNER, who pointed out the papers by TEISSIER quoted in this paper. The authors are also grateful for technical assistance given by MYRTLE WING, MARSHA FRELUND, ANN WILSON and ANN BAZER at North Carolina State University and MAUREEN SCHMIDT, TERRI LESHIKAR and MARY BELK at the University of Texas.

SUMMARY

A series of cage populations with divergent allele frequencies at the Esterase 6 locus were derived from a random mating population which was at gene frequency equilibrium with respect to this locus. Allele frequency changes were recorded for a period of 30 generations. The analysis of selective values associated with the genotypes at this locus during the early generations showed that large selective value differences existed among the three genotypes, forcing the populations to return rapidly to the neighborhood of the original equilibrium frequency. The disadvantageous genotypes were, in general, those which existed in frequencies in excess of the equilibrium frequency. During the latter generations the rate of return to the equilibrium frequency became extremely slow, and analysis of selective values during this stage indicated that approximate selective neutrality existed among the three genotypes.

APPENDIX TABLE 1

Relative fitness estimates (observed/expected) from Cages 1 and 2

Generation transition	Cage 1			Cage 2		
	FF	FS	SS	FF	FS	SS
0-1				1.810	1.706	0.833
1-2	0.953	1.184	1.377	2.877	1.254	0.784
2-3	0.856	1.523	0.874	1.506	1.028	0.935
3-4	1.030	0.858	1.524	1.784	1.051	0.839
4-5	1.073	0.695	2.139	1.045	1.055	0.939
5-6	0.911	1.193	1.086	1.128	0.999	0.968
6-7	0.864	1.183	1.960	0.962	1.101	0.901
7-8	0.953	1.117	0.775	1.209	0.910	1.034
8-9	1.006	0.995	0.973	1.137	1.058	0.880
9-10	0.760	1.243	1.319	0.881	1.025	1.017
10-11	0.985	0.984	1.137	1.046	1.142	0.788
11-12	0.811	1.052	1.441	0.827	1.111	0.929
12-13	0.869	1.109	0.977	1.079	1.018	0.937
13-14	0.919	0.980	1.406	1.081	0.853	1.172
14-15	0.812	1.296	0.543	0.899	1.126	0.867
15-16	0.923	1.061	0.972	1.005	1.196	0.707
16-17	0.820	1.159	0.885	0.766	1.038	1.056
17-18	0.935	1.130	0.780	0.831	0.796	1.343
18-19	0.884	1.039	1.059	1.001	0.924	1.097
19-20	0.746	1.128	1.014	1.004	1.159	0.810
all others*	0.953	0.999	1.038	0.969	1.028	0.979

* These estimates were obtained from the data pooled over all remaining generations.

LITERATURE CITED

- EHRMAN, L., 1966 Mating success and genotype frequency in *Drosophila*. *Animal Behavior* **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.
- HARRIS, H., 1966 Enzyme polymorphisms in man. *Proc. Roy. Soc. Lond. B* **164**: 298-310.
- 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 Samoa. *Proc. Natl. Acad. Sci. U.S.* **56**: 119-125.
- KING, J. L., 1967 Continuously distributed factors affecting fitness. *Genetics* **55**: 483-492.
- KOJIMA, K., and K. M. YARBROUGH, 1967 Frequency dependent selection at the Esterase 6 locus in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.* **57**: 645-649.
- 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., 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. D., 1967 Heterosis as a major cause of heterozygosity in nature. *Genetics* **55**: 493-495.

- PAVLOVSKY, O., and TH. DOBZHANSKY, 1966 Genetics of natural populations. XXXVII. The coadapted system of chromosomal variants in a population of *Drosophila pseudoobscura*. *Genetics* **53**: 843-854.
- PETIT, C., 1958 Le déterminisme génétique et psycho-physiologique de la compétition sexuelle chez *D. melanogaster*. *Bull. Biol. France Belgique* **92**: 248-329.
- SMITH, D., 1965 A formulation and evaluation of methods for estimating fitness values in one- and two-locus genetic systems. North Carolina State University Library, Raleigh.
- SPIESS, E. B., and B. LANGER, 1964 Mating speed control by gene arrangement carriers in *Drosophila persimilis*. *Evolution* **18**: 430-444.
- SVED, J. A., E. REED, and W. F. BODMER, 1967 The number of balanced polymorphisms that can be maintained in a natural population. *Genetics* **55**: 469-481.
- TEISSIER, M. G., 1954a Conditions d'équilibre d'un couple d'allèles et supériorité des hétérozygotes. *Compt. Rend.* **238**: 621-623. — 1954b Sélection naturelle et fluctuation génétique. *Compt. Rend.* **238**: 1929-1931.
- TOBARI, Y., and K. KOJIMA, 1967 Selective modes associated with inversion karyotypes in *Drosophila ananassae*. I. Frequency-dependent selection. *Genetics* **57**: 179-188.
- WRIGHT, T. R. F., 1963 The genetics of an esterase in *Drosophila melanogaster*. *Genetics* **48**: 787-801.