# SELECTION IN EXPERIMENTAL POPULATIONS: 11. COMPONENTS OF SELECTION AND THEIR FLUCTUATIONS IN TWO POPULATIONS OF *DROSOPHILA MELANOGASTER\**

SERGEY POLIVANOV **AND** WYATT W. ANDERSON

*Department of Biology, The Catholic Uniuersity of America, Washington, D. C. 20017*  and *Department of Biology, Yale University, New Haven, Connecticut 06520* 

Received May 13, 1969

POPULATION biologists have long followed gene frequency changes in model populations in attempts to verify various aspects of the mathematical theory of natural selection. The selection has been assumed to be constant in many of these studies, although it is widely recognized (e.g., **TEISSIER** 1954) that the selection in fact fluctuates with the physical and biological elements of the environment. In only a few instances have experiments been specifically designed to reveal how the selection varies from generation to generation (KOJIMA and **YARBROUGH** 1967; **ANDERSON** 1969). Likewise, most experiments on breeding populations have been designed to estimate the overall selective values of the genotypes, and our knowledge of how the selection is partitioned into components such as viability and fertility is notably incomplete. **TEISSIER** (1942) first showed how to carry out such a partition of selection, by scoring the frequencies of the genotypes after selection by viability but before selection by fertility. **WILSON**  (1968) and **ANDERSON** (1969) performed similar analyses. **All** three of these workers assumed constant selection and partitioned into components the selective values estimated over many generations. The purpose of this article is to show how the selection in *each* generation of an experimental population may be divided into components **of** viability and fertility. We shall use this information to demonstrate that selection has significantly fluctuated over a period of thirteen generations in two experimental populations of *Drosophila melanogaster.* 

# **MATERIALS AND METHODS**

TWO replicate populations were begun with 100 pairs of *Drosophila melanogaster* heterozygous for the mutant allele Stubble *(Sb).* This mutant, located at map position *58.2* on the third chromosome, is a dominant marker for reduced bristles and is lethal when homozygous. The founders of **the** populations were from a stock derived from POLIVANOV'S (1964) population, Stubble Mono 1. The Sb allele was introduced into this population six years before the present experiments were begun. Since 1964. the stock, Mono 1, has been maintained in mass cultures; since *Sb* was not balanced by any other lethal, selection for Sb heterozygotes has **been** necessary at each transfer.

The populations were maintained in polyethylene cages  $(22 \text{ cm} \times 33 \text{ cm} \times 6 \text{ cm})$  in a circulating-air incubator at  $25 \pm 1$ °C. The relative humidity was not controlled. Four food cups

\* Dedicated to PROFESSOR TH.DoBzHANsKY-teacher and friend-on his seventieth birthday.

#### TABLE *<sup>1</sup>*

*Frequencies of adults emerging in samples of eggs from the cross*  $Sb/+99 \times Sb/+33$ 



containing 30 to 35 ml of **SPASSKY'S** (1943) cream-of-wheat medium were used in each cage; six drops of a suspension of Fleischmann's yeast were added to each cup before it was put into the cage. The generations were discrete, according to the following procedure.

**1.** Adults were given fresh food and a sample of eggs was taken.

2. All adults were etherized and scored for genotype.

3. The adults were returned to the population and allowed to lay eggs for nine days.

4. The adults surviving the nine days of egg laying were etherized, scored for genotype, and then discarded.

5. Adults **of** the next generation usually began to emerge on the tenth day. Emergence was allowed to continue for a total of nine days, and then the old food cups were discarded.

6. These adults were given fresh food and the cycle beginning with step one was repeated.

The adults at discard were not recorded in the first generation. The adults in generation 5A laid the egg sample in **6A;** the larvae and pupae which would form the adults in 6A were accidentally discarded instead **of** the "old" adults. These "old" adults (the "adults at discard" in 5A) were allowed to lay eggs for a new sample and, thereafter, for a new generation. These "old" adults are listed as generation 5B, and their offspring form the samples in 6B. In generation 7 the adults were not censused before they began to lay eggs for the new generation.

