STABILIZING SELECTION FOR PUPA WEIGHT IN *TRIBOLIUM CASTANEUM*[1]

PAMELA K. KAUFMAN, FRANKLIN D. ENFIELD AND RALPH E. COMSTOCK

Department of Genetics and Cell Biology, University of Minnesota, St. Paul, Minnesota 55108

Manuscript received February 14, 1977
Revised copy received May 27, 1977

ABSTRACT

Ninety-five generations of stabilizing selection for pupa weight in *Tribolium castaneum* resulted in a significant decrease in phenotypic variance, moderate reductions in additive genetic variance, but only slight changes in heritability for the trait. Sterility was significantly lower and the average number of live progeny per fertile mating was significantly higher in populations where stabilizing selection was practiced as compared with random selected populations. The results indicate that more genetic variability is being maintained than would be expected unless a fraction of the genes have a heterozygote advantage on the fitness scale. The reduction in phenotypic variance indicated that the populations with stablizing selection became somewhat more buffered against environmental sources of variation over the course of the experiment.

FOR many metric traits, natural selection appears to maintain an intermediate optimum. The evidence for this is indirect, coming largely from long-term directional selection experiments in the laboratory. Directional selection for a metric trait where population means are changed to extreme values beyond what is usually observed in nature often results in a decline in mean for many fitness components. Relaxation of selection usually results in a return toward the original mean for the metric trait in these cases. (Experiments of REEVE and ROBERTSON 1953; DAWSON 1965; RATHIE and BARKER 1968; and ENFIELD 1977 are among many that serve as examples.)

Genes affecting a metric trait may affect that trait alone, or may be pleiotropic and affect both the metric trait and one or more components of fitness. If we accept the premise that natural selection maintains an intermediate optimum, there is then a need to ascertain the mode of action of genes responsible for the considerable genetic variation that is often observed for the metric trait. Various models that have been proposed for the effects of natural selection for an intermediate optimum lead to different expectations concerning the amount and nature of genetic variation that should persist in a population. LERNER (1954) postulated that variability of a metric trait may persist in the face of natural selection because of a fitness advantage of heterozygotes in the case of genes

[1] Supported by National Science Foundation grant G-1238 and Public Health Service grant GM-16074.

affecting both fitness and the metric trait. BULMER (1973) provided the theoretical framework for considering the simultaneous selection for an intermediate phenotype combined with a small heterozygote advantage on the fitness scale. WRIGHT (1935) and ROBERTSON (1956) have shown that selection for an intermediate optimum for a metric trait should lead to fixation for those genes whose only influence on fitness is through the metric trait itself. Genes of this kind would not then be expected to be segregating, except at low frequencies due to mutation, after a long evolutionary history. If, however, there is a range of intermediate phenotypes that are acceptable in nature, genetic variability may be maintained at some intermediate level because the genes for the metric trait are selectively neutral within this range. WADDINGTON (1957) has proposed that selection for a phenotypic intermediate may be selecting for modifier genes that increase the buffering capacity of the genotype. In this model he suggests that either environmental variance or genetic variance may be reduced without altering allelic frequencies for other than the modifier genes. The WADDINGTON model would require that genotype-environment interactions are significant components of variability for the metric trait.

In their simplest form, with the effects of recurrent mutation ignored, the LERNER, WRIGHT-ROBERTSON, and WADDINGTON models lead to quite different expectations if artificial stabilizing selection for a metric trait is practiced. With the LERNER model of genetic homeostasis, additive genetic variance for the metric trait will be maintained even though additive genetic variance for fitness for these genes should approach zero as the optimum is approached. Thus, artificial stabilizing selection for a metric trait would have no effect on either additive genetic variance or environment variance for a population at its optimum. If the base population in which artificial selection is practiced is not at its optimum, additive genetic variance for the metric trait may change, but its direction could not be predicted. Under the WRIGHT-ROBERTSON model, additive genetic variance for the metric trait should be reduced with artificial stabilizing selection. Phenotypic variance will be reduced only to the extent that genetic variance is reduced. Finally, the WADDINGTON model allows for a reduction in both the genetic and environmental components of variance. The selection of modifier genes can either reduce the effects of differences in the existing environment or neutralize the effects of different genotypes on the metric trait.

No consistent pattern of results has emerged from artificial stabilizing selection experiments to lend strong support to any of the proposed models. In most cases the experimental results have suffered from either (a) effective population sizes being so small that inbreeding and drift have confounded the results, thus preventing critical discrimination between alternative hypotheses or (b) if individual gene effects on the metric trait are small, the experiments were terminated before the effects of selection on variability could adequately be measured, given the large sampling errors usually associated with variance parameter estimates. This problem is not unique to stabilizing selection experiments, but as pointed out by LEWONTIN (1974), is characteristic of much of the literature in experimental evolutionary biology.

