Fitness Sensitivity and the Canalization of Life-History Traits
Author(s): Stephen C. Stearns and Tadeusz J. Kawecki
Published by: Society for the Study of Evolution
Accessed: 18/12/2008 11:18

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at [http://www.jstor.org/page/info/about/policies/terms.jsp](http://www.jstor.org/page/info/about/policies/terms.jsp). JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at [http://www.jstor.org/action/showPublisher?publisherCode=ssevol](http://www.jstor.org/action/showPublisher?publisherCode=ssevol).

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.
FITNESS SENSITIVITY AND THE CANALIZATION OF LIFE-HISTORY TRAITS

STEPHEN C. STEARNS1 AND TADEUSZ J. Kawecki2
Zoology Institute, Rheinsprung 9, CH-4051 Basel, Switzerland

Abstract.—Canalization is an abstract term that describes unknown developmental mechanisms that reduce phenotypic variation. A trait can be canalized against environmental perturbations (e.g., changes in temperature or nutrient quality), or genetic perturbations (e.g., mutations or recombination); this paper is about genetic canalization. Stabilizing selection should improve the canalization of traits, and the degree of canalization should be positively correlated with the traits' impact on fitness. Experiments testing this idea should measure the canalization of a series of traits whose impact on fitness is known or can be inferred, exclude differences among traits in the number of loci and alleles segregating as an explanation for the pattern of variability found, and distinguish between canalization against genetic and environmental variation. These conditions were met by three experiments within which the variation of fitness components among Drosophila melanogaster lines was measured and among which the genetic contribution to the variation among lines was clearly different. The canalization of the traits increased with their impact on fitness and did not depend on the degree of genetic differences among lines. That the flies used had been transformed by a P-element insert suggests that canalization was also effective against novel genetic variation. The results reported here cannot be explained by the classical hypothesis of reduction in the number of loci segregating for traits with greater impact on fitness and confirm that traits with greater impact on fitness are more strongly canalized. This pattern of canalization reveals an underappreciated role for development in microevolution. There is differential genetic canalization of fitness components in D. melanogaster.

Key words.—Canalization, Drosophila, fitness sensitivity, natural selection.

Received May 12, 1993. Accepted April 19, 1994.

For a trait under stabilizing selection, deviations from the optimum by definition reduce fitness. Any developmental mechanism that canalizes the trait by constraining it to be closer to the optimum should be favored by natural selection (Rendel 1967). Such canalization should be more important for traits under strong stabilizing selection than for traits under weak selection. Therefore, traits more important to fitness should be more strongly canalized than traits less important to fitness. The phenotype can be perturbed from the optimum by two sorts of factors—environmental and genetic—and there can be canalization against both (Waddington 1953). In the experiments reported here, we perturbed fly genotypes by inserting a P-element plasmid. We present evidence for canalization against genetic perturbations, showing that traits more important to fitness varied less than traits with less impact on fitness.

The genetic variation of a trait can be reduced either by reducing the number of alleles coding for the trait by fixing them or driving them to low frequency, or by reducing the magnitude of their effects without changing their frequency. The former is the classic view: selection reduces genetic variation directly. It assumes that the average magnitude of allelic effect remains constant, while the number of segregating loci decreases. The latter describes genetic canalization. It assumes that the number of segregating loci remains constant while the average magnitude of allelic effect decreases. In reality, phenotypic variation is probably reduced by a mixture of both effects. The negative correlation of the variation of traits with their impact on fitness did not change when we markedly reduced the number of loci segregating, thus ruling out that explanation of the pattern.

Selection for canalization is illustrated in figure 1. In figure 1A, we sketch the frequency distributions in population of residents and of a more canalized mutant. Because the more canalized mutant produces a frequency distribution of the trait that deviates less from the optimum fitness than that of the residents, it will invade the resident population. Thus, when the deviation of a trait from an intermediate value reduces fitness,
selection should favor an increase in that trait's canalization.

This process will improve canalization, but only up to a point, for the mechanisms that produce canalization should cost something, should be involved in trade-offs, and should themselves be susceptible to erosion by mutations. When the balance between costs and benefits is reached, selection for further canalization will cease. The point at which that occurs will differ among traits and will depend on the strength of selection for canalization, that is, on the sensitivity of fitness to changes in the traits, depicted by the breadth and flatness of their fitness profiles (fig. 1B).