The egg samples were arranged by putting approximately 75 eggs in each of six replicate half-pint bottles containing **SPASSKY'S** medium, and to which four drops **ol** a suspension of Fleischmann's yeast was added. All bottles with eggs were kept in the same incubator where the populalations were maintained. The egg samples were cultured under nearly optimal conditions. The adults obtained from them should represent the two nonlethal genotypes in their zygotic frequencies, adjusted for the absence of the *Sb* homozygotes We can readily infer the zygotic frequencies of all three genotypes. A series of cultures were set up to determine if there was any selection, other than the elimination of the lethal homozygotes, during the culturing of the egg samples. One hundred and twenty pairs of *Sb* heterozygotes from the stock used to initiate the populations were put into a population cage, and egg samples were taken in the usual way. Eighteen bottles of 75 eggs each were set up. Two hundred pairs of *Sb* heterozygotes, obtained as the **F,** offspring of the egg samples taken in the populations at generation 11, were likewise used to set up egg cultures representing the two populations after much **of** the selection had occurred. The expected per cent of the *Sb* heterozygotes was 66.7. The data are summarized in Table 1. There is no indication that the frequencies of genotypes among adults hatching from the egg samples depart from the expected frequencies. We may confidently infer the genotype frequencies among the newly formed zygotes from those among the adults hatching from the egg samples.

Adults were removed from the populations for scoring with a vacuum pump connected to a trap.

## **STATISTICAL ANALYSIS AND RESULTS**

Recorded in [Table](#page-2-0) *2* for each generation are the frequencies of *Sb* heterozygotes

<span id="page-2-0"></span>among the adults hatching from egg samples, among the adults at reproduction, and among the adults at discard. Data for females and males are recorded separately.

# **TABLE** *2*





\* See MATERIALS AND METHODS for explanation. \*\* Sample not taken.



**FIGURE 1 .-Differences in the frequency of Stubble heterozygotes among females and males.**  Upper, Population 1; Lower, Population 2.

We may compare the genotypic frequencies in males and females at the time of reproduction for evidence of differential viabilities in the sexes. Differential fertilities, however, do not produce different genotypic frequencies in the two sexes in any of the three samples per generation. The differential frequencies will exist among the gametes, but the process of mating will reassort the alleles equally between males and females. In Figure 1 are plotted the differences in the heterozygote frequencies in females and males of each sample. The differences among adults hatching from the egg samples are almost wholly due to sampling error and provide a standard of comparison for the samples among the adults. With a few exceptions, the differences among adults at reproduction and at discard are not much larger than those among the adults raised from the egg samples. There seems to be no pattern to variation in the differences; they switch irregularly from positive to negative. We feel justified in combining the genotype frequencies in females and males and in estimating common viabilities for the two sexes.

Our detailed analysis of selection is possible only with alleles which are markers for phenotypic characters when heterozygous but lethal when homozygous. With nonlethal alleles there are too many parameters **to** estimate for the data available over one generation. With lethal alleles the selective value of one genotype, the lethal homozygote, is known, and estimation of all relevant parameters becomes possible.

The viability from the egg stage to the adults at reproduction is estimated from the samples at these two life stages, according to the following model.



The maximum likelihood estimate of V is

$$
\hat{V}(T) = \frac{H_A(T) (1 - H_E(T))}{H_E(T) (1 - H_A(T))}
$$

The true frequencies of the genotypes among the adults at reproduction are known *exactly*, since the entire population was scored. The variance of  $\hat{V}(T)$  depends only on the sampling error in the egg samples. In samples of large size, the variance of  $\hat{V}(T)$  is approximately

 $\frac{\hat{V}(T)^2}{\sum_{i=1}^{n}T_{i}+T_{i}}$ , where  $T_{E}(T)$  is the total number of adults which were counted from the egg sample at generation T.  $\operatorname{Var} \mathring{\operatorname{V}}(\operatorname{T}) = -$ 

The model used to estimate fertility assumes that it is alike in the two sexes. Since there is only one degree of freedom in the data from the egg samples, one parameter is the most we can hope to estimate. If, as is almost certain, the selection via fertility is different in males and females, then our single estimate will be quite close to the *average* of the selections in the two sexes **(ANDERSON** 1969). This average is perhaps the most useful single index of the selection by fertility. Included in fertility are mating frequency of males and fecundity of females, along with any other selective factors which act between the time the adults are first scored and the time the egg sample is taken. The model for estimation is as follows.