In 1963, a series of selection experiments was initiated to investigate the nature of the genetic variation for pupa weight in *Tribolium casteneum* (ENFIELD, COMSTOCK and BRASKERUD 1966). These experiments have included two populations in which stabilizing selection was practiced for 95 generations. These populations and lines derived from them provide the data for this paper. The discussion of these data will be focused on the following questions: (1) what is the effect of artificial stabilizing selection on variability for pupa weight? and (2) does artificial selection for intermediate pupa weight affect the population mean for either pupa weight or measures of reproductive fitness? The uniqueness of our data rests on the number of generations of selection and the large effective population sizes that were maintained, when compared with other stabilizing selection experiments.

MATERIALS AND METHODS

1. *Experimental procedures*

The foundation stock for this experiment consisted of an F_3 population that had been produced from the initial cross of two highly inbred lines of *Tribolium casteneum*. The inbred lines (CSI-5 and CSI-10) were obtained from DR. A. SOKOLOFF. They were derived initially from a synthetic strain produced from the cross of six different wild-type strains marked with sooty. [See LERNER and HO (1961) for a more complete description.] The lines had been brother-sister mated for 38 generations at the time the first F_1 cross was made. The two stablilizing selection populations, which we shall designate C_1 and C_2, were initially considered to serve as control populations for two directional selected populations. During the early generations of the experiment, there was a small but significant decline in the mean of the C populations. We became concerned that stabilizing selection might be changing the genetic mean. As a result, two new populations were initiated from the cross of the same two inbred lines. These two populations, designated R_1 and R_2, have been randomly selected and will serve as the control in measuring the effects of the stabilizing selection. The R populations were established in generation 28. Thus, the total data from the experiment consists of 95 generations in the C populations and 67 generations in the R populations.

Each generation, 36 males and 72 females were selected to serve as parents for the next generation in each population. Each male was mated randomly to two females, except that full-sib matings were avoided. After a mating period of 3–5 days, females were put into individual creamer bottles to lay eggs for 3 days. Twenty-one days from the second day of egg-laying, total progeny counts and total number of pupae were recorded for each full-sib family. When family size was greater than 5 individuals per sex, only 5 of a given sex were randomly selected for weighing. Thus, under ideal conditions the population on which weights were taken would consist of 36 half-sib families with 10 male and 10 female progeny in each family. In the C populations, the most intermediate-weight male and the two most intermediate-weight females in each half-sib family were selected to be parents for the next generation. In actual practice slightly higher numbers were selected in some families to ensure 36 males and 72 females at the time of mating. The critical point of emphasis concerning the selection procedure is that the criterion for selection in the C populations was the *median weight* within a half-sib family. In the R populations, selection was completely random with no attention given to family origin. The two different selection procedures led to different effective population sizes, even though the same number of individuals serve as parents each generation. Effective population size in the C populations averaged about 100 each generation based on calculated inbreeding coefficient data. In the R populations, it was approximately 80.

The experiment was replicated in time with any generation of the second replicate following the corresponding generation of the first by two weeks. A 42-day generation cycle was followed throughout the experiment. Complete pedigree data were available for all populations.

In generation 73 of C_2 (45 for R_2) and generation 74 of C_1 (46 for R_1), samples were taken from both the C and R populations to initiate a new set of populations in which directional selection for heavy pupa weight was practiced. This was done with the idea that if stabilizing selection had been effective in altering the heritability of the trait by differentially affecting genetic and environmental variances, it could be confirmed by a separate experiment using directional selection as a tool for estimating realized heritability. In this experiment, 24 males were each mated to 2 females each generation. Selection for heavy pupa weight was done on an intra-half-sib family basis in each population. Mating and egg laying procedures were the same as for the main experiment with the exception that progeny from only 18 of the 24 possible half-sib families were weighed. The lines in this experiment were designated as follows:

CD: directional selection for heavy pupa weight in a line derived from a stablized selected population

RD: directional selection for heavy pupa weight in a line derived from a random selected population

2. *Statistical analysis*

Data on population means for pupa weight and two measures of reproductive performance, *i.e.*, percent sterility and number of live progeny per fertile mating, were analyzed in two ways. Means of the C and R populations were first compared for the total experiment to determine whether differences existed. When differences were significant, the linear regression of generation means on time was calculated for each population to determine whether the average difference between the selection regimes could be demonstrated to have resulted from time changes.

Heritabilities, additive genetic variances and phenotypic variances were estimated in two ways. The hierarchical breeding structure of the population enabled an analysis of variance in which the sire component of variance provided an estimate of ¼ of the additive genetic variance. The sire component of variance and the corresponding intraclass correlation estimate of heritability in the C populations were adjusted by the method of RAHNEFELD *et al.* (1963), to remove the bias that results from a nonrandom selection of parents contributing to the population. The corrected sire component was obtained using the following expression

$$S' = S[1 - h^2(1 - P'/P)]$$

where S' is the sire component obtained from the analysis, and:

h^2 is the estimated heritability of the trait, P' is the phenotypic variance among selected parents, P is the phenotypic variance among individuals before selection, and S is the unbiased estimate of the sire component without selection.

All analyses of variance were performed on an intra-generation basis separately for the two sexes. Data have been pooled over all generations for both sexes in the final results. Standard errors were calculated empirically, treating each generation estimate as a separate observation of the statistic.