Genetic canalization buffers the phenotype against allelic effects and is thought to evolve through the invasion of modifiers that mask the expression of alleles causing deviations from the optimal trait value (Rendel 1967). Some of the modifiers may be specific, masking the effects of particular alleles that cause deviations from the optimum, but some of the modifiers may be general, effective against many mutations, even ones that have not yet occurred, and possibly against environmental perturbations as well.

Although, on the one hand, genetic and environmental perturbations of the phenotype need to be distinguished, we also note, for purposes of clarity, that genetic canalization and phenotypic plasticity are not mutually exclusive: they combine to form canalized reaction norms (fig. 2). The reaction norm can be genetically canalized (fig. 2A) or uncanalized (fig. 2B); it can be genetically canalized in one environment but not in another (fig. 2C). Failure to appreciate any of these distinctions could lead to misinterpretation.

In these experiments, we perturbed fly genotypes with a P-element plasmid insert in three genetic backgrounds that were inbred (expt. I), then six genetic backgrounds that were even more heterozygous than wild flies (expt. II), and finally a single outbred genetic background into which the P-element was inserted at six different positions (expt. III). The flies transformed with the treatment insert carried an additional gene for elongation factor. Control lines had been similarly manipulated but lacked the extra gene for elongation factor. Elongation factor is present in all organisms and as a gene family in eukaryotes. It docks the incoming amino-acyl-tRNA correctly on the ribosome and catalyzes the transition in which one amino acid is added to the peptide (Darnell et al. 1990). Thus, it is necessary for protein synthesis. In *Drosophila melanogaster* there are two genes for elongation factor. F2 is translated mostly in pupae. F1 is a house-keeping gene needed in all growing cells and is present in only one copy in normal cells (Hovemann et al. 1988). The synthesis of EF-1α protein sharply decreases with age in *Drosophila* and precedes the decrease in total protein synthesis by a few days (Webster and Webster 1983).

An experiment designed to test the hypothesis that the components of fitness are differentially canalized in proportion to their impact on fitness should do at least three things. First, it should...
measure the canalization (or its converse, the variability) of a series of traits whose impact on fitness is known or can be inferred. Second, it should exclude the alternative—variation in the number of alleles segregating rather than in the magnitude of their effects—as an explanation for the pattern of variability found. Third, it should distinguish between canalization against genetic variation and canalization against environmental variation.

The first condition was met by measuring life history traits—age at eclosion, dry weight at eclosion, early fecundity, late fecundity, and life span—and calculating the sensitivity of standard measures of demographic fitness ($R_0$, $r$) for a set of mortality regimes that span the range of conditions that probably have been encountered by *D. melanogaster* in the field and laboratory (Rieseberg and Shorrocks 1987). The degree of canalization can only be measured in laboratory experiments, but the laboratory is no place to measure the correlations of traits with fitness, for it is the relationship between traits and fitness in the environment in which canalization evolved that is relevant. Therefore, we included in our set of mortality regimes several that represent natural environments.

The second condition—ruling out differences among traits in amount of allelic variability present before the perturbation as an explanation of the negative correlation between impact on fitness and amount of variation—was met by measuring variation once across lines among which there was substantial variation in the number of loci segregating (expt. II), then again across similar lines among which there was much less allelic variation (expt. III). The negative correlation between the sensitivity of fitness to changes in a trait and the degree of canalization of the trait was equally strong in both sets of lines and did not change appreciably from the genetically more variable to the genetically more uniform set. If the amount of allelic variation (at all loci) among lines is changed and the correlation between variation in traits and impact on fitness is measured, and that correlation does not change when the allelic variation among lines is changed, then the correlation must be due, at least in part, to differential canalization of the traits. The correlation cannot be solely caused by differences in the number of alleles segregating for the traits.

The third condition—distinguishing between canalization against genetic variation and canalization against environmental variation—was met in three ways. First, we measured the variation in all traits of line means about the mean value across lines within environments. Second, sample sizes for the line means were large, and every effort was made to minimize environmental variation (temperature, food quality, and lar-
val and adult density were all standardized). Thus, the variation among lines within environments should be primarily genetic, not environmental, and differences among traits in degree of variation should reflect canalization primarily against genetic variation. Whether these flies would have a similar pattern of canalization against environmental variation cannot be determined from such a comparison. Third, we calculated the measurements at two temperatures and compared the results. If the traits were canalized better at one temperature than the other, we should have noted differences in the patterns of variation at the two temperatures. We did not.