The frequency of the *Sb* allele,  $Q(T)$ , among those gametes which will combine<br>to form the next generation is  $Q(T) = \frac{H_A(T) \cdot F(T)}{2(H_A(T) \cdot F(T) + N_A(T))}$ . We assume random mating, and we further assume that the mating ability of the male genotypes and the fecundity of the female genotypes do not depend on the genotypes of the mates. Thus, random mating is equivalent to the random combination of the alleles in their frequencies among the gametes. The model proceeds as below. Genotypes  $Sb/+$   $+/+$ Expected frequencies among  $2Q(T)$   $1-Q(T)$ <br>adults hatching from  $2Q(T)$   $1-Q(T)$ adults hatching from  $\frac{2Q(T)}{1+Q(T)}$   $\frac{1-Q(T)}{1+Q(T)}$ Observed frequencies among adults hatching from the egg samples  $H_E(T+1)$   $N_E(T+1)$ 

The maximum likelihood estimate of F is

 $\mathbf{\hat{F}}(T) = \frac{2H_E(T+1) (1-H_A(T))}{H_A(T) (2-3H_E(T+1))}$ , and in samples of large size the variance

of  $\mathbf{\hat{F}}(\mathbf{T})$  will be approximately

$$
\text{Var } \hat{F}(T) = \frac{4 \, \hat{F}(T)^2 \, (1 - H_E(T+1))}{H_E(T+1) \, T_E(T+1) \, (2 - 3H_E(T+1))^2}.
$$

The total selection over any one generation is the product of its components; W=V.F. The maximum likelihood estimate **of** W is

$$
\hat{\mathbf{W}} = \hat{\mathbf{V}} \cdot \hat{\mathbf{F}}, \text{ or } \hat{\mathbf{W}}(\mathbf{T}) = \frac{2\mathbf{H}_{\mathbf{E}}(\mathbf{T} + 1) (1 - \mathbf{H}_{\mathbf{E}}(\mathbf{T}))}{\mathbf{H}_{\mathbf{E}}(\mathbf{T}) (2 - 3\mathbf{H}_{\mathbf{E}}(\mathbf{T} + 1))}
$$

The overall selection depends only on the genotypic frequencies in egg samples one generation apart, as we should expect, since W measures the selection over an entire generation. In large samples the variance of  $W(T)$  is approximately  $\text{Var } \hat{W}(T) = \hat{F}(T)^2 \text{Var } \hat{V}(T) + \hat{V}(T)^2 \text{Var } \hat{F}(T)$ . The formulas for  $\hat{W}$  and its variance in terms only of the frequencies in egg samples one generation apart were given by **ANDERSON** (1969), who utilized Monte Carlo simulations to demonstrate that they are reliable estimators in samples of at least 200 individuals.

The adults were allowed to continue mating and to lay eggs for nine days after they were censused and the egg samples taken. We should like to know if there was differential mortality between the two adult genotypes during this time. The adult viability can be determined according to the following model.

Genotypes 
$$
Sb/+
$$
  $+$ /+  
Observed frequencies among  
adults at reproduction  $H_A(T)$   $N_A(T)$ 

Adult viability 1

$$
\mathrm{M}(\mathrm{T})
$$

Observed frequencies among adults at discard **HD** $(\text{T})$ 

 $N_{\rm p}(T)$ 

 $\mathrm{H_{p}(T)}$ The adult viability is  $M(T) = \frac{H_D(T) (1 - H_A(T))}{H_A(T) (1 - H_D(T))}$ . Since the genotypic fre-

quencies among the adults at reproduction and at discard are known exactly from total counts of the populations, the true values of M are known and no estimates of the variances are necessary.

In order to visualize the course of selection in the two populations, the frequencies of *Sb* heterozygotes in the adults hatching from egg samples and in the adults at reproduction are graphed in Figure 2. The irequencies of the heterozygotes drop and seem to level off in both populations. In population 1 the frequency abruptly plunges in the last egg sample.