Four kinds of parent-offspring regressions, *i.e.*, sire-male progeny, sire-female progeny, dam-male progeny, and dam-female progeny, were calculated each generation. These results were then pooled over all generations for each population to provide a single heritability estimate from the parent-offspring data. Each type of estimate was given equal weight in the pooled estimate. When it was observed that parent-offspring estimates of heritability were significantly higher than analysis of variance estimates in the C populations, possible sources of bias were investigated. A theoretical investigation of the expectation of the parent-offspring regression in the C populations revealed an important upward bias that is a function of the selection process. This was not expected, since it is commonly assumed that selection *per se* does not bias the parent-offspring regression estimate of heritability. In this case, the bias is a function of the selection having been practiced within families. The theoretical expectations of the parent-offspring regression in this case are presented in detail in RESULTS.

RESULTS

1. *Population means*

A comparison of generation means for the C and R populations is shown in

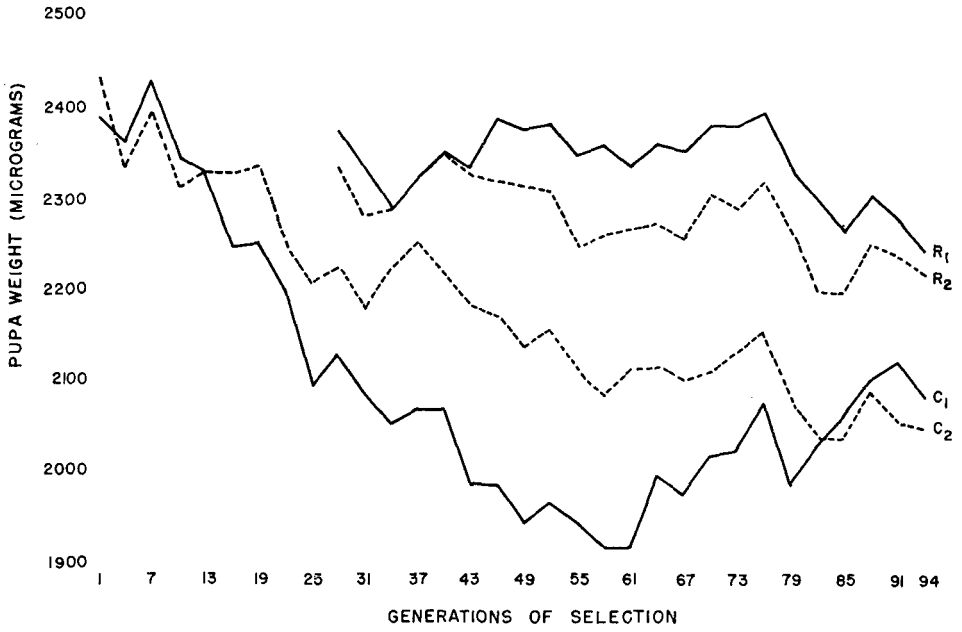

FIGURE 1.—Mean pupa weight (micrograms) for the C and R population.

Figure 1. The points on the graph represent 3-generation means in each case. There was a significantly greater decline in the C than in the R populations when measured by the linear regression of single generation means on time (-3.5 ± 0.3 in the C as compared with -1.3 ± 0.2 in the R populations). A significant quadratic effect was observed in C_1, with the mean reaching its low point by generation 60, followed by a modest rise after that point. The two replicates were very close in mean performance in both the C and R populations at the end of the experiment. While the decline in both the C and R populations was small, it is of interest to evaluate the potential causes for the decline in both, and for the greater decline with stabilizing selection. Selection differentials, inbreeding depression, environmental trends and correlations with reproductive performance all need consideration in interpretation of the results.

Mean selection differentials in the C and R populations were 5.4 ± 0.5 and 1.5 ± 1.4 per generation for the total experiment and thus do not provide an explanation for the observed differences in mean performance. Since the selection differential in both selection systems would normally be expected to be zero, the significantly positive value in the C populations deserves comment. Rather than following a strictly normal distribution, pupa weights are slightly skewed toward lighter weights. Given a slight skewness and the selection procedure followed in the C populations (selection for the median in the half-sib family), a positive selection differential would be expected since the median will be greater than the mean.

Very little reduction in the mean would be expected as a function of inbreed-

ing depression. The initial cross of the two inbred lines resulted in about a 250 microgram heterotic effect when the F_1 was compared with the mean of the two parental inbred lines (ENFIELD, COMSTOCK and BRASKERUD 1966). The average increase in level of inbreeding is about 0.005 per generation which, assuming a linear regression of mean pupa weight on inbreeding depression, would result in an expected decline of about 0.62 micrograms per generation.

A deterioration of the environment as it relates to pupa weight is a potential cause for decline in both populations. The only evidence for an environmental effect comes from comparing the means of the base populations in the C and R populations. The R populations were established from the cross of the same inbred lines as the C populations, but 28 generations later. The base population means were about 50 micrograms lower in the R than the C populations. This difference is not significant and thus is only weak supporting data for environmental deterioration.