The experiments reported here were more complex than would be necessary to test this idea, for they were performed for other reasons (cf. Stearns et al. 1993; Stearns and Kaiser 1993). That the flies used had been transformed by a P-element insert makes the point that the canalization mechanisms suggested by our results were effective against novel genetic variation, but the use of genetically transformed flies is not central to the logic of this paper. Because the experiments were done for other purposes, we did not have a set of control lines with no inserts. The results reported here rule out, for these data, the hypothesis that differential amounts of allelic variation account for the pattern of variation in fitness components and confirm that traits more strongly correlated with fitness are more strongly canalized.

**MATERIALS AND METHODS**

In three experiments of similar design, we measured the canalization of five fitness components: age at eclosion, dry weight at eclosion, life span, early fecundity, and late fecundity. Their degree of genetic canalization was estimated from the variation among replicate-line means. Because all measurements were made at two temperatures, we could also test whether the genetic canalization of the traits differed between the two temperatures. The logic of this paper relies on the answers to three questions: (a) Did the traits differ in how much they varied among lines, and was the degree of variation correlated with impact on fitness? (b) If so, did that pattern change among the three experiments? (c) Did it change between the two temperatures? The materials and methods are organized to provide the relevant information for each of those questions.

The experiments were planned originally to measure the impact on fitness components of inserting an additional copy of the gene for elongation factor, EF-1α, into the *Drosophila* genome. The nature of the genetic transformation required us to make measurements both for C-inserts (controls without the additional gene for elongation factor) and for T-inserts (treatments with the additional gene for elongation factor) and at both 25°C and at 29.5°C. Simpler experiments could have measured the degree of canalization of a series of fitness components.

**Differences among Treatments**

T-lines consisted of *D. melanogaster* transformed with a P-element containing the ry+ marker plus the F1 copy of the gene for elongation factor EF-1α, flanked by initiation and termination sequences for heat-shock protein (Shepherd et al. 1989). C-lines had been similarly manipulated but lacked the gene for elongation factor; their insert was shorter and at a different position in the *Drosophila* genome (Stearns and Kaiser 1993). The map positions of the inserts on the chromosomes was determined for the lines used in experiment I, in experiments II and III, all inserts were on chromosome III, but their map position on that chromosome was not determined.

**Differences among Experiments**

Experiment I was done with inbred flies. Three T-and three C-lines (provided by W. Gehring) were first maintained on our standard medium at 25°C for four generations. In experiment I, all flies were held at 25°C until eclosion; adults were tested either at 25°C or at 29.5°C. Longevity and fecundity were measured with 40 flies per line, age and weight at eclosion with 24 flies per line.

Experiment II was done on outbred flies with six different genetic backgrounds. They were even more heterozygous than wild flies. The T-lines and C-lines were obtained from a jumpstart cross in which one arranges for P-elements already inserted in the genome to jump to new insert positions (Cooley et al. 1988; Robertson et al. 1988; Bellen et al. 1989). The use of balancers ensured that the only surviving flies carried the plasmid at some position on the third chromosome. After one fly with the desired marker combinations was crossed again with the balancer stock, a brother-sister cross between flies with the correct markers yielded flies homozygous for the plasmid construct.

The test lines in experiment II were obtained by outcrossing the one inbred T-and one inbred C-line with six different inbred lab stocks to yield six pairs of outbred T-and C-lines with the ge-
netic backgrounds to be tested. Genetic backgrounds were provided by isofemale lines held in the laboratory since 1987 and maintained by full-sib mating for about 60 generations. In experiments II and III, larvae and pupae were reared at the same temperature as the adults.

Experiment III was done on outbred flies with one genetic background, that of line 5 in experiment II. The test flies were obtained by outcrossing six inbred T-and six inbred C-lines from the jumpstart cross with one inbred lab stock to yield six heterozygous treatment and six heterozygous control lines with the insert positions to be tested. All inserts were on the third chromosome, but the C-inserts were not in the same positions on the third chromosome as the T-inserts.

Fly Culture

All measurements were made at 25°C and at 29.5°C (on different flies at each temperature). Flies were reared from first instar larvae gathered within 8 h of hatching and then kept at standard density. Ages were calculated from the midpoint of egg laying, which lasted 4 h. When experiments had to be run in sequence rather than parallel, corresponding treatment and control lines were run together, so that differences among sequences were absorbed into the effects of replicates (individuals or vials) within lines. Here we discuss only the traits of mated females, for male fitness is difficult to measure. To measure the life span of mated females, we established 4 vials per line in experiment I and 10 vials per line in experiments II and III with 10 2-day-old virgin females and 15 2-day-old males. Three times a week the flies were transferred to new vials, and the number of dead flies was recorded until the last fly died. Males were replaced by younger males (about 1 wk adult age) from the same line if there were fewer males than females in the vial. Females had access to appropriate substrate for oviposition and did lay eggs.