The total numbers of adults at the beginning of egg laying are pictured in Figure *3.* The size of the populations varied from about 300 to 600 individuals. The correlation between the numbers in the two populations is 0.83, which is very high indeed. Clearly the numbers of adults in these two populations have undergone very similar variations, most likely due to common environmental fluctuations.

The results of the analysis are given in Table 3. The successive estimates of the overall selective values, the W's, are correlated **(ANDERSON** 1969), and it may be shown that the correlation coefficient is roughly  $-0.5$ . Since the frequencies of the genotypes in the initial batches of eggs are known exactly, no variances are given for the egg-to-adult viabilities in generation 1. The adult viabilities are in general close to one, and they indicate that little differential mortality occurred between the adult genotypes during the nine days they were permitted to lay eggs. In fact, the effect of adult viability was even less than the values in Table *3*  indicate. Although the adults were allowed to lay eggs for nine days, only eggs deposited during the first two or three days had much chance to develop into the adults included in the next generation. PECK and **RITTER** (unpublished) found that  $97.4 \pm 1.0\%$  of all adults in any generation developed from eggs laid during the first two days; the mortality of adults in these first two days was negligible.

The estimates of the egg-to-adult viabilities and of the fertilities of the *Sb*  heterozygotes are plotted for each generation in Figure **4.** The estimates vary rather differently in the two populations. The standard errors of the estimates **of**  the viabilities and of the fertilities are large compared with the sizes of the estimates themselves, as shown in Table *3.* The estimates seem to vary widely, but we should like to know whether this variation is real or whether it is simply a reflection of sampling error. The estimates of viability in each population are statistically independent, as are the estimates of fertility. Successive estimates of these components of selection will utilize a common set of genotype frequencies among the adults, but these frequencies are known without sampling error, since



FIGURE 2.--Frequencies of Stubble heterozygotes. Upper, Population 1; Lower, Population 2.

**the entire populations were censused in each case. We may then test the** hy**pothesis that our estimates for the viabilities and for the fertilities depart from common true values by sampling error alone. We assume that the estimates are normally distributed. The information for each estimate**  $(I_v(T)$  **and**  $I_F(T)$ **) is the reciprocal of its variance.** 

## **TABLE 3**

Gener-	Egg-to-Adult Viability + $S.E.$			Fertility $S, E$ . $\div$		Overall Selective Value, $W + S.E.$		
stion	Pop. $1$	Pop. $2$	Pop. $1$	Pop. 2	Pop. 1	Pop. 2		Pop. 1 Pop. 2
ı	1,21	, 92	$1.26 + .35$	$2.05 + .68$	$1.52 + .42$	1,89+.63		
$\mathbf{2}$	$-86 + .09$	$-64 + 07$	$1.28 + .24$	$,99+,15$	$1.10 + .24$	$.63 + .12$	-99	.97
з	$-79 + .09$	$1,00+,11$	$,86+,12$	$.92 + .13$	$,68+,12$	$-93 + 17$	.94	.91
4	$.97 + .11$	$1.44 + .16$	$-90 + 13$	$.90+.14$	$.87 + .16$	1,30+.25	1,00	,90
5A	$,81+,10$	$1,06+.13$	$.82 + .12$	$2.03 + .31$	$,66+13$	$2, 16 + 42$	1,20	1,01
5B			$,90+,14$	$.95 + .14$			,86	,89
6	$1.16 + .16$	$,93+,11$	$.73 + .10$	$.91-.13$	$.84 + .17$	$,85+,16$	,88	,87
7					$1.68 + .32$	$1.50 + .27$		
8	$1, 36 + .16$	$.87 + .09$	$1.12 + .14$	$.75-.11$	$1,52+,27$	$,66+,12$	.49	.89
9	$-83 + 09$	$1.09 + .14$	$1.27 + .21$	$-69 - 12$	$1.06 + .21$	$.76 + .16$	1,05	.97
10	$-67 + .09$	$1.20 + 19$	$.96 + .14$	$.88 + .14$	$.65 + .13$	$1.06 + .24$	1.05	.89
11	1,06+.14	$.81-.12$	$.81 + .13$	$.90+.16$	$,86+,18$	$.73 + .17$	96ء	.96
12	$1.52 + .23$	$1,61-.28$	$1,07+.16$	$,78+,15$	1,62+,35	$1, 26 + 33$	.90	,83
13	$1.08 + .15$	$1,32 + .24$	$.12-.04$	$.81-.15$	$,13+,04$	1,08+.28	.94	.92