A final consideration in comparing the means of the C and R populations is the relationship of pupa weight to reproductive fitness. In the base populations, there was a negative genetic relationship between pupa weight and measures of reproductive performance (ENFIELD, COMSTOCK and BRASKERUD 1966). Thus, it could be argued that the base populations were not at their optimum mean initially, and that natural selection was operating to lower the mean for pupa weight and to increase fitness. Specific data on this point will be discussed in the section on reproductive performance.

2. *Heritabilities, genetic and phenotypic variances*

Table 1 gives the estimates of heritability, additive genetic variance and total phenotypic variance obtained from the pooled analysis of variance for the total experiment for each population. Estimates of heritability for the four types of parent-offspring regression are presented in Table 2. The analysis of variance data are first presented separately by replication to show the consistency between the replicates within a selection system, and then the data are pooled for the two replicates.

The comparison of the pooled analysis of variance estimates for the C and R

TABLE 1

Estimates of phenotypic variance, additive genetic variance and heritability for pupa weight from the analysis of variance

Population	$\hat{\sigma}^2_p$	$\hat{\sigma}^2_g$	$\hat{h}^2$
C_1	34553 ± 554	6762 ± 927	0.195 ± 0.021
C_2	38591 ± 608	8300 ± 812	0.215 ± 0.020
Pooled C	36572 ± 411	7531 ± 616	0.205 ± 0.014
R_1	43399 ± 1150	10742 ± 2124	0.247 ± 0.035
R_2	43981 ± 600	9888 ± 1708	0.224 ± 0.035
Pooled R	43690 ± 648	10315 ± 1363	0.236 ± 0.025
Pooled (R-C)	7118 ± 767	2784 ± 1496	0.031 ± 0.028

TABLE 2

Parent-offspring regression estimates of heritability

Population*	Type of regression: Sire-male offspring	Sire-female offspring	Dam-male offspring	Dam-female offspring	Average
C_1	0.30	0.34	0.40	0.36	0.35 ± 0.020
C_2	0.36	0.28	0.38	0.34	0.34 ± 0.020
Pooled C	0.33	0.31	0.39	0.35	0.345 ± 0.014
R_1	0.20	0.22	0.30	0.28	0.25 ± 0.015
R_2	0.26	0.28	0.30	0.32	0.29 ± 0.020
Pooled R	0.23	0.25	0.30	0.30	0.27 ± 0.012

* Estimates for the C populations are biased upward because of the method of selection employed.

populations shows a significant difference in phenotypic variance, no significant differences in the heritability and a difference in additive genetic variance that approaches significance at the 0.05 level of probability. The lack of significance in the observed difference in additive genetic variance reflects the tremendous amounts of data required when estimating parameters from components of variance. In this case, the estimated additive genetic variance is 27 percent lower in the C and R populations.

A second method of comparing the data from the analysis of variance is the regression of the mean for each parameter on generations of selection. The linear regression of phenotypic variance on generations of selection is significantly negative for the pooled C population data (-118 ± 13), indicating that selection has decreased variability. The regression was not significant in the R populations (25 ± 35). The regression of additive genetic variance on time is also significantly negative in the C populations when the two replicates are pooled (-79 ± 22). The same estimate in the R populations is positive but not significant, with very large sampling errors (99 ± 100). No significant time trends could be detected for heritability estimates in either the C or R populations.

The parent-offspring regression estimates of heritability (Table 2) are higher than the analysis of variance estimates in all populations. The difference was particularly striking in the C populations. The reason for the higher estimates in the C populations was not apparent to us until the expectation of the regression estimate, given the selection system employed, was examined.

When parents are selected on the basis of their phenotype, but without reference to pedigrees, the expectation of the regression of offspring on parents is not altered by the selection. This is because such selection affects the variance among parents and the covariance of parents with offspring in the same degree. However, when intra-family selection is practiced, as in this experiment, the within-family portions of (1) the phenotypic variance among selection parents and (2) the parent-offspring covariance, are altered by the selection, but between-family portions are not. In consequence, the over-all regression is modified by the selection unless, in absence of selection, the within-family and between-family regressions would have been equal.

Given the family structure of this experiment and the fact that selection was practiced within half-sib families in the C populations,

$$\sigma_b^2 = \sigma_s^2 + \frac{1}{n}\sigma_d^2 \cong 0.3875\ \sigma_g^2 + \sigma_{e_b}^2$$

where σ_b^2 is the among half-sib family variance,
σ_s^2 is the sire component of variance,
n is the average number of dams per sire, (1.8 in our data)
σ_d^2 is the dam component of variance,
σ_g^2 is the additive genetic variance, and
$\sigma_{e_b}^2$ is the environmental and non-additive genetic variance among families.

$$\sigma_w^2 = 0.6125\ \sigma_g^2 + \sigma_{e_w}^2$$

where σ_w^2 is the variance within half-sib families
and $\sigma_{e_w}^2$ is the environmental and non-additive genetic variance within families.

$$\sigma_{P_s}^2 = k\ \sigma_w^2 + \sigma_b^2$$

where $\sigma_{P_s}^2$ is the total phenotypic variance among *selected* individuals
and k is a constant determined by the intensity of selection.

