INTEGR. COMP. BIOL., 43:387–395 (2003)

Mutation Accumulation, Performance, Fitness[1]

RAYMOND B. HUEY,[2,*] GEORGE W. GILCHRIST,† KIMIORA WARD,‡ LISA MAVES,§ DAVID PEPIN,§§ AND DAVID HOULE§§§

*Department of Biology, Box 351800, University of Washington, Seattle, Washington 95195-1800
†Department of Biology, Box 8795, College of William & Mary, Williamsburg, Virginia 23187
‡Mt. Baker-Snoqualmie National Forest, 42404 SE North Bend Way, North Bend, Washington 98045
§Institute of Neuroscience, University of Oregon, Eugene, Oregon 97403-1254
§§Department of Biology, Campus Box 1137, Washington University, St. Louis, Missouri 63130
§§§Department of Biological Science, Florida State University, Tallahassee, Florida 32306

SYNOPSIS. The morphology-performance-fitness paradigm is usually explored by determining whether natural or "phenotypically engineered" variation among individuals in morphology (physiology) or performance covaries with an index of fitness such as survival. Here we study between-line covariation between performance and fitness for 44 lines of flies that had undergone mutation accumulation (in the absence of natural selection) on the second chromosome for 62 generations, plus 13 control lines. These mutation accumulation (MA) lines were known to have reduced competitive fitness and life history scores, and to have positive between-line covariances among life history traits. We measured several performance traits of larvae and adults (and a life history trait), examined covariances among those trait means, and also examined covariances of traits with competitive fitness. MA lines had significantly lower performances than did control lines in most traits. However, because control lines had been unknowingly contaminated, a conclusion that MA reduces performance must be tentative. Correlations among performance traits were highly variable in sign, suggesting that MA does not negatively affect all traits equivalently. Even so, correlation matrices for MA and for control lines were very similar. In bivariate comparisons, only one performance trait (a "get-a-grip index," which measures the ability of a falling fly to catch itself on baffles) was positively correlated with competitive fitness. Multivariate analyses again suggested the importance primarily of get-a-grip. Two main patterns emerge from this study. First, MA negatively affects diverse aspects of physiological performance, but does so differentially across traits. Second, except for GAG, MA-induced variation in performance is at best weakly correlated with competitive fitness.

INTRODUCTION

The morphology (physiology)-performance-fitness paradigm (Arnold, 1983) has become a central component of evolutionary physiology. This paradigm is usually investigated by scoring natural or phenotypically engineered variation among individuals in a trait and then determining whether an individual's performance correlates with its survival (Jayne and Bennett, 1990; Sinervo *et al.*, 1992; Schmitt *et al.*, 1999). A complementary way to explore associations between traits and fitness involves using genetic protocols to allow mutations to accumulate in independent replicate lines across generations and then to determine whether among-line variation in traits such as performance correlates with variation in fitness (Houle *et al.*, 1994*b*). This method has two main advantages over traditional approaches. First, it may provide greater power to test the relationship between performance traits and fitness because mutation accumulation can accentuate natural phenotypic variation. Second, it provides an opportunity to explore the role of mutation in the evolution of physiological performance.

Here we study the performance and fitness of a large set of lines of *Drosophila melanogaster* that underwent spontaneous mutation accumulation ("MA") for 62 generations in the absence of natural selection. We examine whether MA lines have both reduced performance and reduced competitive fitness relative to control lines (*i.e.*, lines not accumulating mutations), as would be expected if most mutations are deleterious.

Mutations are the ultimate source of new genetic variation upon which selection can operate; yet most mutations are thought to have deleterious effects on fitness, presumably by reducing an individual's ability to perform key activities. However, the impact of individual mutations is technically difficult to measure because most mutations are thought to be partially recessive and of small effect. Not surprisingly, the empirical linkages among mutation, performance, and fitness are not well delineated. If, however, spontaneous mutations are allowed to accumulate over generations, the composite effect of multiple mutations might eventually be large and measurable. In *Drosophila melanogaster*, such mutation accumulation is readily accomplished via a multi-generation crossing scheme whereby any new mutations (except dominant lethals, of course) on the large, second chromosome are shielded from natural selection and recombination and thus will necessarily accumulate across generations ("mutation accumulation": Dobzhansky *et al.*, 1952; Wallace, 1956; Mukai, 1964; Simmons and Crow, 1977). Eventually the second chromosomes can be extracted from the MA lines and made homozygous. The combined

[1] From the Symposium *Selection and Evolution of Performance in Nature* presented at the Annual Meeting of the Society for Integrative and Comparative Biology, 4–8 January 2003, at Toronto, Canada.

[2] Corresponding author; E-mail: hueyrb@u.washington.edu

impact of multiple mutations on traits of interest can be studied in such lines.