To measure developmental time and dry weight at eclosion, we placed 12 larvae in each of 10 vials per line. When all larvae had pupated, the vials were placed in an eclosion fractionator (Stearns et al. 1987) that collected freshly eclosed flies at 6-h intervals. After sex was determined, the flies were dried at 50°C for 3 h and weighed to 0.01 mg.

To measure lifetime egg production, we established 30 vials per line with one female and two males per vial. The laying surface was replaced daily by a new one with a drop of fresh yeast. Dead males were removed and replaced by a young virgin male from the same line. The number of eggs laid by each female in 24 h was counted daily until the last female died. For this analysis, fecundity early in life was defined as the number of eggs laid per day from the fourth to the fourteenth day after eclosion. This was 13–23 days after birth at 25°C and 11–21 days after birth at 29°C. Fecundity late in life was defined as the number of eggs laid from the fifteenth to the fifth day before the death of the last female. This was 32–42 days after birth at 25°C and 23–33 days after birth at 29.5°C. If a female died, her fecundity was treated as a missing value, not as zero.

Sample Sizes

Experiment I.—For life span, 40 individuals from each line were tested at each temperature: 3 backgrounds × 2 treatments × 2 temperatures = 480 mated females. For fecundity, 40 mated females from each line, temperature, and treatment were tested (480 total). For age and weight at eclosion, 24 larvae from each line were tested at each temperature (288 total).

Experiments II and III.—For life span, 100 individuals for each line were tested at each temperature: 6 backgrounds × 2 treatments × 2 temperatures = 2400 mated females. For fecundity, 30 mated females from each line, temperature, and treatment were tested (720 total). For age and weight at eclosion, 120 larvae from each line were tested at each temperature (2880 total, of which about 2400 survived to eclosion).

Comparison of Experimental Designs

The major differences among experiments I, II, and III were these: (a) experiments II and III were better replicated; there were three C-lines and three T-lines in experiment I, six C-lines and six T-lines in experiments II and III. (b) Genetic backgrounds differed to an unknown but probably small degree in every line in experiment I; there were six backgrounds with one T and one C line per background in experiment II, and just one background in experiment III. (c) Insert positions differed in every line in experiment I with some on chromosome II and some on chromosome III; there was one T position and one C position in experiment II, both on chromosome III; there were six T positions and six C positions.
in experiment III, all on chromosome III. (d) The flies were inbred in experiment I and highly outbred in experiments II and III. (e) The larvae and pupae were reared at 25°C in experiment I, 25°C and 29.5°C in experiments II and III. (f) Sample sizes were small in experiment I and large in experiments II and III.

**Statistics**

**Measures of variation.**—For each trait in each experiment for each temperature-by-treatment combination, we calculated two measures of variation among the means of replicate lines, the coefficient of variation and the standardized range. They are not statistically independent. We made the calculations using both to demonstrate that the conclusion did not depend on the measure of variation used. Line means were calculated from vial means for age and size at eclosion and life span, and from individual means for fecundity.

**Problems with the assumption of normality.**—The means and the variances of the CVs and ranges were correlated. Therefore, when we calculated linear regressions of measures of trait variation on sensitivity of fitness to changes in traits, we calculated them with and without log transformations of the measures of variation. It is not clear that the data used in those linear regressions fulfill all the assumptions of that method. Therefore, we also calculated Kruskal-Wallis ANOVAs to check whether the results would change when a nonparametric test was used.

**Effects of experiment, temperature, and treatment on the measures of variation.**—To test whether the patterns of variation were affected by experiment (I, II, or III), treatment (control insert or elongation factor insert), or temperature (25°C or 29.5°C), we did a three-way ANOVA, SAS Type III sums of squares, using the three-way interaction term as the denominator in the F tests. The ANOVA was done on untransformed and on log-transformed CVs and ranges to see whether the transformation made any difference.

