# Environment Dependence of Mutational Parameters for Viability in *Drosophila melanogaster*

**James D. Fry*,[1] and Stefanie L. Heinsohn†**

**Department of Biology, University of Rochester, Rochester, New York 14627-0211 and †Department of Genetics, North Carolina State University, Raleigh, North Carolina 27695-7614*

Manuscript received October 17, 2001
Accepted for publication March 29, 2002

ABSTRACT

The genomic rate of mildly deleterious mutations ($U$) figures prominently in much evolutionary and ecological theory. In *Drosophila melanogaster*, estimates of $U$ have varied widely, from <0.1 to nearly 1 per zygote. The source of this variation is unknown, but could include differences in the conditions used for assaying fitness traits. We examined how assay conditions affect estimates of the rates and effects of viability-depressing mutations in two sets of lines with accumulated spontaneous mutations on the second chromosome. In each set, the among-line variance in egg-to-adult viability was significantly greater when viability was assayed using a high parental density than when it was assayed using a low density. In contrast, the proportional decline in viability due to new mutations did not differ between densities. Two other manipulations, lowering the temperature and adding ethanol to the medium, had no significant effects on either the mean decline or among-line variance. Cross-environment genetic correlations in viability were generally close to one, implying that most mutations reduced viability in all environments. Using data from the low-density, lower-bound estimates of $U$ approached the classic, high values of Mukai and Ohnishi; at the high density, $U$ estimates were similar to recently reported low values. The difference in estimated mutation rates, taken at face value, would imply that many mutations affected fitness at low density but not at high density, but this is shown to be incompatible with the observed high cross-environment correlations. Possible reasons for this discrepancy are discussed. Regardless of the interpretation, the results show that assay conditions can have a large effect on estimates of mutational parameters for fitness traits.

THE genomic rate and distribution of effects of deleterious mutations are critical parameters in much evolutionary theory (Lynch *et al.* 1999). High rates of mildly deleterious mutations could provide an advantage for sexual reproduction (Kondrashov 1988; Charlesworth 1990; Peck 1994), promote the evolution of outcrossing mechanisms (Charlesworth *et al.* 1990), reduce neutral genetic variation (Charlesworth *et al.* 1993), and endanger small populations (Gabriel and Bürger 1994; Lande 1995; Lynch *et al.* 1995).

Several experiments have been performed to attempt to put a lower bound on the rate of mutations with negative effects on egg-to-adult viability in *Drosophila melanogaster* (Mukai 1964; Mukai *et al.* 1972; Ohnishi 1977; García-Dorado *et al.* 1998; Fry *et al.* 1999; Chavarrías *et al.* 2001). In these "mutation-accumulation" (MA) experiments, spontaneous mutations are allowed to accumulate in very small populations, in which the efficacy of natural selection is reduced. Three early studies (Mukai 1964; Mukai *et al.* 1972; Ohnishi 1977) reported estimates of the minimum number of new deleterious mutations per zygote of 0.30–0.85. In contrast, estimates from more recent studies have been in the range 0.01–0.1 (García-Dorado *et al.* 1998; Fry *et al.* 1999; Chavarrías *et al.* 2001). While mutation rates near 1 would be sufficient for many of the hypothesized evolutionary consequences of deleterious mutations, mutation rates ~0.1 would not. Thus the Drosophila data are frustratingly ambiguous on the evolutionary significance of deleterious mutation.

One problem with interpreting the Drosophila data is that the experiments have differed in several ways that could potentially affect the estimates: These include the method for accumulating mutations, the method for estimating base population (control) fitness, the fitness traits assayed, and the assay conditions (medium recipe, humidity, larval density, etc.). Coming to any generalizations about genomic mutation rates in Drosophila will be difficult without understanding how these factors may influence the estimates.

The goal of this study is to examine how differences in assay conditions affect estimates of the rates and effects of deleterious mutations in *D. melanogaster*. This was also one of the goals of our earlier study (Fry *et al.* 1999), in which viability of a set of MA lines was assayed in four environments differing in temperature, rearing density, and medium composition. (The rationale for

[1]*Corresponding author:* Department of Biology, University of Rochester, Rochester, NY 14627-0211. E-mail: jfry@mail.rochester.edu

the choice of environments is given below.) The analysis reported in that article was incomplete, however, and an important feature of the results was overlooked. In this article, we give a more thorough analysis of genotype × environment interactions in the Fry *et al.* (1999) dataset and report a similar analysis of a new MA experiment that followed the same design as that of Fry *et al.* (1999). By comparing the results of two parallel experiments, we hope to be able to form tentative generalizations about the effects of assay conditions on mutational parameter estimates.

The most widespread method for estimating the genomic rate of deleterious mutations, $U$, is by the formula of Bateman (1959) and Mukai (1964), which gives a lower bound:

$$U \geq U_{BM} = c\frac{[\Delta M]^2}{\Delta V}. \quad (1)$$

Here, $\Delta M$ is the per generation rate by which fitness declines in a set of MA lines, and $\Delta V$ is the increase in among-line variance per generation; the "$c$" scales the estimate to the entire genome. Effects of assay conditions on $U_{BM}$ could arise through effects on $\Delta M$, $\Delta V$, or both. Because statistical comparisons of $\Delta M$ and $\Delta V$ between treatments are relatively straightforward, while comparisons of $U_{BM}$ are not, the focus of this article is on how $\Delta M$ and $\Delta V$ are affected by assay conditions.

A second goal of the work reported here is to estimate cross-environment genetic correlations in viability among the MA lines. Low correlations of mutational effects across environments would suggest the possibility that many mutations that are deleterious in some environments are approximately neutral in others. This would imply that the rate of mutations that are deleterious in at least one environment could be much higher than the rate measured in any given environment. In addition, such conditionally deleterious mutations could contribute to the evolution of specialization (Kawecki 1994; Fry 1996; Kawecki *et al.* 1997) and speciation (Kawecki 1997). Only a few studies have reported estimates of cross-environment correlations in fitness traits in MA lines (Fry *et al.* 1996; Fernández and López-Fanjul 1997; Vassilieva *et al.* 2000).

## MATERIALS AND METHODS

**Mutation-accumulation experiments:** Spontaneous mutations were accumulated on *D. melanogaster* second chromosomes shielded from recombination and selection by the method of Mukai (1964). Details of the procedure may be found in Fry *et al.* (1999). The new MA experiment (hereafter, experiment 2) used identical methods as the experiment of Fry *et al.* (1999; hereafter, experiment 1), except that a different progenitor chromosome was used, and viability assays were performed at slightly different generation numbers (see below). In both experiments, three control populations were established by making the original chromosome homozygous. The control populations were maintained at an effective population size larger than that of the MA lines (~100 *vs.* 1) and with a longer generation interval (4 weeks *vs.* ~2.5), both of which should have slowed the accumulation of deleterious mutations relative to the MA lines. For the purpose of comparing $\Delta M$ between treatments, it is not necessary that the control lines accumulated no deleterious mutations, only that they accumulated fewer than the MA lines.

**Viability assays:** Viability was assayed using the "*Curly*" method. In this method, $Cy/+_i$ females and males are intercrossed, where $+_i$ denotes a wild-type chromosome derived from the $i$th MA line, and *Cy* denotes a balancer chromosome bearing the dominant markers *Curly* and *Rough eye* (Fry *et al.* 1999). In the absence of differential viability, this cross is expected to produce $+_i/+_i$ and $Cy/+_i$ progeny in a 1:2 ratio. The departure of the ratio from 1:2 is used as a measure of the relative viability of the mutant homozygotes to balancer heterozygotes:

$$\text{RV} = \frac{2 \times (\text{no. wild-type progeny})}{(\text{no. Curly progeny}) + 1}. \quad (2)$$

The 1 in the denominator is a slight bias correction (Haldane 1956).

Viability assays were performed at generations (G)27 and G33 in experiment 1 and at G31 and G35 in experiment 2. In each experiment, approximately one-half the MA lines were assayed at the earlier time, and the remainder were assayed at the later time. All three control populations were assayed at both times.

Four environmental treatments were used for the viability assays. In the "standard" treatment, six pairs of Cy/+ flies were allowed to lay eggs on cornmeal-molasses-dead yeast-agar medium at 25°. Each of the other three treatments differed in one respect from the standard treatment: Two pairs were allowed to lay eggs in the "low-density" treatment; the temperature was reduced to 18° in the "low-temperature" treatment; and the medium was supplemented with 10% ethanol in the "ethanol" treatment. In the first set of assays in each experiment (G27 and G31), viability was measured in the standard, low-density, and low-temperature treatments, with 6 and 12 replicate crosses per treatment for each MA and control line, respectively. In the second set of assays (G33 and G35), viability was measured in the standard and ethanol treatments, with 8 MA and 16 control replicates. Crosses produced an average of 300 flies (range 21–531), with only 1% producing <100. Additional details may be found in Fry *et al.* (1999).