It then follows that

$$\sigma_{P_s \cdot O} \cong \frac{1}{2}(0.6125\ k\ \sigma_g^2 + .3875\ \sigma_g^2)$$

where $\sigma_{P_s \cdot O}$ is the covariance between selected parents and their offspring
and

$$E(\hat{h}^2) = \frac{2\ \sigma_{P \cdot O}}{\sigma_{P_s}^2} \cong \frac{(0.6125\ k + 0.3875)\sigma_g^2}{k\ \sigma_w^2 + \sigma_b^2} \qquad (1)$$

where $\hat{h}^2$ is the estimated heritability in a population where parents are selected within families.

Estimates from the analysis of variance of the C population data are as follows:

$\hat{\sigma}_b^2 = 4094$ $\qquad$ $\hat{\sigma}_w^2 = 32478$
$k = 0.299$ $\qquad$ $\hat{\sigma}_g^2 = 7531$
$\qquad$ $\hat{\sigma}_{P_s}^2 = 11530$

Substituting in (1) gives 0.34 as the expectation of the regression estimate of heritability when an unbiased estimate would be 0.22. This is in almost exact agreement with the estimate of 0.345 obtained from pooling the four types of regressions.

In the R populations, the average heritability estimate from the parent-offspring regressions is slightly higher than the analysis of variance estimate. This appears to be a function of a bias due to maternal effects in regressions involving the female parent. In both the C and R populations, the dam-offspring regressions estimates are higher (0.05 and 0.06) than the corresponding sire-offspring estimates. With this bias removed, the two methods of estimating heritability in the R populations are in complete agreement. Again, there was

no evidence for a time trend in the heritability estimates over the course of the experiment in either the C or R populations.

Considering the data in total, the greatest effect of stabilizing selection has been in the reduction of phenotypic variance. The difference between the C and R populations for the total period is highly significant (7118 ± 767), as is the regression of phenotypic variance on generation of selection in the C populations (−118 ± 13). While the data are not totally consistent internally, there is a strong indication that both the genetic and environmental portions of the phenotypic variance were reduced by stabilizing selection. The mean difference in the R and C populations for additive genetic variance approached significance (2784 ± 1496). The highly significant negative regression of additive genetic variance on generations of selection in the C populations (−79 ± 22) provides additional support for the view that the reduction was real. Taking this regression at face value, additive genetic variances in the base populations for the C populations would be estimated to be 11244, slightly higher, but certainly not significantly different from the pooled estimate of additive genetic variance in the R populations. Using the same approach, the amount of additive genetic variance remaining in the C populations after 95 generations of selection would be 3818 ± 1034. This represents a 66 percent reduction for the total period. The fraction of the total variation that is due to environmental variance and nonadditive genetic variance can be obtained from the differences in σ_P^2 and σ_g^2 in the two populations (Table 2). The difference between the R and C populations is highly significant (4338 ± 1680). However, if only the regression of σ_P^2 (−118 ± 13) and σ_g^2 (−79 ± 22) in the C populations are considered, the difference is not significant (−39 ± 26).

One of the potential explanations for the decline in phenotypic variance in the C populations is that it is a scaling effect associated with the decline in the mean. Since both the mean and the variances decline with time in the C populations, our only real test for a scaling effect is to examine the relationship between means and variances in the R populations where the generation-to-generation variation in the phenotypic variance showed no significant time trend, and the decline in the mean over time was slight. The regression of phenotypic variance on generation mean was 0.6 ± 10.8 in the R populations, thus providing no evidence for a change in phenotypic variance as a function of mean changes in the unselected populations. Since we are not dealing with large differences in the means in the C and R populations, it seems unlikely that the changes in the means are a major factor in the reduction in phenotypic variance.

The results of the directional selection experiment in the populations derived from the C and R populations are shown in Figure 2. Estimates of realized heritability were 0.19 ± 0.08 and 0.20 ± 0.08 in C_1D and C_2D and 0.31 ± 0.09 and 0.05 ± 0.09 in R_1D and R_2D. Pooling the replicates gave estimates of 0.19 ± 0.06 in the lines derived from the C populations and 0.18 ± 0.06 in the lines derived from the random bred populations. These estimates are consistent with those obtained from the analysis of variance in the main experiment, and even though the sampling errors are large, they do provide additional support

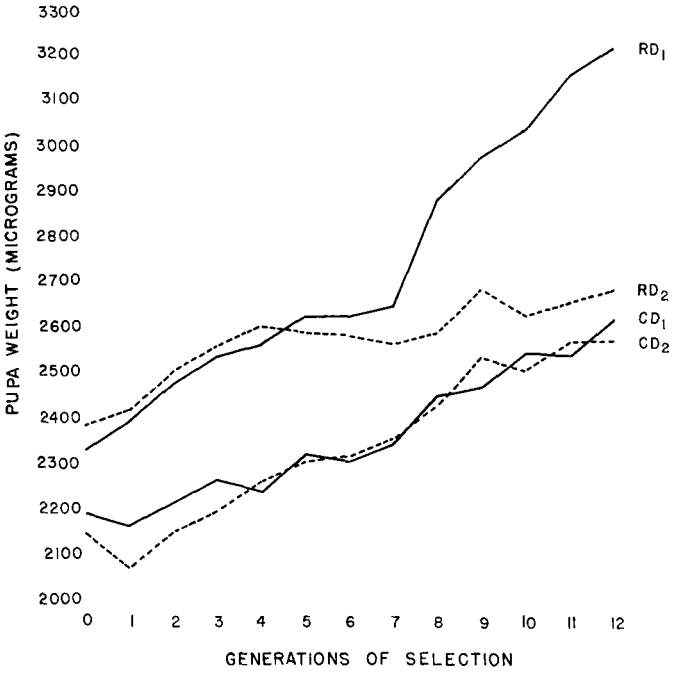

FIGURE 2.—Mean pupa weight (micrograms) in populations derived from the C and R populations.