*Estimates of the egg-to-adult viability, fertility, overall selective value, and adult viability of* **Sb**  *heterozygotes in each generation. Also given are the standard errors (S.E.) estimated from the sampling variances* 



**FIGURE 3.-Total numbers of adults.** 



FIGURE 4. - Viabilities and fertilities of Stubble heterozygotes; their standard errors are given in Table 3 and discussed in the text. Upper, Population 1; Lower, Population 2.

The estimates of the common values are

$$
\widetilde{V} = \frac{\sum\limits_{\mathbf{r}} \hat{V}(\mathbf{T}) \mathbf{I}_{\mathbf{r}}(\mathbf{T})}{\sum\limits_{\mathbf{r}} \mathbf{I}_{\mathbf{v}}(\mathbf{T})} \text{ and } \widetilde{F} = \frac{\sum\limits_{\mathbf{r}} \hat{F}(\mathbf{T}) \mathbf{I}_{\mathbf{r}}(\mathbf{T})}{\sum\limits_{\mathbf{r}} \mathbf{I}_{\mathbf{r}}(\mathbf{T})}
$$

The weighted sums of squares about these weighted means are distributed as  $x^2$ 

#### **TABLE** *4*

*Results of tests for homogeneity of the estimates of viability and fertility in different generations* 



\* **Statistically significant at** *.05* **level.** 

\*\* **Statistically significant at** *.005* **level.** 

with N-1 degrees of freedom, where N is the number of independent estimates of each parameter. Thus

$$
\begin{array}{l} \displaystyle \frac{\Sigma}{\Gamma} \, \left( \hat{\mathrm{V}}(\mathrm{T}) - \widetilde{\mathrm{V}} \right)^2 \, \mathrm{I}_\mathrm{v}(\mathrm{T}) = x^2_{\,\mathrm{K-1}} \, \, \mathrm{and} \\[0.2cm] \displaystyle \frac{\Sigma}{\Gamma} \, \left( \hat{\mathrm{F}}(\mathrm{T}) - \widetilde{\mathrm{F}} \right)^2 \, \mathrm{I}_\mathrm{F}(\mathrm{T}) = x^2_{\,\mathrm{L-1}} \, . \end{array}
$$

These chi-squares provide a test for homogeneity of the estimates. The chi-squares from our data are listed in Table **4,** and in all cases they are too large to be accounted for by sampling error alone. Selection has not been constant but has significantly fluctuated during the course of thirteen generations.

The correlations between estimates of viability and of fertility in each generation were small and far from statistical significance: in population 1,  $r_{10} = .024$ ; and in population 2,  $r_{10} = .020$ . The estimates of the overall selective values in the various generations are graphed in Figure 5; again, the standard errors (given in Table 3) are large. The W's in the two populations at each generation are not significantly correlated. The selective values in each population show very little correlation with the frequencies of the *Sb* heterozygotes. They are weakly correlated with the number of adults at reproduction, but the correlations are **in**  opposite directions in the two populations: in population 1,  $r_{10} = .17$ ; in population 2,  $r_{10} = -.31$ .

## **DISCUSSION**

It is clear that both viability and fertility have varied significantly in each **of**  the two populations. The curves which show the frequencies of *Sb* heterozygotes as functions of time (Figure 2) seem at first glance not too irregular. They are not too different from the gene frequency curves reported by other workers and analyzed under the assumption of constant selection (e.g., CHUNG 1967). The techniques we have used, however, permit a closer look at the operation of selection and reveal something of the complexity of the process.

Both the viability and fertility components of selection are important, and neither factor has predominated in determining the changes in gene frequency. The lack of correlation between the estimates of viability and fertility in the various generations emphasizes the importance of investigating *both* factors. It is not possible to gain an accurate picture of the overall selection by studying viability alone, although this procedure is sometimes the only technically feasible one. The estimates of viability and of fertility are sometimes heterotic, and some-





times not; there appears to be no consistent pattern. The selection has fluctuated in a rather irregular way from generation to generation.