The standard and low-density treatments were designed to match roughly the conditions (medium, temperature, and parental density) used by Mukai (1964) and Ohnishi (1974, 1977), respectively. Although eggs were not counted, larval density differed visibly between the treatments, and the average number of flies emerging was greater in the standard treatment (363) than in the low-density treatment (310). The ethanol and low-temperature treatments were chosen because of their relevance to wild *D. melanogaster* populations (Gibson *et al.* 1981; Jones *et al.* 1987). These treatments resulted in slower development and fewer emerging flies (means of 198 and 272, respectively) than did the standard treatment.

Lines classified as lethal (<5% relative viability) on the basis of the first two crosses in the standard treatment were not considered further. Lethal mutation rates were 1.0% per generation (95% C.I.: 0.7–1.5%) in experiment 1 (Fry *et al.* 1999) and 0.9% (0.6–1.3%) in experiment 2 (J. D. Fry, unpublished data).

After each set of viability assays, each control population and 5–10 of the assayed MA lines were checked for contamination by scoring polytene second chromosomes for insertion sites of the transposable elements *roo* [experiment (exp.) 1]

or *copia* (exp. 2). All lines in exp. 1 shared a common set of 19 *roo* sites (Fry *et al.* 1999), and all lines in exp. 2 shared a common set of 8 *copia* sites (S. V. Nuzhdin, personal communication, 1998), confirming lack of contamination (*cf.* Charlesworth *et al.* 1992).

**Statistical analysis:** Means, among-line, and within-line variances, as well as standard errors for these statistics, were estimated for each treatment and line type (MA or control) separately by restricted maximum likelihood (REML), using the MIXED procedure in SAS version 8 (SAS Institute 1992; Littell *et al.* 1996). Blocks were included as a random effect (see Fry *et al.* 1999 for details of block structure). Likelihood-ratio tests were used to test whether among-line variances differed from zero. Under the null hypothesis, twice the difference in log-likelihoods between the full model and a model with the line effect left out should have a chi-square distribution with 1 d.f. As the test in this case is one-tailed, the resulting probabilities were halved.

Likelihood-ratio tests were also used to determine whether $\Delta V$ differed among treatments, using data for each treatment pair within each set of MA lines. In the full model, among-line variances were allowed to differ between treatments ("unstructured" covariance option on "RANDOM" statement); in the constrained model, the variances were constrained to be equal ("Toeplitz" covariance option). Both the constrained and unconstrained models allowed residual variances to differ between treatments ("GROUP =" option on "RESIDUAL" statement) and included random block and block × treatment interactions. The test in this case is two-tailed. Because comparisons of variances are likely to be sensitive to violations of the normality assumption, the comparisons were done only for "quasinormal" lines. These are defined as lines with more than one-half the viability of the controls (*cf.* Mukai 1964; Mukai *et al.* 1972; Ohnishi 1977). As is seen below, although few in number, the "severely deleterious" (*i.e.*, not quasinormal) lines induced a strong left skew to the distributions of line means and greatly inflated among-line variances.

Genetic correlations across treatments and their standard errors were estimated by repeating the above unconstrained analyses using the "UNR" covariance option. Likelihood-ratio tests of the hypotheses $r_g = 0$ and $r_g = 1$ were performed by constraining the correlations with the "UN(1)" covariance option and the "PARMS" statement, respectively. The test of $r_g = 0$ is two-tailed, while that of $r_g = 1$ is one-tailed. If the point estimate of $r_g$ was 1, no standard error could be calculated; instead, a lower bound for $r_g$ was found as the smallest value that did not cause a significant ($P < 0.05$) increase in the log-likelihood of the model.

To determine whether $\Delta M$ differed among treatments, PROC MIXED was used to perform an analysis on the line means, with fixed effects of line type, treatment, and their interaction. A significant interaction would imply that the difference in viability between MA lines and controls depends on treatment and hence that $\Delta M$ differs among treatments. (Attempts to perform a similar analysis on the raw values rather than on means were unsuccessful, due to lack of convergence and exceedingly long run times.) The analysis took into account the correlations of MA line means across treatments, using the unstructured and heterogeneous Toeplitz covariance options for datasets with two and three treatments, respectively. These covariance structures also permit variances among lines to differ between treatments. Separate covariance matrices were estimated for MA and control lines, using the GROUP = option; this allowed variances to differ between MA lines and controls. To increase the power of the analysis by eliminating extraneous parameters, correlations across treatments for the control lines were assumed to be zero, as expected if most of the variation among the controls was due to sampling error. This assumption did not significantly reduce the fit of the model (likelihood-ratio tests, $P > 0.15$), except in the case of the G33 dataset of experiment 1 ($P < 0.01$). For this dataset, results are reported with and without the correlation constrained.

Estimates of the mutational parameters $\Delta M$ and $\Delta V$ were made for each treatment in each experiment. $\Delta M$ was calculated as the difference in mean viability between MA lines and controls, divided by the generation number and control mean. $\Delta V$ was calculated as the variance among MA lines, divided by the generation number and the square of the control mean. For the standard treatment, in which viability was measured in both sets of lines in each experiment, the estimates reported are the averages of the two estimates, weighted by the number of lines in each set. $U_{BM}$ was calculated by Equation 1, with $c = 5$. The Bateman-Mukai upper-bound estimate of the average mutational effect, $S_{BM}$, was calculated as the ratio of $\Delta V$ to $\Delta M$. Both $U_{BM}$ and $S_{BM}$ were estimated with severely deleterious lines excluded; such estimates are likely to be the most informative (Crow and Simmons 1983). Because the main goal of this work is to determine the qualitative effects of environment on mutational parameters, an issue that is addressed by the statistical tests above, only point estimates of the mutational parameters are given.

If a subset of MA lines had viability-increasing mutations, $U_{BM}$ would severely underestimate $U$ (*cf.* Shaw *et al.* 2000). To test this possibility, the mean of each MA line was compared to the mean of the pooled controls for each treatment separately, using Dunnet's method to adjust for multiple comparisons and including a block effect in the analysis. This would be a conservative test for adaptive mutations if the control lines had themselves increased in viability due to adaptive mutations. This possibility seems unlikely, however, for reasons given below.

## RESULTS

Mean viabilities of the MA and control lines are given in Table 1, and viabilities of the individual lines are depicted in Figures 1 and 2. MA line means were always lower than control means (Table 1); in the mixed-model analysis, effects of line type were significant in each of the four datasets (Table 2). Treatment effects were less strong. In the first sets of assays (G27 and G31), means were lower in the standard treatment than in the low-density and low-temperature treatments (Table 1); the differences approached significance only in experiment 1, however (Table 2). In the second set of assays in experiment 1 (G33), viability was significantly or almost significantly lower in the ethanol treatment than in the standard treatment (Tables 1 and 2). In contrast, in experiment 2 (G35), the pattern of means was reversed, but the differences were not significant. In no case was the interaction between treatment and line type significant (Table 2); thus there is no evidence that the difference in viability between control and MA lines (and hence $\Delta M$) depended on treatment.

In both the G27 (exp. 1) and G31 (exp. 2) datasets, two MA lines classified as severely deleterious in the low-temperature treatment were nonetheless quasinormal in the other two treatments (Figure 1). In the G35 dataset (exp. 2) were four severely deleterious lines, two of which were severely deleterious in the ethanol treatment

**TABLE 1**

**Viability means**

| Experiment | Generation | Treatment | Control mean (SE) | MA lines | | |
|---|---|---|---|---|---|---|
| | | | | Lines included | *N* | Mean (SE) |
| 1 | 27 | Low dens. | 0.874 (0.023) | QN | 38 | 0.806 (0.011) |
| | | Standard | 0.803 (0.028) | QN | 38 | 0.791 (0.016) |
| | | Low temp. | 0.882 (0.040) | QN | 36 | 0.818 (0.015) |
| | | Low temp. | | QN + SD | 38 | 0.787 (0.026) |
| 2 | 31 | Low dens. | 0.863 (0.032) | QN | 37 | 0.824 (0.017) |
| | | Standard | 0.792 (0.060) | QN | 37 | 0.747 (0.029) |
| | | Low temp. | 0.881 (0.028) | QN | 35 | 0.796 (0.025) |
| | | Low temp. | | QN + SD | 37 | 0.768 (0.032) |
| 1 | 33 | Ethanol | 0.789 (0.027) | QN | 34 | 0.692 (0.036) |
| | | Standard | 0.806 (0.018) | QN | 34 | 0.763 (0.021) |
| 2 | 35 | Ethanol | 0.853 (0.021) | QN | 28 | 0.808 (0.021) |
| | | Standard | 0.837 (0.034) | QN | 30 | 0.728 (0.040) |
| | | Ethanol | | QN + SD | 32 | 0.736 (0.041) |
| | | Standard | | QN + SD | 32 | 0.699 (0.044) |

QN, quasinormal; SD, severely deleterious (see text for definitions).

only (Figure 2). Excluding the severely deleterious lines had little effect on inferences about line type or treatment effects (Table 2).