**Sensitivity of Fitness to Changes in Traits**

For each trait, we performed numerical analysis to make seven estimates of the percent reduction in fitness caused by a 10% detrimental change in the trait. These seven scenarios spanned a range of plausible mortality regimes. For the first five estimates fitness was measured as the lifetime reproductive success \( R_0 \), which was calculated as

\[
R_0 = e^{-m_\omega} \int_0^\infty e^{-m_\omega x} B(x) \, dx,
\]

where \( m_\omega \) and \( m_\omega \) are larval and adult mortality, \( \alpha \) is age at eclosion (10 d), \( \omega \) is life span (40 d), and \( B(x) \) is fecundity at age \( x \). For scenarios 1, 2, 3, 6, and 7, fecundity was calculated from the function proposed by Roff (1981) following MacMillan et al. (1970):

\[
B(x) = \frac{1}{2} cw [1 - e^{-b(x-20)}] e^{-ax},
\]

where \( w \) is dry weight and \( a, b, \) and \( c \) are constants: \( a = 0.12; b = 0.45; c = 320; \) the first eggs were laid on the third day of adult life. We assumed that dry weight was 0.25 mg; this choice does not affect fitness sensitivity in scenarios 1–5, but it does in scenarios 6 and 7. For scenarios 4 and 5,

\[
B(x) = cw,
\]

that is, fecundity was constant from the third day of adult life until death. For scenarios 6 and 7, we assumed that fitness was measured by \( e^{15r} \), a finite rate of increase per approximate generation (for the assumptions involved, see Stearns 1992). We also calculated, but do not report, the sensitivities of fitness to changes in each trait with \( r \) as the fitness definition. They did not differ much from the sensitivities reported here.

**RESULTS**

First, we report the measures of variation of a series of traits whose impact on fitness is known or can be inferred (table 1): age at eclosion, weight at eclosion, early fecundity, late fecundity, life span. In all three experiments, for both treatments, and at both temperatures, there was a general increase in variation from age at eclosion to weight at eclosion to early fecundity to late fecundity. Life span was generally less variable than late fecundity.

Second, we inferred the sensitivity of fitness to those traits by calculating the percentage change in fitness caused by a 10% change in the trait for seven different scenarios spanning a range of plausible mortality regimes. The sensitivities are reported in table 2A, and the seven scenarios are described in table 2B. In general, fitness was most
sensitive to changes in age and weight at eclosion, rather sensitive to changes in early fecundity, and not very sensitive to changes in late fecundity or life span.

Third, we calculated the relation between the sensitivity of fitness to changes in the trait and the variation of the trait among lines within temperature and treatment. Table 3 reports the results of regressions of trait variation on fitness sensitivity and of Kruskal-Wallis ANOVAs for all seven scenarios for both temperatures (25°C and 29.5°C). The Kruskal-Wallis ANOVAs establish whether the traits differed significantly in their canalization with no normality assumptions; the regressions establish whether those differences in canalization were positively or negatively associated with impact on fitness. We report only the results for coefficients of variation. The results for ranges were the same, and the cases with the best and worst fit (highest and lowest $r^2$) (fig. 3) were both found in the results for CVs. Table 3 also reports only the results for log-transformed data. In the untransformed data, the $P$ values for the regressions were all $\leq 0.0001$, the $P$ values for the Kruskal-Wallis tests were all $\leq 0.0345$; the coefficients of determination ranged from 0.20 to 0.43; and the slopes of the regressions were all negative.

The most striking feature of table 3 is that the relationship is significantly negative for all scenarios. The coefficient of determination ranged from 0.25 to 0.72. Figure 3 presents the best and worst fits for the relationship between coefficients of variation and fitness sensitivities.

Fourth, we excluded genetic variation prior to the manipulation as an explanation for the pat-

### Table 1. Coefficients of variation among lines. EF, elongation factor insert; Control, control insert.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Temp °C</th>
<th>Age at eclosion (d)</th>
<th>Dry weight (mg)</th>
<th>Fecundity (eggs/day)</th>
<th>Life span (d)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Early</td>
<td>Late</td>
</tr>
<tr>
<td>Exp. I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>1.94</td>
<td>6.86</td>
<td>49.27</td>
<td>79.48</td>
</tr>
<tr>
<td>Control</td>
<td>25</td>
<td>0.50</td>
<td>3.65</td>
<td>60.82</td>
<td>86.54</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>1.01</td>
<td>5.88</td>
<td>9.38</td>
<td>57.96</td>
</tr>
<tr>
<td>Exp. II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF</td>
<td>25</td>
<td>1.34</td>
<td>5.16</td>
<td>7.46</td>
<td>32.98</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>1.28</td>
<td>4.71</td>
<td>19.18</td>
<td>31.51</td>
</tr>
<tr>
<td>Control</td>
<td>25</td>
<td>1.21</td>
<td>5.90</td>
<td>9.88</td>
<td>26.01</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>1.00</td>
<td>6.11</td>
<td>31.03</td>
<td>57.85</td>
</tr>
<tr>
<td>Exp. III</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF</td>
<td>25</td>
<td>0.69</td>
<td>5.87</td>
<td>8.22</td>
<td>40.96</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>2.43</td>
<td>8.73</td>
<td>7.35</td>
<td>75.51</td>
</tr>
<tr>
<td>Control</td>
<td>25</td>
<td>2.10</td>
<td>4.02</td>
<td>11.08</td>
<td>75.49</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>2.07</td>
<td>3.06</td>
<td>6.16</td>
<td>70.27</td>
</tr>
</tbody>
</table>