II IV I<sup>I</sup> i<sup>1</sup> Paior B.

Our attempts to relate the selective values to gene frequencies, genotype frequencies, **or** to population densities gave negative results. The numbers of adults in the two populations were very highly correlated, even though the densities fluctuated widely. We may assume that many environmental factors-at least those affecting population numbers—were similar in the replicate populations. Yet the selections in the two populations were different. Thus, a *regular* variation in the selective values with major environmental factors seems unlikely. Part of the variation in selection may be an erratic fluctuation with the multitude of micro-environmental factors, such as the microbial floras. Another source of the fluctuations may be changes in allele frequencies at other loci and the formation and breaking up **of** different complexes of linked genes. The importance of the genetic background and of polygenic complexes associated with mutant alleles to the course of selection has been demonstrated by POLIVANOV (1964).

The overall selection in natural or experimental populations is undoubtedly a combination of these several types of selection. Erratic fluctuations are probably superimposed on any directional changes in gene frequency. Which form of selection predominates will depend on the particular circumstances in the population being considered. Since we have shown that selection can fluctuate rather widely under the fairly uniform environment in the laboratory, then it must surely do so in nature, where the environment is continually changing. In natural populations the selective values are probably in constant flux, creating a concomitant **flux** in gene frequencies. The boundaries of this **flux** will be determined by the variation in the environment and by the relative strengths of the various types of selection which are present.

Fluctuations in selective values also affect the long-term fate of genes. Very high or very low selective values have a great influence on the long-term change in gene frequencies; the overall selection is closer to the geometric mean of the selective values in the various generations than to the arithmetic mean. In small populations this effect is even greater. **HALDANE** and **JAYAKAR** (1962) showed that two alleles at an autosomal locus will be maintained in a large population with fluctuating selection if the geometric mean of the selective values for each homozygote, relative to **the** heterozygote, is less than one. These conditions can be met in several ways, and consistent superiority **of** the heterozygotes is not required. The arithmetic mean of the selective values in the homozygotes may be larger than the average selective value for the heterozygotes; yet if the selective values fluctuate widely in the homozygotes but less so in the heterozygotes, then conditions will be favorable for retention of the polymorphism.

We are grateful to DRS. TH. DOBZHANSKY and W. WATT for their helpful comments on the **paper.** 

#### **SUMMARY**

The frequencies of heterozygotes for the *Sb* allele were followed among the zygotes and among the adults at the time of reproduction in two experimental populations of *Drosophila melanogaster.* These data were analyzed by techniques which partition the selcction in each generation into components of viability and of fertility and which assess the reliability of the estimates **of** the components. Both the viability and the fertility in the two populations fluctuated significantly. Attempts to relate the selective values in the thirteen generations with gene frequencies, genotype frequencies, or population densities were unsuccessful. Viability and fertility were about equally important in determining the changes in gene frequency; the sizes of these two components of selection were not significantly correlated. Possible causes of the variable selection are considered, and the importance of fluctuations in selective values is discussed.

#### **LITERATURE CITED**

- **ANDERSON, W. W., 1969 Selection in experimental population. I. Lethal genes. Genetics** *42:*  **653-672.**
- **CHUNG,** Y. J., 1967 **Persistence of a mutant gene in populations of different genetic background. Genetics 57: 957-967.**
- HALDANE, J. B. S. and S. D. JAYAKAR, 1962 Polymorphism due to selection of varying direction. **J. Genet. 58: 237-242.**
- **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.
- POLIVANOV, S., 1964 Selection in experimental populations of *Drosophila melanogaster* with different genetic backgrounds. Genetics 50: 81-100.
- **SPASSKY, B.,** 1943 Cream of wheat-molasses **fly** medium. Drosophila information Service **17:** 67.
- TEISSIER, M. G., 1942 Persistance d'un gène léthal dans une population de Drosophiles. Compt. Rend. 214: 327-330. -, 1954 Sélection naturelle et fluctuation génétique. Compt. Rend. **238:** 1929-1931.
- Experimental determination of fitness interactions in *Drosophila melanogaster*  **WILSON, J.,** 1968 by the method of marginal populations. Genetics 59: 501-511.