Variances among the control lines were low and nonsignificant in 9 of 10 instances, the exception being the standard treatment of experiment 2 at G31 (Table 3, Figures 1 and 2). The same three control lines did not differ when retested at G35 (Table 3, Figure 2), suggesting that the difference at G31 had a temporary, environmental cause. In contrast, variances among the MA lines were highly significant except in the low-density treatment, where they nonetheless approached significance (Table 3).

Variances among the MA lines were significantly greater in the standard treatment than in the low-density treatment in each experiment (Tables 3 and 4). The difference was more than fivefold in each case. In experiment 2 but not experiment 1, the among-line variance in the low-temperature treatment was significantly higher than that in the low-density treatment. No other differences in among-line variances were significant. Because REML comparisons of variances are sensitive to nonnormality, we repeated the comparisons between the standard and low-density treatments using a nonparametric bootstrap procedure. Viability means were calculated for each line and treatment combination, and 10,000 bootstrap samples of the resulting bivariate distributions were made by sampling lines with replacement. For each sample, the variance of line means was calculated for each treatment. Variances in the standard treatment exceeded those in the low density treatment in all but 3.1 and 1.4% of the samples in experiments 1 and 2, respectively. This test is conservative because both among- and within-line variances contribute to the sample variance of line means, and the latter were higher in the low-density treatment (Table 3). The results thus confirm the treatment effect on among-line variances.

In an analysis of variance on data from the standard

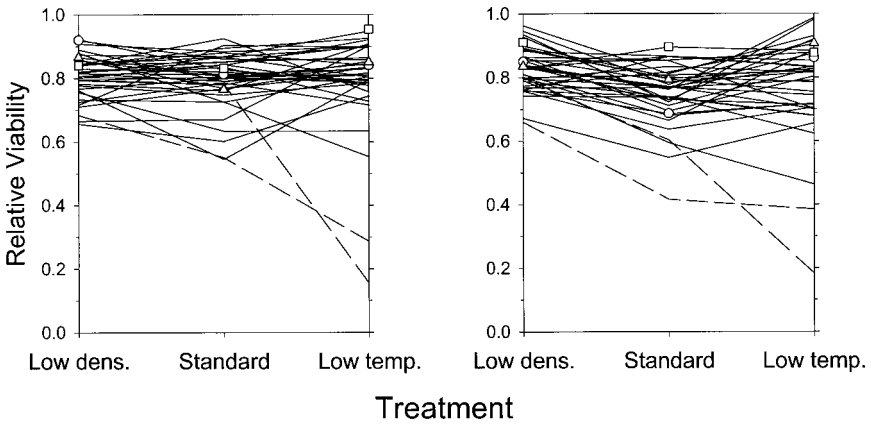


FIGURE 1.—Mean viabilities at generation 27 in the first experiment (left) and at generation 31 in the second experiment (right). Viabilities of the three control populations are represented by open symbols. MA lines classified as "severely deleterious" in the low-temperature treatment are represented by dashed lines; all others are represented by solid lines.

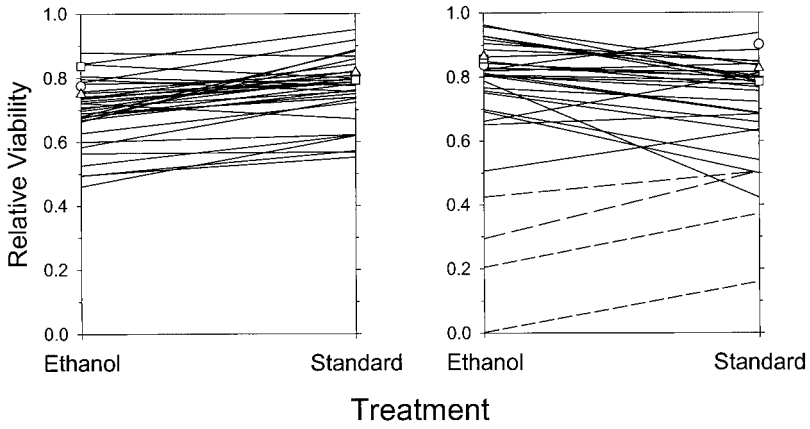


FIGURE 2.—Mean viabilities at generation 33 in the first experiment (left) and at generation 35 in the second experiment (right). Viabilities of the three control populations are represented by open symbols. MA lines classified as severely deleterious in both treatments are represented by short-dashed lines, those severely deleterious in the ethanol treatment only by long-dashed lines, and all others by solid lines.

and low-density treatments in experiment 1, there was no significant interaction between line and treatment (Table 2 in FRY *et al.* 1999). This led FRY *et al.* (1999) to overlook the difference in among-line variances between the treatments. A similar analysis of variance on data from experiment 2 (not shown) likewise shows no significant line × treatment interaction ($P > 0.3$). Although differences in among-line variance contribute to the genotype × environment interaction variance (COCKERHAM 1963), REML apparently provides a more powerful method for detecting such differences.

All cross-environment correlations in viability were positive (Table 5, Figures 1 and 2). Of eight correlations estimated with severely deleterious lines excluded, four were equal to one, two were not significantly less than one ($P > 0.15$), and two approached being significantly less than one ($P \approx 0.06$; Table 5). Inclusion of the severely deleterious lines in some cases decreased, but in others increased, the correlations.

With severely deleterious lines excluded, the range in $\Delta M$ among treatments was 0.10–0.37% in experiment 1 and 0.15–0.31% in experiment 2 (Table 6). Including the severely deleterious lines increased the estimates, although not dramatically. Differences in $\Delta M$ between treatments were not consistent across experiments, as expected given that they were within the bounds of sampling error (see above). Averaging across experiments, $\Delta M$ with severely deleterious lines excluded showed a small range among treatments, from 0.19% in the standard treatment to 0.29% in the low-temperature treatment.

Estimates of $\Delta V$ with severely deleterious lines excluded were fairly consistent between experiments (Table 7), showing less than twofold variation in each treatment. Averaging across the standard, low-temperature, and ethanol treatments, $\Delta V$ with severely deleterious (SD) lines excluded was $\sim 3.4 \times 10^{-4}$. (For simplicity, these are referred to as the "high-density" treatments, even though larval density was not necessarily higher in the latter two treatments than in the low-density treatment; see MATERIALS AND METHODS). Average $\Delta V$ in the low-density treatment was $0.55 \times 10^{-4}$ (note that no lines were classified as severely deleterious in the low-density treatment). Including SD lines greatly increased the $\Delta V$

**TABLE 2**

**Effects of treatment, line type (MA or control), and their interaction on viability**

| Experiment | Generation | Lines included | Treatment | | Line type | | Treatment × line type | |
|---|---|---|---|---|---|---|---|---|
| | | | $F$ (d.f.$_1$, d.f.$_2$) | $P$ | $F$ (d.f.$_1$, d.f.$_2$) | $P^a$ | $F$ (d.f.$_1$, d.f.$_2$) | $P$ |
| 1 | 27 | QN | 4.29 (2, 5.62) | 0.074 | 5.63 (1, 8.89) | 0.021 | 1.99 (2, 5.62) | 0.22 |
| | | QN + SD | 3.80 (2, 6.31) | 0.083 | 7.81 (1, 13.2) | 0.008 | 2.35 (2, 6.31) | 0.17 |
| 2 | 31 | QN | 2.31 (2, 2.71) | 0.26 | 4.20 (1, 4.48) | 0.051 | 1.50 (2, 2.71) | 0.36 |
| | | QN + SD | 2.69 (2, 3.64) | 0.19 | 6.14 (1, 6.07) | 0.024 | 2.35 (2, 3.64) | 0.22 |
| 1 | 33 | QN | 9.33 (1, 3.28) | 0.049 | 10.98 (1, 12.3) | 0.003 | 3.30 (1, 3.28) | 0.16 |
| | | QN[b] | 6.45 (1, 2.52) | 0.10 | 14.00 (1, 21.3) | 0.001 | 2.28 (1, 2.52) | 0.24 |
| 2 | 35 | QN | 3.97 (1, 3.96) | 0.12 | 7.50 (1, 8.93) | 0.012 | 1.53 (1, 3.96) | 0.28 |
| | | QN + SD | 1.65 (1, 4.35) | 0.26 | 11.49 (1, 24.8) | 0.001 | 0.30 (1, 4.35) | 0.61 |

Results of Satterthwaite approximate $F$-tests (LITTELL *et al.* 1996) are shown.