### Table 2. The percent reduction in fitness resulting from a 10% detrimental change in the trait. No trade-offs assumed.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Trait</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at eclosion</td>
<td>9.52</td>
<td>18.13</td>
<td>4.87</td>
<td>9.52</td>
<td>9.52</td>
<td>17.9</td>
<td>25.8</td>
</tr>
<tr>
<td>Dry weight</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.00</td>
<td>10.30</td>
<td>9.38</td>
</tr>
<tr>
<td>Early fecundity</td>
<td>5.92</td>
<td>5.92</td>
<td>6.20</td>
<td>4.41</td>
<td>2.66</td>
<td>5.40</td>
<td>6.50</td>
</tr>
<tr>
<td>Late fecundity</td>
<td>0.00</td>
<td>0.00</td>
<td>0.31</td>
<td>0.00</td>
<td>2.59</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Lifespan</td>
<td>0.00</td>
<td>0.00</td>
<td>0.83</td>
<td>0.00</td>
<td>10.35</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>B. Mortality values</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Larval</td>
<td>0.100</td>
<td>0.200</td>
<td>0.050</td>
<td>0.100</td>
<td>0.100</td>
<td>0.100</td>
<td>0.100</td>
</tr>
<tr>
<td>Adult</td>
<td>0.400</td>
<td>0.400</td>
<td>0.000</td>
<td>0.400</td>
<td>0.001</td>
<td>0.400</td>
<td>0.000</td>
</tr>
<tr>
<td>Fitness measure</td>
<td>$R_0$</td>
<td>$R_0$</td>
<td>$R_0$</td>
<td>$R_0$</td>
<td>$e^{15r}$</td>
<td>$e^{15r}$</td>
<td></td>
</tr>
</tbody>
</table>
Table 3. The relationships of coefficients of variation of traits among lines to the sensitivity of fitness to changes in those traits. KW, Kruskal-Wallis.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. 25°C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td>0.57</td>
<td>0.69</td>
<td>0.25</td>
<td>0.64</td>
<td>0.33</td>
<td>0.69</td>
<td>0.70</td>
</tr>
<tr>
<td>Slope</td>
<td>-0.115</td>
<td>-0.069</td>
<td>-0.123</td>
<td>-0.127</td>
<td>-0.096</td>
<td>-0.071</td>
<td>-0.049</td>
</tr>
<tr>
<td>$P$</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>$P(KW)$</td>
<td>0.0007</td>
<td>0.0003</td>
<td>0.0017</td>
<td>0.0007</td>
<td>0.0087</td>
<td>0.0043</td>
<td>0.0043</td>
</tr>
<tr>
<td>B. 29.5°C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r^2$</td>
<td>0.60</td>
<td>0.72</td>
<td>0.28</td>
<td>0.68</td>
<td>0.48</td>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td>Slope</td>
<td>-0.118</td>
<td>-0.079</td>
<td>-0.129</td>
<td>-0.130</td>
<td>-0.115</td>
<td>-0.071</td>
<td>-0.050</td>
</tr>
<tr>
<td>$P$</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0001</td>
</tr>
<tr>
<td>$P(KW)$</td>
<td>0.0048</td>
<td>0.0004</td>
<td>0.0141</td>
<td>0.0048</td>
<td>0.0016</td>
<td>0.0350</td>
<td>0.0035</td>
</tr>
</tbody>
</table>

The ANOVAs in Table 4 lack power because there were only two degrees of freedom for the denominator sum of squares. However, even in a large design it is unlikely that we would have detected a significant effect of experiment, temperature, or treatment on the coefficients of variation of any of the traits except late fecundity, because for each of those traits at least one of the two-way interactions had a larger mean square than any of the three main effects. A larger design might have detected significant differences among experiments in the coefficients of variation and ranges of late fecundity.

Experiment II and experiment III are compared in figure 4A (25°C) and figure 4B (29.5°C). The points overlap completely at 25°C; at 29.5°C there is a slight, nonsignificant tendency for the variation among lines to be lower in experiment III (intercept, 1.32; SE, 0.17) than experiment II (intercept, 1.49; SE, 0.13). That small difference could be attributed to a reduction in overall genetic variation among lines.