for the argument that stabilizing selection has done little to alter the heritability for the trait.

3. *Reproductive performance*

Table 3 provides a comparison of the two sets of populations for percent sterility. Percent sterility is measured in this case as the number of females exposed to a male that have no live progeny at 21 days of age, divided by the total number of females exposed. Significant differences exist in the mean per-

TABLE 3

Mean sterility and linear regression of sterility on generations of selection

Population	Number of generations	Mean (percent)	Regression on time
C_1	96	21.8 ± 1.0*	—0.037 ± 0.035*
R_1	68	29.8 ± 1.3	0.132 ± 0.067
C_2	96	16.9 ± 1.1*	0.016 ± 0.039
R_2	68	34.0 ± 1.4	0.080 ± 0.072
Pooled C		19.6 ± 0.7*	—0.012 ± 0.026*
Pooled R		31.8 ± 0.9	0.107 ± 0.049

* Indicates a significant difference from the corresponding R population.

formance. The linear regression coefficients suggest that at least part of the difference is a function of increasing sterility over time in the R populations, which did not occur in the C populations. Since most measures of reproductive performance would be expected to decline with a build up of inbreeding effects, the increase was not surprising in the R populations. It is significant to note that the average inbreeding coefficient in the C populations was nearly 0.5 after 95 generations, but that, if anything, the sterility may have decreased.

Table 4 shows the same kind of analysis for the number of live progeny per fertile mating for a three-day egg-laying period. Once again there is a significant difference favoring the C populations. The observed difference in this case, however, is not as readily associated with time trends. Rather, it may have simply been the case from the outset that the individuals closest to the population mean will tend to have the higher reproductive rate when fertile. A comparison of only those generations where the C and R populations were contempories in time, *i.e.*, elimination of the first 28 generations of the C populations, gives nearly identical results, with the mean for the C populations being 27.8 ± 0.4 rather than 27.9 ± 0.3 (Table 4).

DISCUSSION

There are several additional kinds of data on pupa weight in Tribolium that we have obtained from our directional selection experiments that should be summarized before attempting to relate the results from the stabilizing selection experiment to existing theory. First, directional selection for heavy pupa weight was effective for more than 75 generations in populations derived from the cross of the same inbred lines as those used for the base populations in the C populations (ENFIELD 1977). The number of genes required to obtain the total observed response has been conservatively estimated at 150 (ENFIELD 1973). Second, even though the populations responded to selection for much longer than most directional selection experiments, there is considerable evidence that genetic variability persists in the plateaued populations. Finally, relaxed selection after 52 generations of directional selection has resulted in a steady return of the popu-

TABLE 4

Mean number of progeny per fertile mating and linear regression of progeny number on generations of selection

Population	Mean number of progeny	Regression on time
C_1	28.7 ± 0.4*	—0.037 ± 0.015
R_1	22.1 ± 0.3	—0.008 ± 0.016
C_2	27.1 ± 0.4*	0.053 ± 0.013*
R_2	23.7 ± 0.2	0.008 ± 0.015
Pooled C	27.9 ± 0.3*	0.011 ± 0.011
Pooled R	23.1 ± 0.2	0.000 ± 0.011

* Indicates a significant difference from the corresponding R population.

lations toward the original mean. This regression toward the natural mean has continued for more than 50 generations without as yet leveling out at any point above the original base populations. The mean has declined from approximately 4700 micrograms to 3600 micrograms over this period. This decline amounts to nearly 50 percent of the prior total response to selection.

Next, the nature of the base populations deserves comment. Selection was initiated from the cross of two highly inbred lines, and thus initial allelic frequencies for those genes which have segregated in the C and R populations should have been near 0.5, except for genes that may have kept segregating in the inbred lines because of a heterozygote advantage. Obviously, allele differences between the two inbred lines and any segregation within these lines had to be associated with genes that were segregating in the wild populations which were sampled in the development of the synthetic from which the inbreds were derived. Existing theory concerning the effects of stabilizing selection suggests that those genes may have been of two kinds, those affecting the metric trait that have a heterozygote advantage with respect to fitness, and those of the kind postulated by ROBERTSON and WRIGHT. The latter kind may or may not be segregating depending upon the magnitude of gene effects on the trait, but more importantly, on the range of phenotypic expression that is tolerable for the intermediate optimum. There may be a range of pupa weights within which the genes effecting the trait are selectively neutral. Under this hypothesis, selection may operate to eliminate genetic variability up to a certain point, but beyond that point fixation or loss of variation becomes a function of effective population size.