[a] Probability values for line type were halved to produce a one-tailed test of the prediction that MA lines should have lower viability than controls.

[b] Model with no constraint on control correlations (see text).

**TABLE 3**

**Among- and within-line variances**

| | | | Control lines | | MA lines | |
|---|---|---|---|---|---|---|
| Experiment | Generation | Treatment | Among-line variance | Within-line variance | Among-line variance | Within-line variance |
| 1 | 27 | Low dens. | 0.03 (0.17) | 1.50 (0.38) | 0.12* (0.11) | 1.50 (0.38) |
| | | Standard | 0.01 (0.12) | 1.34 (0.36) | 0.61***(0.19) | 1.33 (0.36) |
| | | Low temp. | 0.13 (0.40) | 3.15 (0.84) | 0.35***(0.16) | 1.74 (0.19) |
| | | Low temp.[a] | | | 2.18***(0.57) | 1.70 (0.18) |
| 2 | 31 | Low dens. | 0 | 1.85 (0.50) | 0.12** (0.12) | 2.09 (0.22) |
| | | Standard | 0.98***(1.09) | 1.28 (0.32) | 0.67***(0.22) | 1.62 (0.17) |
| | | Low temp. | 0 | 1.09 (0.30) | 0.75***(0.29) | 2.65 (0.29) |
| | | Low temp.[a] | | | 2.14***(0.61) | 2.58 (0.27) |
| 1 | 33 | Ethanol | 0 | 3.25 (0.68) | 0.77***(0.25) | 1.98 (0.19) |
| | | Standard | 0 | 1.28 (0.28) | 0.76***(0.24) | 1.66 (0.15) |
| 2 | 35 | Ethanol | 0 | 2.02 (0.43) | 0.65***(0.29) | 3.02 (0.31) |
| | | Standard | 0.17 (0.34) | 2.42 (0.55) | 1.38***(0.43) | 1.98 (0.20) |
| | | Ethanol[a] | | | 4.60***(1.26) | 2.81 (0.27) |
| | | Standard[a] | | | 2.66***(0.74) | 1.90 (0.18) |

Variances and their standard errors (in parentheses) have been multiplied by 100. Among-line variances: *$P = 0.095$; **$P = 0.12$; ***$P < 0.001$. All others $P > 0.2$. Dens., density.

[a] Including severely deleterious lines (see Table 1).

estimates, with the differences between the low-density treatment and the other treatments becoming larger.

Because $\Delta M$ and $\Delta V$ did not vary significantly among the high-density treatments, we pooled estimates of these quantities from these treatments to estimate $U_{BM}$ and $S_{BM}$. Using the $\Delta M$ values in Table 6 for quasinormal lines, the weighted average from the high-density treatments is 0.210% in experiment 1 and 0.254% in experiment 2. Combining these with the above $\Delta V$ values (using $c = 5$ in Equation 1) gives $U_{BM} = 0.071$ and $S_{BM} = 0.148$ for experiment 1 and $U_{BM} = 0.088$ and $S_{BM} = 0.145$ for experiment 2. These mutation rate estimates are similar to those reported by Fry *et al.* (1999; note that the mutation rate estimates in that article are not scaled to the entire genome) and much lower than the "classic" estimates of 0.30–0.85 (Mukai 1964; Mukai *et al.* 1972; Ohnishi 1977). The $S_{BM}$ estimates are likewise similar to those reported by Fry *et al.* (1999) and much higher than the classic estimates of 0.023–0.030. In contrast, in the low-density treatment, using the values in Tables 5 and 6 gives $U_{BM} = 0.72$ and $S_{BM} = 0.020$ for experiment 1 and $U_{BM} = 0.21$ and $S_{BM} = 0.035$ for experiment 2. The first $U_{BM}$ estimate is well within the range of the classic estimates, while the second is a bit lower. Both $S_{BM}$ estimates are remarkably similar to the classic estimates.

To test whether some MA lines had viability-increasing mutations, which would cause $U_{BM}$ to severely underestimate $U$, means of each MA line in each treatment were contrasted to that of the pooled controls. The lack of significant variation among the control lines in most

**TABLE 4**

**Comparisons of among-line variances between treatments (quasinormal lines only)**

| Experiment | Generation | Treatments compared | $-2\log(\lambda)$[a] | $P$ |
|---|---|---|---|---|
| 1 | 27 | Standard–low dens. | 8.07 | 0.005 |
| | | Standard–low temp. | 0.44 | 0.51 |
| | | Low dens.–low temp. | 1.85 | 0.17 |
| 2 | 31 | Standard–low dens. | 7.28 | 0.007 |
| | | Standard–low temp. | 2.54 | 0.11 |
| | | Low dens.–low temp. | 9.36 | 0.002 |
| 1 | 33 | Standard–ethanol | 0.005 | 0.94 |
| 2 | 35 | Standard–ethanol | 1.30 | 0.25 |

[a] Twice the difference in log-likelihoods between constrained and unconstrained models. Dens., density.

**TABLE 5**

**Estimates of genetic correlations across treatments**

| Experiment | Generation | Treatments | Lines included | Correlation (SE or LB[a]) | *P* for $H_0$: $r_g = 0$ | *P* for $H_0$: $r_g = 1$ |
|---|---|---|---|---|---|---|
| 1 | 27 | Standard–low dens. | QN | 1 (LB = 0.79) | <0.001 | — |
| | | Standard–low temp. | QN | 0.69 (0.22) | 0.009 | 0.061 |
| | | Low dens.–low temp. | QN | 0.83 (0.46) | 0.062 | 0.37 |
| | | Standard–low temp. | QN + SD | 0.54 | — | — |
| | | Low dens.–low temp. | QN + SD | 0.67 | — | — |
| 2 | 31 | Standard–low dens. | QN | 1 (LB = 0.55) | <0.001 | — |
| | | Standard–low temp. | QN | 0.81 (0.21) | 0.003 | 0.17 |
| | | Low dens.–low temp. | QN | 1 (LB = 0.74) | 0.001 | — |
| | | Standard–low temp. | QN + SD | 0.84 | — | — |
| | | Low dens.–low temp. | QN + SD | 1 | — | — |
| 1 | 33 | Standard–ethanol | QN | 1 (LB = 0.81) | <0.001 | — |
| 2 | 35 | Standard–ethanol | QN | 0.78 (0.17) | 0.002 | 0.065 |
| | | Standard–ethanol | QN + SD | 0.91 | — | — |

QN, quasinormal; SD, severely deleterious. Standard errors and significance tests for datasets with severely deleterious lines are not shown, due to violation of the normality assumption.

[a] Lower bound; see MATERIALS AND METHODS.

instances (Table 3) justified pooling them; in the one case where significant variation among the controls was present, the comparisons are liberal. No MA line had viability significantly higher than that of the controls ($P < 0.05$) after adjusting for multiple comparisons. In contrast, an average of four (range 0–8) quasinormal lines had viability significantly lower than that of the controls in the 10 datasets.

## DISCUSSION

In this study, the mutational variance for viability was strongly affected by the parental density of the viability assays in each of two MA experiments, being sevenfold higher in the higher-density standard treatment than in the low-density treatment. In contrast, the rate of mutational decline was unaffected by density, within the limits of detection. Two other environmental manipulations, reducing the temperature and adding ethanol to the medium, had no effect on either the mutational variance or mean decline, after a small number of lines with severely deleterious mutations were excluded.

**TABLE 6**

**Estimates of per generation rate of decline in viability, $\Delta M$ (%)**

| | Standard | Low dens. | Low temp. | Ethanol |
|---|---|---|---|---|
| Quasinormal lines only | | | | |
| Exp. 1 | 0.10 | 0.29 | 0.27 | 0.37 |
| Exp. 2 | 0.27 | 0.15 | 0.31 | 0.15 |
| Average | 0.19 | 0.22 | 0.29 | 0.26 |
| Quasinormal and severely deleterious lines | | | | |
| Exp. 1 | 0.10 | 0.29 | 0.40 | 0.37 |
| Exp. 2 | 0.32 | 0.15 | 0.41 | 0.39 |
| Average | 0.21 | 0.22 | 0.41 | 0.38 |

Our previous lower-bound estimate of the rate of deleterious mutations in experiment 1 was $U_{BM} = 0.105$ (FRY *et al.* 1999, Table 1, after scaling to the entire genome), while the upper-bound estimate of the average mutational effect was $S_{BM} = 0.113$. These estimates were based on pooling data from the low-density and standard treatments at G27 and from the ethanol and standard treatments at G33. While the analysis reported here supports the decision to pool data from the latter two treatments, data from the low density and standard treatments should not have been pooled. Nonetheless, the revised estimates reported here for the high-density treatments in experiment 1, $U_{BM} = 0.071$ and $S_{BM} = 0.148$, are not much different from those reported by FRY *et al.* (1999). This is not surprising, because the low-density treatment contributed only one-quarter of the pooled data in that article. In addition, the new mutational parameter estimates from the high-density treatments in experiment 2 are almost the same as those from experiment 1. Both sets of estimates suggest a lower rate of deleterious mutations, but with stronger individual effects, than in the classic studies (MUKAI 1964; MUKAI *et al.* 1972; OHNISHI 1977). In contrast, when estimates of rates and effects of deleterious mutations are calculated separately for the low-density treatment, a surprising result emerges: The estimates, especially those from experiment 1, are similar to the classic estimates. Possible interpretations of the differences in $U_{BM}$ and $S_{BM}$ between treatments are discussed below.