Fifth, the results of the three-way ANOVA and the plots of the regression lines for each temperature (fig. 3 A,B) suggest that there was no effect of temperature on the degree of canalization of the traits. Thus, a trait like age at eclosion had a relatively canalized reaction norm, corresponding to figure 2A, and a trait like late fecundity had a relatively uncanalized reaction norm, corresponding to figure 2B, but none of the traits was significantly better canalized at one temperature than at the other (fig. 2C).

Discussion

The Canalization of Fitness Components

Does genetic canalization exist at all? If it does, are traits with greater impact on fitness more strongly canalized against genetic perturbations? The results from these experiments on Drosophila melanogaster suggest that genetic canalization exists and that fitness components are canalized in relation to their impact on fitness. The reliability of this conclusion rests both on the implausibility of alternative explanations of the pattern and on the pattern’s robustness.

The significantly negative relationship be-
The relationship between the variation among lines and the sensitivity of fitness to changes in traits for all three experiments combined. Each point represents an estimate of variation among lines within one treatment and one temperature. Open symbols: 25°C; filled symbols: 29.5°C. A. Best case: Scenario 2, fitness measure $R_e$. 25°C: slope $= -0.069$, $r^2 = 0.69$, $P(KW) = 0.0003$ (Kruskal-Wallis). 29.5°C: slope $= -0.070$, $r^2 = 0.72$, $P(KW) = 0.0004$. B. Worst case: Scenario 3, fitness measure $R_e$. 25°C: slope $= -0.123$, $r^2 = 0.25$, $P(KW) = 0.0017$. 29.5°C: slope $= -0.129$, $r^2 = 0.28$, $P(KW) = 0.0141$. Triangles, age at eclosion; circles, dry weight at eclosion; squares, early fecundity; circles, late fecundity; triangles, life span.

Fig. 4. A representative comparison of the results for experiment II (open symbols) and experiment III (closed symbols) at 25°C and 29.5°C. Triangles, age at eclosion; circles, dry weight at eclosion; squares, early fecundity; circles, late fecundity; triangles, life span.

Could the pattern be explained by different numbers of loci segregating for each of the traits measured? To support that position, one would have to argue that the genetic variation among lines for these traits remained similar in experiments I, II, and III, despite the large differences in the genetic constitution of the flies used in those experiments. That seems particularly implausible given the lack of change in the pattern from experiment II, in which each line had a different genetic background, to experiment III, in which each line had the same genetic background. The only alternative is that the presence
of plasmid inserts on six different places on the third chromosome in experiment III, five more than in experiment II, compensated for the reduction from six to one different genetic backgrounds. Even if the larger number of plasmid inserts had exactly that effect, it would have been an effect on the overall expression of genetic variation rather than a compensatory increase in the number of loci segregating for each trait. We find that alternative implausible and conclude that, in these experiments, the greater the contribution of a trait to fitness, the stronger was the genetic canalization of that trait.

**Canalization and Heritability**

The heritability of fitness is a central concept in evolutionary genetics. Because fitness itself is always under strong selection, additive genetic variation for fitness should be steadily reduced and would eventually vanish if there were no mutation or gene flow (Bulmer 1985, chap. 9). Although natural selection depletes genetic vari-
ation by eliminating deleterious alleles, some additive genetic variation in traits under selection is widely observed (Mousseau and Roff 1987; Roff and Mousseau 1987). The explanations include mutation-selection balance, overdominance induced by genotype-by-environment interactions, frequency-dependent selection, spatial variation, and pleiotropy (Barton and Turelli 1989).

The pattern on which we have focused occurs among traits varying in their impact on fitness. Robertson (1955, 1968) suggested that heritabilities should decrease in traits with stronger correlations to fitness because stronger selection eliminates more deleterious alleles. Traits more closely correlated with fitness do tend to have lower heritabilities (Gustafsson 1986; Roff and Mousseau 1987), supporting the genetic explanation of this pattern of variation in fitness components. Here we provide a complementary explanation for low genetic variation in traits important to fitness. A negative relation between the variation of traits and the sensitivity of fitness to those traits may have two causes that are not mutually exclusive. It can be caused by a reduction in the number of loci segregating, by an increase in canalization, or by both, as fitness sensitivity increases. We emphasize that the two explanations are not exclusive, and that we ruled out Robertson's explanation only in the results reported here, not in natural populations, where we expect both effects to be present.