Our artificial stabilizing selection (1) was more intense than would have occurred in nature and (2) was done in the laboratory under a very narrow range of environmental conditions compared with what is encountered in nature. Because of the intensity of selection, the range of acceptable phenotypes was restricted to a greater extent than by natural selection. Secondly, because of the narrow range of environments in the laboratory, segregation required in nature to allow for a wide range of adaptability may have been eliminated in the laboratory.

Given this background, we have interpreted the experimental results in the following way. First, although the results are not completely conclusive, it appears that there was a reduction in environmental variance from stabilizing selection. The experiment was conducted under constant environmental conditions throughout the total experiment to the extent that such is possible. This, together with the absence of any downward trend of environmental variance in the R populations, strongly suggests that environments did not become more uniform during the course of the experiment. Thus, the conclusion to be drawn is that the genotypes became more adapted to the existing set of environmental conditions. The question can be posed as to why artificial stabilizing selection was effective in reducing environmental variability if this is a phenomenon that should have already been accomplished by natural selection. First, as already mentioned, our environmental conditions in the laboratory span a much narrower range of temperature and humidity conditions, both of which are import-

ant for growth and development, than would be true under natural conditions. Thus, stabilizing selection in the laboratory could be selecting for a much narrower range of adaptability than would be required in nature. Second, the artificial stabilizing selection was much more intense than would occur naturally (as measured by the R populations) and could therefore, contribute to the reduction in variance that was demonstrated.

Another significant demonstration that selection can operate to minimize the effects of variable environmental effects was the work of WADDINGTON and ROBERTSON (1966). Their experiment was specifically designed to select for increased or decreased sensitivity to temperature variation on eye size in Drosophila. In several stabilizing selection experiments where a decrease in phenotypic variance has been noted, at least a fraction of the decline has been attributed to a decline in environmental variance. (GIBSON and BRADLEY 1974; SCHARLOO 1964; and SCHARLOO, HOOGMOED and TER KUILE 1967). On the other hand, PROUT (1962) found no decline in environmental variance with stabilizing selection for developmental time in Drosophila.

We can now focus on the effects of stabilizing selection in changing additive genetic variance and relate these results to expectations under alternative models of gene action. We are assuming that the 95 generations of stabilizing selection led to a reduction, but certainly not an exhaustion, of additive genetic variance for pupa weight. Unfortunately this result is not incompatible with any of the gene action models that have been proposed. Had we been able to demonstrate that stabilizing selection was effective in eliminating most or all of the additive genetic variance, it would have been strong evidence against the heterozygote advantage on the fitness scale. On the other hand, since we know there are a large number of genes that affect pupa weight, each with a relatively small effect, we need not necessarily expect a great reduction in additive genetic variance from the stabilizing selection given the selection intensity and the length of our experiment. There are, however, some aspects of the data to suggest that more variability is being maintained than would be expected if it were not for a heterozygote advantage in fitness.

Given an effective population size of 80 in the R populations, 100 in the C populations and initial allelic frequencies of 0.5 for segregating genes in both populations, we can pose the question of how much reduction in additive genetic variance should occur in each population as a function of drift with no selection. Additive genetic variance can be expressed as

$$2 \sum_i \bar{q}_i \, (1 - \bar{q}_i) [1 + (1 - 2\bar{q}_i) a_i]^2 \, u_i^2$$

where $\bar{q}_i$ is the allelic frequency for the allele for heavier weight at the i^{th} gene, a_i is the measure of level of dominance of that allele, u_i is 1/2 the difference in mean genotypic value of the two homozygous genotypes, and $\sum_i$ indicates summation over all segregating genes affecting pupa weight in the base populations. $\sum_i \bar{q}_i(1 - \bar{q}_i)$ will be maximum at $\bar{q}_i = 0.5$ and will be reduced in advancing generations as a function of drift. In generation t, the expectation of $\bar{q}_{it}(1 - \bar{q}_{it}) = \bar{q}_{io}(1 - \bar{q}_{io}) \left(1 - \frac{1}{2N}\right)^t$ where t is the number of generations of random mating

without selection, $\bar{q}_{i0}$ is the allelic frequency in generation 0 and N is effective population size (ENFIELD *et al.*, 1969). With the assumption of trivial effects due to dominance [average level of dominance was estimated to be $\bar{a} = 0.13$ in an earlier paper (ENFIELD *et al.*, 1969)] the expected percent reduction in additive genetic variance in the R populations as a function of drift after 67 generations, assuming an effective population size of 80, would be 34 percent. In the C populations, with 95 generations and an effective population size of 100, the expected reduction would be 38 percent of the original additive genetic variance in the base population. As a first approximation, these figures translate into a reduction of 0.51 percent per generation in the R population and 0.40 percent in the C populations, assuming linearity over time. We are then confronted with determining the best estimate of additive genetic variance in our base populations. We have chosen to use the mean estimates from the analysis of variance adjusted to a first generation basis by the linear regression of additive genetic variance on time. This gives us a pooled base generation estimate of 10820. Thus, the expected decline per generation as a function of drift would be −55 and −43 in the R and C populations respectively.