Our results suggest that the variation among Drosoph-

**TABLE 7**

**Estimates of mutational variance, $\Delta V$**

| | Standard | Low dens. | Low temp. | Ethanol | High density treatments, average[a] |
|---|---|---|---|---|---|
| | | | Quasinormal lines only | | |
| Exp. 1 | 3.52 | 0.59 | 1.66 | 3.77 | 3.11 |
| Exp. 2 | 4.42 | 0.51 | 3.12 | 2.56 | 3.67 |
| Average | 3.97 | 0.55 | 2.39 | 3.17 | 3.39 |
| | | | Quasinormal and severely deleterious lines | | |
| Exp. 1 | 3.52 | 0.59 | 10.39 | 3.77 | 5.39 |
| Exp. 2 | 6.89 | 0.51 | 8.88 | 18.07 | 10.01 |
| Average | 5.20 | 0.55 | 9.63 | 10.92 | 7.70 |

Estimates have been multiplied by $10^4$.

[a] Weighted by the number of lines assayed in each treatment.

ila studies in estimates of mutational parameters for viability could have been caused partly by differences in assay density. A review of $\Delta V$ estimates from studies that used the *Curly* method for estimating viability gives some support for this conclusion (Table 8; see APPENDIX). The estimates from the four studies that used a relatively high density vary widely, from 0.69 to 3.97 $\times$ $10^{-4}$, and overlap those from the two studies using a lower density (0.55–0.83). Among the high-density studies, however, is that of CHAVARRÍAS *et al.* (2001), who used a different method for accumulating mutations, and whose experiment was of much longer duration, than the other studies. One or both of these factors could account for their rather low $\Delta V$ estimate. For example, there is evidence that an appreciable fraction of the mildly deleterious mutations detectable in two Drosophila MA experiments (including exp. 1 here but not exp. 2) was caused by movement of the retrotransposable element *copia* (S. V. NUZHDIN, D. HOULE and J. D. FRY, unpublished data). Because *copia* transposes only in males (PASYUKOVA *et al.* 1997), $\Delta V$ estimated from full-sib MA lines, in which genes spend one-half their time in females, might be expected to be lower than that estimated from Mukai-type lines, in which second chromosomes are passed exclusively through males. With the CHAVARRÍAS *et al.* (2001) dataset excluded, there is some evidence for a density effect: The smallest estimate at high density, that of MUKAI *et al.* (1972), is more than twice the largest estimate at low density, that of OHNISHI (1974, 1977).

On the other hand, density variation did not significantly affect $\Delta M$ in this study, and much of the variation in $U_{BM}$ estimates is due to differences in $\Delta M$ estimates (*e.g.*, CHAVARRÍAS *et al.* 2001, 0.1%; this study, high-density treatments, 0.23%; MUKAI 1964, 0.6%). FRY (2001) discusses factors that may have contributed to this variation; they include different methods for estimating base population viability, strain differences, and sampling error. Most notably, Fry's reanalysis shows that the high $\Delta M$ estimates of MUKAI (1964) and MUKAI *et al.* (1972) were probably not inflated by failure to distinguish mutational and nonmutational changes, as has been argued (KEIGHTLEY 1996; GARCÍA-DORADO 1997; FRY *et al.* 1999).

All conclusions about $\Delta M$ in this study, and hence about $U_{BM}$ and $S_{BM}$, are subject to the caveat that they

**TABLE 8**

**Summary of estimates of $\Delta V$ for second chromosome relative viability (standard Drosophila medium, 25°)**

| Experiment | Type of MA lines | Generations | Density | $\Delta V$ |
|---|---|---|---|---|
| MUKAI (1964), MUKAI and YAMAZAKI (1968) | Chrom. 2 | 32 | High | 2.21 |
| MUKAI *et al.* (1972) | Chrom. 2 | 40 | High | 1.90 |
| OHNISHI (1974, 1977) | Chrom. 2 | 40 | Low | 0.83 |
| CHAVARRÍAS *et al.* (2001) | Full-sib | 250–255 | High | 0.69[a] |
| This study—standard treatment | Chrom. 2 | 27–35 | High | 3.97 |
| This study—low density treatment | Chrom. 2 | 27–31 | Low | 0.55 |

All estimates have been multiplied by $10^4$. Estimates from the Mukai and Ohnishi studies were recalculated as described in the APPENDIX. Chrom., chromosome.

[a] Mean of the estimates for generations 250 and 255 in Table 5. This article used a slightly different method for calculating $\Delta V$ (see APPENDIX).

depend on the validity of the control populations. If the control populations had accumulated adaptive or deleterious mutations, $\Delta M$ would have been overestimated or underestimated, respectively. Adaptive mutations would have had to have large effects to reach high frequencies in the 17–26 generations that elapsed between when the control lines were established and when the viability assays were performed. It seems unlikely that large-effect adaptive mutations would have arisen and escaped stochastic loss in the first few generations in all three control lines; adaptive mutations in one or two lines only would have caused divergence between the lines in viability, which was not observed. Some decline in control viability due to deleterious mutations could have occurred, but if the declines were approximately equal in the different treatments, they would not have affected the conclusions about treatment effects on $\Delta M$. One could hypothesize that, because the control populations were maintained at 18° (Fry *et al.* 1999), mutations neutral at this temperature but deleterious at 25° may have accumulated, biasing the $\Delta M$ estimates at 25° downward. Such mutations do not appear to occur at a high rate, however (Figure 1). Fry *et al.* (1999) give additional evidence that mutations in the control lines probably had little effect on the conclusions of experiment 1. Probably the most serious limitation of the controls is that only three control populations were kept in each experiment; as a result, control means were estimated with higher standard errors than MA line means (Table 1).

Cross-environment correlations of viability were generally close to one. Our results therefore do not support the conclusion that there is a high rate of mutations with conditionally deleterious effects, as required by some models of the evolution of ecological specialization (Kawecki 1994; Fry 1996; Kawecki *et al.* 1997). We realize, however, that the range of environments we used is small compared to that in nature. As is shown below, the high cross-environment correlations give information on how much $U$ may have differed between densities.

**Does density affect *U*?:** Having two sets of estimates of $U_{BM}$, one for each density, raises the question of which set is to be preferred when drawing conclusions about the evolutionary impact of deleterious mutations. We consider three possible interpretations of the disparate estimates. The interpretations are based on the well-known relationships (*e.g.*, Mukai *et al.* 1972) between the estimable quantities $\Delta M$ and $\Delta V$ and the underlying parameters $U$, $\bar{s}$ and $\sigma_s^2$, where the latter two parameters are the mean and variance, respectively, of mutational effects,

$$\Delta M = U\bar{s} \tag{2a}$$

and

$$\Delta V = U(\bar{s}^2 + \sigma_s^2) = U\bar{s}^2(1 + C^2). \tag{2b}$$

Here, $C$ is the coefficient of variation of mutational effects $\sigma_s/\bar{s}$.

The first possibility is that the point estimates of $U_{BM}$ and $S_{BM}$ are correct for all practical purposes. This requires that mutations had constant effects ($C = 0$) within a given environment; otherwise, $U_{BM}$ will underestimate $U$, and $S_{BM}$ will overestimate $\bar{s}$. The problem with this hypothesis is that it is not possible to reconcile the large difference in mutation rates between environments implied by the $U_{BM}$ estimates with the observed high cross-environment correlations for viability (Table 5). If $U_i$ and $s_i$ are rates and effects of mutations as measured in environment $i$ (1 = low density, 2 = high density), then the highest cross-environment correlation occurs when the mutations that affect viability in the high-density treatment are a subset of those that affect viability in the low-density treatment. In this case, $U_2$ is the rate of mutations with effects in both treatments, while $U_1 - U_2$ is the rate of mutations that affect viability only in the low-density treatment. The cross-environment covariance among MA lines is then $U_2 s_1 s_2$, and the cross-environment correlation is

$$r_g = \frac{U_2 s_1 s_2}{\sqrt{U_1 s_1^2 U_2 s_2^2}} = \sqrt{\frac{U_2}{U_1}}. \tag{3}$$

For experiment 1, the predicted $r_g$ is 0.31, well below the lower bound of 0.79 for $r_g$ between the low-density and standard treatments (Table 5). For experiment 2, the predicted correlation is 0.65, which is above the lower bound for $r_g$ between the low-density and standard treatments, but below that between the low-density and low-temperature treatments (Table 5).