At first sight, an increase in the canalization of a trait would appear to increase, not decrease, its heritability, because canalization causes a reduction in phenotypic variation, and \( h^2 = V_A / V_p \), where \( V_A \) = additive genetic variance and \( V_p \) = total phenotypic variance. If the numerator does not change, a reduction in the denominator should increase the heritability. However, the denominator \( V_p = V_A + V_e \). Genetic canalization reduces the additive genetic variance, \( V_A \), by reducing the magnitude of allelic effects, whereas environmental canalization reduces the environmental variance, \( V_e \). Whether the evolution of canalization increases or decreases the heritability depends on the quantitative balance of those two effects. It is possible that the same developmental mechanisms buffer traits against both genetic and environmental perturbations. If they do—and whether they do or not is an important open research question—then the evolution of canalization will have a complex mixture of effects on both the numerator and the denominator of the expression for heritability.

It takes very large experimental designs to discriminate a heritability of 0.1 from zero (Falconer 1989). If genetic canalization reduces the additive genetic variance into the region of measurement error, then it will in practice reduce the number of heritabilities detected as statistically significant. Thus, some of the pattern attributed to differences among traits in heritability, usually thought to be caused by differences in current selection pressures, may also reflect the degree of canalization of the traits involved (cf. Gustafsson 1986; Roff and Mousseau 1987).

**Future Experiments**

These experiments did not measure the number of loci segregating for each trait, which is needed to reliably estimate the magnitude of the average effects of alleles on the traits. One future challenge is to identify the loci segregating for a series of traits that appear to be differentially canalized, then to construct, with an appropriate breeding scheme, test groups with different numbers of loci segregating for each trait, and through that manipulation to estimate the average effects of alleles on each trait. This approach to studying canalization, more reliable and direct than the one used in this paper, awaits additional knowledge of quantitative trait loci (QTLs) (cf. Tanksley 1993).

The comparisons used in this study were primarily among traits expressed at different ages. One could argue that increasing variation of traits with age was not the product of reduced selection at greater age but rather of the accumulation of errors with age. Kirkwood's (1987) work on the evolution of age-specific investment in maintenance removes some of the force of this alternative, but it would nevertheless be helpful to have experiments comparing the canalization of several traits expressed at the same age but with different impact on fitness. In these experiments, age at eclosion did appear to be more canalized than dry weight at eclosion (fig. 3A), but whether age or weight at eclosion had more impact on fitness depended on the mortality scenario (fig. 3B). The data presented here are not sufficient to refute that alternative.

**Microevolution, Canalization, and Macroevolution**

Developmental mechanisms may evolve more slowly than patterns of heritabilities. When they do, then the pattern of phenotypic variation reported here would be caused by past selection no longer acting on a developmental system that
continued to affect the expression of genetic variation in the present. However, canalization is known in some cases to evolve rapidly. Maynard Smith and Sondhi (1960), for example, succeeded in artificially selecting increased canalization of a morphological pattern in seven generations. We have no general answer to the question, Are deleterious alleles eliminated more rapidly by selection than canalization against them can evolve? The results reported here emphasize the importance of answering that question.

Development may play a more intimate role in microevolution than had previously been suspected by modulating the expression of genetic variation of fitness components themselves. The microevolution of the developmental system should be ultimately responsible for the origin of the macroevolutionary patterns we call historical or phylogenetic. While all DNA sequences mutate, which implies that no genetic information is ever completely fixed, some traits are more canalized than others, and the end point of canalization is developmental—not genetic—fixation of the trait. In that state, the trait does not vary at all within the population. Canalization may be the key to understanding the conversion of genetic variation within populations into traits developmentally fixed within lineages. Such fixed traits form the macroevolutionary patterns (Stearns 1994) that suggest phylogenetic and developmental constraints (Maynard Smith et al. 1985; Scharloo 1991).

ACKNOWLEDGMENTS

To M. Kaiser, A. Mislin, B. Sykes, and H. Zingerli, our heartfelt thanks. They made these experiments possible. Comments by N. Barton, M. Doebeli, G. de Jong, I. Olivieri, F. Kjellberg, two anonymous referees, and especially by R. Huey and R. Grosberg improved the analysis and presentation. Our research was supported by the Swiss Nationalfonds (no. 31-28511.90).

LITERATURE CITED


Cooley, L., R. Kelley, and A. Spradling. 1988. Inser-


Stearns, S. C., and M. Kaiser. 1993. The effects of enhanced expression of elongation factor EF-1α on


Corresponding Editor: R. Grosberg