The fact that no decline in additive genetic variance could be demonstrated in the R populations suggests that forces were operating to maintain more variability than would be expected under a "no selection" model with drift. The most plausible explanation for this is that at least a fraction of the genes segregating which have an effect on pupa weight have a heterozygote advantage on the fitness scale. The decline in fitness with directional selection, the return toward the original mean with relaxed selection, and the failure to exhaust additive genetic variance with directional selection are independent results, each of which would also be expected under the genetic homeostasis model.

On the other hand, the homeostatic model by itself provides no explanation for the decline in additive genetic variance that occurred in the C populations. The reduction that might occur as the result of shifts in allele frequencies from original values near 0.5 toward equilibrium values in the case of genes with heterozygote advantage for fitness should also have occurred in the R populations. It appears that the difference between the C and R populations with respect to time trend in additive genetic variance must be explained in terms of genes that affect pupa weight, but not fitness except through their effects on the metric trait. To explain segregation of such genes in our base population and in natural populations, we have proposed that in nature there is a range within which any pupa weight is essentially optimum and that the same would have been true for the R populations. The corollary is that a mild reduction in additive genetic variance arising from drift could have occurred in the R populations. In this connection, it should be remembered that the regression estimate for additive variance in the R populations 99 ± 100, does not exclude some decline as the expectation under random selection. In the C populations, the effect of the stabilizing selection practiced would have been to narrow the range of pupa weight that is optimum so that, in accord with the theory by WRIGHT and ROBERTSON, movement toward homozygosity and elimination of additive genetic variance would be promoted by the selection as well as by drift.

The authors are indebted to OVE BRASKERUD for the many years of technical assistance in the collection of data for this experiment. Thanks also to NANCY HARTUNG for her assistance during the preparation of the manuscript.

LITERATURE CITED

BULMER, M. G., 1973 The maintenance of genetic variability of polygenic characters by heterozygous advantage. Gen. Res. **22**: 9–12.

DAWSON, P., 1965 Genetic homeostasis and development rate in Tribolium. Genetics **51**: 873–885.

ENFIELD, F. D., 1973 Long term selection for pupa weight in Tribolium. Gene action and gene numbers. Genetics **74**: s72. ——, 1977 Selection experiments in Tribolium designed to look at gene action issues. Proc. Int. Conference on Quant. Genet. (In press).

ENFIELD, F. D., R. E. COMSTOCK and O. BRASKERUD, 1966 Selection for pupa weight in *Tribolium castaneum*. I. Parameters in base populations. Genetics **54**: 523–533.

ENFIELD, F. D., R. E. COMSTOCK, R. GOODWILL and O. BRASKERUD, 1969 Selection for pupa weight in *Tribolium castaneum*. II. Linkage and level of dominance. Genetics **62**: 849–857.

GIBSON, J. B. and B. P. BRADLEY, 1974 Stabilizing selection in constant and fluctuating environments. Heredity **33**: 293–302.

LERNER, I. M., 1954 *Genetic homeostasis*. Oliver and Boyd, Edinburgh.

LERNER, I. M. and F. K. Ho, 1961 Genotype and competitive ability of Tribolium species. Am. Naturalist **95**: 329–343.

LEWONTIN, R. C., 1974 *The genetic basis of evolutionary change*. Columbia Univ. Press, New York.

PROUT, T., 1962 The effects of stabilizing selection on the time of development in *Drosophila melanogaster*. Genet. Res. **3**: 364–382.

RAHNEFELD, G. W., W. J. BOYLAN, R. E. COMSTOCK and M. SINGH, 1963 Mass selection for post-weaning growth in mice. Genetics **48**: 1567–1583.

RATHIE, K. S. and J. S. F. BARKER, 1968 Effectiveness of regular cycles of intermittent artificial selection for a quantitative character in *Drosophila melanogaster*. Aust. J. Biol. Sci. **21**: 1187–1213.

REEVE, E. C. R. and F. W. ROBERTSON, 1953 Studies in quantitative inheritance. II. Analysis of strains of *Drosophila melanogaster* selected for long wings. J. Genet. **51**: 276–316.

ROBERTSON, A., 1956 The effect of selection against extreme deviants based on deviations or on homozygosis. J. Genet. **54**: 236–248.

SCHARLOO, W., 1964 The effect of disruptive and stabilizing selection on a cubitus interruptus mutant in Drosophila. Genetics **50**: 553–562.

SCHARLOO, W., M. S. HOOGMOED and A. TER KUILE, 1967 Stabilizing and disruptive selection on a mutant character in Drosophila. I. Phenotypic variance and its components. Genetics **56**: 709–726.

WADDINGTON, C. H., 1957 *The strategy of the genes*. George Allen and Unwin Ltd., London.

WADDINGTON, C. H. and E. ROBERTSON, 1966 Selection for developmental canalisation. Genet. Res. **7**: 302–312.

WRIGHT, S., 1935 The analysis of variance and the correlation between relatives with respect to deviations from an optimum. J. Genet. **30**: 234–256.

Corresponding editor: R. C. LEWONTIN