A second possibility is that the point estimates of $\Delta M$ and $\Delta V$ are accurate, but that mutation effects were not constant, leading to underestimation of $U$ and overestimation of $\bar{s}$ at one or both densities. To see whether this hypothesis can be reconciled with the high cross-environment correlations, we make the nonrestrictive assumption that $U$ is the same between densities; if a subset of mutations has no effect at one density, this can be reflected in lower $\bar{s}$ and higher $C$. The cross-environment correlation can be derived by considering an expression analogous to (2b) for the mutational covariance across environments. After some rearrangement, we obtain

$$r_g = \frac{1 + r_m C_1 C_2}{\sqrt{(1 + C_1^2)(1 + C_2^2)}}. \tag{4}$$

Here, $r_m$ is the cross-environment correlation of effects of individual mutations. The mutational variances provide information on $C_1$ and $C_2$. Under the assumption that there was no difference in $\Delta M$ between densities, $\bar{s}$ at the different densities must have been equal (by Equation 2a); it follows that any differences in $\Delta V$ were due to differences in $C$ (Equation 2b). Specifically, the ratio $\Delta V_2/\Delta V_1$ is given by $(1 + C_2^2)/(1 + C_1^2)$. Values of

this ratio from Table 7 (5.3 for experiment 1, 7.2 for experiment 2) can be used to give an expression for $C_2$ in terms of $C_1$. After substituting this into Equation 4 and taking $r_m = 1$, we solved numerically for the minimum value of $C_1$ consistent with the lower bounds for $r_g$ from Table 5 (0.79 for experiment 1, 0.74 for experiment 2). This gives $C_1 = 0.58$ and 0.55 for the two experiments. Because $U_{BM}$ estimates $U/(1 + C^2)$, under this model $U$ in the low-density treatment would have to be higher than the estimated $U_{BM}$ values; the corrected $U$ values are 0.96 and 0.27 in experiments 1 and 2, respectively. These are also estimates of $U$ in the high-density treatments under this model; the great disparity between $U$ and $U_{BM}$ for these treatments results from the high values of $C_2$ implied by the ratio $\Delta V_2/\Delta V_1$.

Equation 4 generalizes Equation 7 of Keightley *et al.* (2000), which was derived by assuming that mutational effects have a gamma distribution. The equivalence can be seen by noting that β, the shape parameter for the gamma distribution, is equal to $C^{-2}$. Keightley *et al.* (2000) correctly point out that $r_g$ will be positive even when mutational effects are uncorrelated across traits or environments ($r_m = 0$); this occurs because MA lines differ in the number of mutations they carry. Those authors incorrectly stated, however, that $r_g$ will exceed $r_m$ as long as $r_m > 0$. From (4), it is apparent that the relationship between $r_g$ ($Y$) and $r_m$ ($X$) is linear, with $Y$-intercept and slope each >0 and <1. The line passes through the point $r_g = r_m$ when $r_m$ equals

$$r_m^* = \frac{1}{-C_1C_2 + \sqrt{(1 + C_1^2)(1 + C_2^2)}}. \qquad (5)$$

When $r_m < r_m^*$, $r_g > r_m$, but when $r_m > r_m^*$, $r_g < r_m$. Note that if $C_1 = C_2$, $r_m^* = 1$, so $r_g$ will never be $< r_m$. When $C_1$ and $C_2$ are very unequal, however, it is possible for $r_g$ to be $\ll r_m$. An intuitive explanation for this is that, if $C_2 \gg C_1 \approx 0$, much of the variance among MA lines in environment 2, but not environment 1, will be due to variation in mutational effects (see Equation 2b). This component of variation in environment 2 will be only weakly correlated with variation in environment 1, which will be caused mostly by variation in the numbers of mutations carried by the different MA lines.

A third possibility is that $\Delta M$ and/or $\Delta V$ were estimated with substantial error, in such a way as to inflate the differences in $U_{BM}$ and $S_{BM}$ between treatments. The $\Delta V$ estimates seem fairly robust; they were consistent between experiments (Table 7), and the estimates for the standard treatment reported here are quite similar to estimates calculated for the same lines using the covariance between viabilities of lines at widely separated times (Fry 2001). (That these covariance estimates were similar to those in Table 7 gives evidence that variances among the MA lines were not inflated by environmental factors, as seemed to occur in one case with the control lines). In contrast, the $\Delta M$ estimates were less consistent between experiments (Table 6) and were based on only three control populations. The weakest link may be the $\Delta M$ estimates in the low-density treatment, which were based on smaller datasets than those from the standard or combined high-density treatments. Overestimation of $\Delta M$ may have contributed to the surprisingly high $U_{BM}$ estimate for the low-density treatment of experiment 1.

In summary, if the estimates of $\Delta M$ and $\Delta V$ are assumed to be accurate, the results support the high estimates of deleterious mutation rates of Mukai and Ohnishi. An alternative possibility, more consistent with the conclusions of Fry *et al.* (1999), is that sampling error inflated the $U_{BM}$ estimates in the low-density treatment.

**Previous work on genotype × environment interaction from new mutations:** Genotype × environment interaction for fitness from new mutations has been investigated in several previous studies, of Drosophila and other organisms. The limited data give no support for a general effect of density on $\Delta V$, but indicate that other types of environmental stress can sometimes increase $\Delta V$ estimates.

Kondrashov and Houle (1994) estimated the relative fitness of two MA lines and a selected control in a series of environments varying in density, dilution of the medium, and temperature. The fitness difference between the MA lines and the control increased substantially with increasing density and dilution. Kondrashov and Houle's results give no information on $\Delta V$, but imply that $\Delta M$ increases with density, contrary to the results reported here. There are at least two possible explanations for this difference in results. The range of densities used by Kondrashov and Houle was 16-fold, compared to only 3-fold in this study, giving them greater power to detect density effects on $\Delta M$. Furthermore, the MA lines had accumulated mutations for about twice as long as in the current study, again probably resulting in greater power.

Studying *D. melanogaster*, Fry *et al.* (1996) found no evidence that density affected $\Delta V$ for a fitness measure that integrated viability and fertility. Mean-standardized $\Delta V$ was greater when assays were performed at low temperature, but this was because of lower mean performance at the low temperature, rather than an increase in the among-line variance. Because the MA lines had a population size of 10 pairs, selection is likely to have played a much more important role in filtering new mutations in this experiment than in most MA experiments.

Fernández and López-Fanjul (1997) tested full-sib MA lines in four treatments, measuring noncompetitive viability and fecundity under uncrowded conditions. In contrast to Fry *et al.* (1996), they found no tendency for mean-standardized $\Delta V$ to be higher in the environments with lower means (high salt, dilute food). Chavarrías *et al.* (2001) measured second chromosome competitive viability of the same set of lines many generations later

and compared their results to the earlier results on noncompetitive viability (Fernández and López-Fanjul 1996, 1997). Estimates of $\Delta M$ and $\Delta V$ were similar between the competitive and noncompetitive datasets, giving no support for a general effect of density on $\Delta V$. Nonetheless, it is possible that their results would have been different had they assayed MA lines simultaneously in the different treatments, particularly since many of the lines assayed by Fernández and López-Fanjul (1996, 1997) had gone extinct by the time of the Chavarrías *et al.* (2001) study.

Vassilieva *et al.* (2000) compared mutational parameters for several fitness traits in the nematode *Caenorhabditis elegans* between a stressful temperature (12°) and a nonstressful temperature (20°). $\Delta M$ for the intrinsic rate of increase was almost 2-fold higher at the stressful temperature, and $\Delta V$ was >25-fold higher. As a result of the proportionately much larger increase of $\Delta V$, $U_{BM}$ was lower at the more stressful temperature, and $S_{BM}$ was higher. These results thus parallel the results reported here, except that we observed no treatment effects on $\Delta M$.

Szafraniec *et al.* (2001) investigated the effects of deleterious mutations in yeast at a stressful (38°) and nonstressful (30°) temperature. Mutations were induced by allowing single mismatch-repair-deficient cells to replicate for 35–40 generations; the accumulation of mutations was then halted by reintroducing the mismatch repair gene on a plasmid. They observed a dramatically greater mean decline and increase in variance of fitness at the stressful temperature (see also Korona 1999). Interpretation of these results, however, is complicated by the possible role of selection in the 35–40 generations of exponential growth during which mutations were occurring. Kibota and Lynch (1996) show that mutations with relatively large effects ($s \geq 0.1$) are likely to be greatly underrepresented at the end of such an exponential growth phase. In Szafraniec *et al.*'s (2001) experiment, mutations with relatively large effects at 30°, the propagation temperature of the MA lines, would therefore have been filtered by selection. In contrast, mutations with large effects at 38° but not 30° would have been less affected by selection. It is not clear, therefore, whether Szafraniec *et al.*'s (2001) results accurately portray the difference in the distribution of mutational effects at the two temperatures.

Only three of the above studies have reported cross-environment genetic correlations among the MA lines. In the studies of Fry *et al.* (1996) and Vassilieva *et al.* (2000), almost all of the cross-environment correlation estimates were closer to 1 than to 0, in keeping with the estimates reported here. In contrast, Fernández and López-Fanjul's (1997) 15 estimated cross-environment correlations included one significantly negative value, and only two values above 0.33. While this result would appear to give new life to theories of ecological specialization that rely on high rates of conditionally deleterious mutations, Fernández and López-Fanjul's study is the only one of the four in which MA lines were tested nonsimultaneously in the different treatments, with from 7 to 56 generations separating pairs of treatments. It is possible that this downwardly biased the genetic correlation estimates.

## CONCLUSION

The results reported here, as well as those of Vassilieva *et al.* (2000) on *C. elegans*, show that estimates of genomic rates and effects of deleterious mutations can be strongly influenced by assay conditions. In this study, estimated mutation rates and effects from viability assays using a relatively low larval density were similar to the classic Drosophila estimates (Mukai 1964; Mukai *et al.* 1972; Ohnishi 1977) and markedly different from estimates from high-density viability assays (Fry *et al.* 1999; Fry 2001).

Our results indicate that future MA experiments will be most informative when fitness traits are assayed in multiple environments. Proper design and analysis of such experiments is essential. Comparisons of mutational parameters between treatments are most meaningful when MA lines are assayed simultaneously in the different treatments. As the experience here shows, when cross-environment correlations are close to one, biologically significant differences in among-line variances may be present without giving rise to a significant genotype × environment interaction term in an analysis of variance. Likelihood methods or the bootstrap are therefore needed for comparing mutational variances between treatments.

We thank S. V. Nuzhdin for the polytene analysis and S. Durham and R. Miller for their SAS insights. R. Shaw and two anonymous reviewers made helpful comments on an earlier draft. This work was supported by National Science Foundation grants DEB-9317754 and DEB-9707470.

## LITERATURE CITED

Bateman, A. J., 1959 The viability of near-normal irradiated chromosomes. Int. J. Radiat. Biol. **1:** 170–180.

Charlesworth, B., 1990 Mutation-selection balance and the evolutionary advantage of sex and recombination. Genet. Res. **55:** 199–221.

Charlesworth, B., A. Lapid and D. Canada, 1992 The distribution of transposable elements within and between chromosomes in a population of *Drosophila melanogaster*. I. Element frequencies and distribution. Genet. Res. **60:** 103–114.

Charlesworth, B., M. T. Morgan and D. Charlesworth, 1993 The effect of deleterious mutations on neutral molecular variation. Genetics **134:** 1289–1303.

Charlesworth, D., M. T. Morgan and B. Charlesworth, 1990 Inbreeding depression, genetic load, and the evolution of outcrossing rates in a multi-locus system with no linkage. Evolution **44:** 1469–1489.

Chavarrías, D., C. López-Fanjul and A. García-Dorado, 2001 The rate of mutation and the homozygous and heterozygous mutational effects for competitive viability: a long-term experiment with *Drosophila melanogaster*. Genetics **158:** 681–693.

Cockerham, C. C., 1963 Estimation of genetic variances, pp. 53–94 in *Statistical Genetics and Plant Breeding*, edited by W. D. Hanson

and H. F. Robertson. National Academy of Sciences-National Research Council, Washington, DC.
Crow, J. F., and M. J. Simmons, 1983 The mutation load in Drosophila, pp. 1–35 in *The Genetics and Biology of Drosophila*, Vol. 3c, edited by M. Ashburner, H. L. Carson and J. N. Thompson, Jr. Academic Press, New York.
Fernández, J., and C. López-Fanjul, 1996 Spontaneous mutational variances and covariances for fitness-related traits in *Drosophila melanogaster*. Genetics **143:** 829–837.
Fernández, J., and C. López-Fanjul, 1997 Spontaneous mutational genotype-environment interaction for fitness-related traits in *Drosophila melanogaster*. Evolution **51:** 856–864.
Fry, J. D., 1996 The evolution of host specialization: Are trade-offs overrated? Am. Nat. **148:** S84–S107.
Fry, J. D., 2001 Rapid mutational declines of viability in *Drosophila*. Genet. Res. **77:** 53–60.
Fry, J. D., S. L. Heinsohn and T. F. C. Mackay, 1996 The contribution of new mutations to genotype-environment interaction for fitness in *Drosophila melanogaster*. Evolution **50:** 2316–2327.
Fry, J. D., P. D. Keightley, S. L. Heinsohn and S. V. Nuzhdin, 1999 New estimates of the rates and effects of mildly deleterious mutation in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. USA **96:** 574–579.
Gabriel, W., and R. Bürger, 1994 Extinction risk by mutational meltdown: synergistic effects between population regulation and genetic drift, pp. 69–84 in *Conservation Genetics*, edited by V. Loeschke, J. Tomiuk and S. K. Jain. Birkhauser, Basel, Switzerland.
García-Dorado, A., 1997 The rate and effects distribution of viability mutation in *Drosophila*: minimum distance estimation. Evolution **51:** 1130–1139.
García-Dorado, A., J. L. Monedero and C. López-Fanjul, 1998 The mutation rate and the distribution of mutational effects of viability and fitness in *Drosophila melanogaster*. Genetica **102/103:** 255–265.
Gibson, J. B., T. W. May and A. V. Wilks, 1981 Genetic variation at the alcohol dehydrogenase locus in *Drosophila melanogaster* in relation to environmental variation: ethanol levels in breeding sites and allozyme frequencies. Oecologia **51:** 191–198.
Haldane, J. B. S., 1956 The estimation of viabilities. J. Genet. **54:** 294–296.
Jones, J. S., J. A. Coyne and L. Partridge, 1987 Estimation of the thermal niche of *Drosophila melanogaster* using a temperature-sensitive mutant. Am. Nat. **130:** 83–90.
Kawecki, T. J., 1994 Accumulation of deleterious mutations and the evolutionary cost of being a generalist. Am. Nat. **144:** 833–838.
Kawecki, T. J., 1997 Sympatric speciation driven by deleterious mutations with habitat-specific effects. Evolution **51:** 1751–1763.
Kawecki, T. J., N. H. Barton and J. D. Fry, 1997 Mutational collapse of fitness in marginal habitats and the evolution of ecological specialisation. J. Evol. Biol. **10:** 407–429.
Keightley, P. D., 1996 Nature of deleterious mutation load in Drosophila. Genetics **144:** 1993–1999.
Keightley, P. D., E. K. Davies, A. D. Peters and R. G. Shaw, 2000 Properties of ethylmethane sulfonate-induced mutations affecting life-history traits in *Caenorhabditis elegans* and inferences about bivariate distributions of mutational effects. Genetics **156:** 143–154.
Kibota, T. T., and M. Lynch, 1996 Estimate of the genomic mutation rate deleterious to overall fitness in *E. coli*. Nature **381:** 694–696.
Kondrashov, A. S., 1988 Deleterious mutations and the evolution of sexual reproduction. Nature **336:** 435–441.
Kondrashov, A. S., and D. S. Houle, 1994 Genotype-environment interactions and the estimation of the genomic mutation rate in *Drosophila melanogaster*. Proc. R. Soc. Lond. Ser. B **258:** 221–227.
Korona, R., 1999 Genetic load of the yeast *Saccharomyces cerevisiae* under diverse environmental conditions. Evolution **53:** 1966–1971.
Lande, R., 1995 Mutation and conservation. Conserv. Biol. **9:** 782–791.
Latter, B. D. H., and J. A. Sved, 1994 A reevaluation of data from competitive tests shows high levels of heterosis in *Drosophila melanogaster*. Genetics **137:** 509–511.
Littell, R. C., G. A. Milliken, W. W. Stroup and R. D. Wolfinger, 1996 *SAS System for Mixed Models*. SAS Institute, Cary, NC.
Lynch, M., J. Conery and R. Burger, 1995 Mutation accumulation and the extinction of small populations. Am. Nat. **146:** 489–518.
Lynch, M., J. Blanchard, D. Houle, T. Kibota, S. Schultz *et al.*, 1999 Perspective: spontaneous deleterious mutation. Evolution **53:** 645–663.
Mukai, T., 1964 The genetic structure of natural populations of *Drosophila melanogaster*. I. Spontaneous mutation rate of polygenes controlling viability. Genetics **50:** 1–19.
Mukai, T., and T. Yamazaki, 1968 The genetic structure of natural populations of *Drosophila melanogaster*. V. Coupling-repulsion effect of spontaneous mutant polygenes controlling viability. Genetics **59:** 513–535.
Mukai, T., S. I. Chigusa, L. E. Mettler and J. F. Crow, 1972 Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster*. Genetics **72:** 335–355.
Ohnishi, O., 1974 Spontaneous and ethyl methanesulfonate induced polygenic mutations controlling viability in Drosophila melanogaster. Ph.D. Dissertation, University of Wisconsin, Madison, WI.
Ohnishi, O., 1977 Spontaneous and ethyl methanesulfonate-induced mutations controlling viability in *Drosophila melanogaster*. II. Homozygous effect of polygenic mutations. Genetics **87:** 529–545.
Pasyukova, E., S. Nuzhdin, W. Li and A. J. Flavell, 1997 Germ line transposition of the *copia* retrotransposon in *Drosophila melanogaster* is restricted to males by tissue-specific control of *copia* RNA levels. Mol. Gen. Genet. **255:** 115–124.
Peck, J. R., 1994 A ruby in the rubbish: beneficial mutations, deleterious mutations and the evolution of sex. Genetics **137:** 597–606.
SAS Institute, 1992 *SAS Technical Report P-229, SAS/STAT Software: Changes and Enhancements, Release 6.07*. SAS Institute, Cary, NC.
Shaw, R. G., D. L. Byers and E. Darmo, 2000 Spontaneous mutational effects on reproductive traits in *Arabidopsis thaliana*. Genetics **155:** 369–378.
Szafraniec, K., R. H. Borts and R. Korona, 2001 Environmental stress and mutational load in diploid strains of the yeast *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **98:** 1107–1112.
Vassilieva, L. L., A. M. Hook and M. Lynch, 2000 The fitness effects of spontaneous mutations in *Caenorhabditis elegans*. Evolution **54:** 1234–1246.

Communicating editor: R. G. Shaw

## APPENDIX

There are at least two problems in comparing $\Delta V$ estimates among the Drosophila second chromosome MA studies.

First, two measures of viability have been used, either relative viability as given by Equation 2 (Mukai *et al.* 1972; Fry *et al.* 1999; Chavarrías *et al.* 2001; Fry 2001) or the percentage of wild-type flies (Mukai 1964; Ohnishi 1977). As Latter and Sved (1994) point out, the latter underestimates the true fitness difference between genotypes. For example, consider two wild-type chromosomes, with homozygous viability relative to *Cy*/+ heterozygotes of 0.8 and 1.0, respectively. The expected proportions of wild-type flies emerging from crosses between *Cy*/+ heterozygotes will be $(0.8/3)/(0.8/3 + 2/3) = 0.286$ for the first chromosome and $1/3 = 0.333$ for the second chromosome. The ratio of these two proportions is $\sim$0.86, giving the first chromosome higher relative fitness than it should have. Because the percentage of wild-type flies leads to underestimation of $\Delta V$, the estimates presented by Mukai (1964) and Ohnishi (1977) need to be converted to the relative viability scale.

**TABLE A1**

**Conversion of Mukai's (1964) data to the relative viability scale**

| Generation | No. of QN lines | Mean, %WT | Mean, RV | $V_{lines}$, %WT | $V_{lines}$, RV |
|---|---|---|---|---|---|
| 10 | 98 | 31.60 | 0.924 | 0.698 | 0.00111 |
| 15 | 97 | 27.89 | 0.774 | 2.092 | 0.00269 |
| 20 | 89 | 30.93 | 0.896 | 1.848 | 0.00282 |
| 25 | 84 | 28.36 | 0.792 | 3.662 | 0.00484 |
| 32[a] | 80 | 28.04 | 0.779 | 5.647 | 0.00733 |

QN, quasinormal; %WT, percentage wild type; RV, relative viability; $V_{lines}$, variance component among lines.

[a] The generation 32 data come from Mukai and Yamazaki (1968).

A second problem is that the various authors have not been consistent in the practice of standardizing $\Delta V$ estimates by the mean viability. We have standardized our $\Delta V$ estimates by the mean of the controls (Fry *et al.* 1999; Fry 2001; this article). In contrast, Chavarrías *et al.* (2001) log transformed their data before analysis, which in effect causes their $\Delta V$ estimates to be standardized by the mean of the MA lines (see below). Finally, the other authors (Mukai 1964; Mukai *et al.* 1972; Ohnishi 1977) did not standardize their $\Delta V$ estimates at all.

Using data available in the previous articles, we recalculated $\Delta V$ for relative viability, standardizing by the control mean, as described here.

**Mukai (1964):** Means and among-line variances for the proportion of wild-type flies are shown in Table A1. The means can be converted to means of relative viability by the formula

$$\overline{\mathrm{RV}} = \frac{2\overline{\%\mathrm{WT}}}{100 - \overline{\%\mathrm{WT}}} \qquad \text{(A1)}$$

(Fry 2001). To derive an approximate conversion for the among-line variances, we used a Taylor series expansion around $\overline{\%\mathrm{WT}}$ to express RV as an approximate linear function of %WT. Taking variances of both sides yields the relationship

$$\mathrm{VAR(RV)} \cong \frac{4\ \mathrm{VAR(\%WT)}}{[10 - \overline{\%\mathrm{WT}}/10]^4}. \qquad \text{(A2)}$$

Applying (A2) to the datasets in this article and to Ohnishi's (1974) data (see below) shows that it tends to overestimate variances of RV by ~15% (range 11–18%). Consequently, to convert Mukai's (1964) variances to the RV scale, Equation A2 was applied, and the results were divided by 1.15. While this conversion is only approximate, it is far better than no conversion, which results in ~40% underestimation of $\Delta V$ (J. D. Fry, unpublished results).

Results of both conversions are shown in Table A1. Regressing among-line variances of relative viability against generation number, forcing the regression through the origin, gives a slope of $1.95 \times 10^{-4}$. On the RV scale, the mean of the "control" lines chosen by the order method (see Fry 2001, Table 1) is 0.939, giving mean-standardized $\Delta V = 2.21 \times 10^{-4}$.

**Ohnishi (1974):** Although Ohnishi (1977) reported mutational parameter estimates on the basis of the percentage of wild-type flies, in his thesis (Ohnishi 1974, Table 11) he obtained $U_{BM} = 0.365$ and $S_{BM} = 0.0337$ using an index proportional to relative viability. (We have expressed the latter estimate as a fraction of the control mean, estimated from Ohnishi's Table 2-a.) These numbers imply $\Delta V = 0.829 \times 10^{-4}$.

**Mukai *et al.* (1972):** Mukai *et al.* (1972) did three 40-generation MA experiments, starting with 50 lines each. Estimated rates of increase of among-line variance were 0.447, 1.167, and 2.287 per generation (all $\times 10^{-4}$). Standardizing these by the three control means (1.0129, 0.9962, and 0.7474) gives $\Delta V$ estimates of 0.44, 1.18, and 4.09 (all $\times 10^{-4}$). On the assumption that the variation among these estimates is due largely to sampling error, we have used their mean in Table 8.

**Chavarrías *et al.* (2001):** These authors reported $\Delta V$ estimates using log-transformed relative viability. For low $\Delta V$ values, this transformation should give estimates similar to those from untransformed data, but standardized by the MA line mean rather than by the control mean. For higher $\Delta V$ values, such as those from the high-density treatments in exps. 1 and 2, it gives substantially higher estimates than do untransformed data (J. D. Fry, unpublished data). The $\Delta V$ estimate reported by Chavarrías *et al.* (2001) is lower than most of those reported by other workers (Table 8), so their method of analysis cannot be an explanation.

Another issue that affects $\Delta V$ estimates is the choice of threshold for declaring a line "quasinormal." Mukai (1964) and Ohnishi (1977) excluded lines with relative viability $<0.5$. Their controls had relative viability $>0.90$, so their criterion is roughly similar to the one used here, which was to eliminate lines with less than one-half the viability of the controls. Mukai *et al.* (1972) used a slightly different criterion, that of excluding lines with viability less than the mean of the MA lines by $>0.3$ (in their case the MA line mean was predicted from the regression of viability against generation number). Applying this slightly more stringent criterion to the data in exp. 1 results in no additional lines being excluded. In contrast, in exp. 2, one additional line would have been excluded from each dataset, except that for the low-density treatment. The resulting $\Delta V$ estimates average nearly 50% lower than those in Table 6, although still more than threefold higher than those in the low-density treatment. While this would not change the major conclusions of this article, it illustrates how $\Delta V$ estimates can be strongly affected by seemingly minor differences in the definition of quasinormal.