In an early set of studies, Terumi Mukai and colleagues (Mukai, 1964; Mukai *et al.*, 1972; see also Simmons and Crow, 1977) generated many independent MA lines, estimated mutation rates, and found that MA had major deleterious effects on life history traits and on fitness. Brian Charlesworth subsequently founded a new set of MA lines to test theories of life history evolution (Medawar, 1952; Charlesworth, 1994). The resulting lines were initially screened after 33 and 44 generations of MA, and second chromosomes were extracted from each line (Houle *et al.*, 1994*b*) and made homozygous. Each line was then assayed for several life history traits (fecundity, longevity, productivity, male mating ability) as well as for relative competitive fitness against a marker stock (Sved, 1971). Houle *et al.* (1992) reported that MA lines had lower competitive fitness than the control lines. However, they later questioned this conclusion after discovering that some control lines had been inadvertently contaminated with another second chromosome (Houle *et al.*, 1994*a*, *b*, p. 775). Nevertheless, Houle *et al.* (1994*b*) found strongly positive co-variances among life-history traits and between all life-history traits and competitive fitness for the MA lines alone, which were probably uncontaminated. This pattern suggests that life history scores are closely associated with competitive fitness. Further studies of a small subset of these lines at generation 52 showed that high densities greatly accentuated the detrimental effect of MA on fitness (Kondrashov and Houle, 1994).

Previous studies of MA lines have focused largely on fitness and on life history traits (Houle *et al.*, 1994*a*; Kondrashov and Houle, 1994). Our studies focus instead on various ''performance'' traits (*e.g.*, larval feeding rate, adult walking speed, see below), thereby allowing us to test presumed functional links between organismal performance and fitness (Bartholomew, 1964; Huey and Stevenson, 1979; Arnold, 1983). We designed experiments to investigate two questions. First, does MA negatively affect all performance traits in a given line, as might occur if MA leads to an overall reduction in organismal vitality? Houle *et al.* (1994*b*) found strong and positive covariances among line scores in diverse life history traits for the MA lines and concluded that the negative impact of mutations was general across key life history traits. Second, does performance map onto competitive fitness? Thus does a MA line with relatively low performance also have relatively low competitive fitness, as would be expected if performance influences fitness? Houle *et al.* (1994*b*) found strong support for a link between life history and competitive fitness, as these traits were positively correlated among lines.

We screened several performance traits (of larvae or adults) that have diverse ecological significance. *Larval feeding rate* influences larval growth rate and can be a good index of larval competitive fitness (Joshi and Mueller, 1988). For populations evolving stress resistance, however, selection can favor reduced feeding rates (L. Mueller, personal communication, see Joshi and Mueller, 1996; Fellowes *et al.*, 1999; Borash *et al.*, 2000; Mueller and Joshi, 2000). *Larval crawling speed* and *adult walking speed* are measures of physiological vigor (Crill *et al.*, 1996; Gilchrist, 1996; Gilchrist *et al.*, 1997), and adult male walking speed may influence mating success (Partridge, 1991). *Adult knock-down temperature* (T_{kd}) is the upper temperature at which a fly falls from a Weber column (Huey *et al.*, 1992; Crill *et al.*, 1996; Gilchrist and Huey, 1999). Knock-down temperature is an index of the ability—or perhaps the willingness—of a fly to hang on to a substrate at high temperature. However, neither knock-down temperature (nor a related measure, ''knock-down time'') correlates positively with ''heat tolerance'' (Hoffmann *et al.*, 1997; Gilchrist and Huey, unpublished), as scored by the ability to survive an acute heat shock. *Get-a-grip* index (GAG) is a new index that measures the ability of a falling fly to catch itself on baffles when tossed into the top of a ''Weber column'' at normal temperature. We assume it is a measure of overall coordination, but other factors could be involved. Finally, we measured *egg viability*, which is a key life-history trait, because of its contribution to fitness (Roff, 1992; Stearns, 1992).

We measured most traits on 44 MA and on 13 control lines. For reasons given below, we started measuring GAG partway through the experiments and thus measured GAG on only a subset of lines (N = 20 and 7, respectively).

After completing all experiments, we learned that some of the control lines—but apparently not the MA lines—had been contaminated (see Houle *et al.*, 1994*b*, p. 775). Consequently, any conclusions based on our control-line data are suspect: high performance and fitness scores for control lines (relative to MA lines) might thus in part reflect out-crossing enhancement of performance in the control lines (previously inbred) rather than mutational degradation of the MA lines. Accordingly, although we report patterns for both control and MA lines, we focus primarily on patterns involving only variation among the mutation accumulations.

METHODS AND MATERIALS

We obtained 44 lines of flies that had accumulated mutations on the second chromosome, and 13 lines of control stocks. The derivation and maintenance of these lines is detailed in Houle *et al.* (1994*b*). At generation 62, the second chromosome was extracted twice from each line and used to found replicate extraction sublines (''yellow,'' ''white'') for each line. For most traits we scored only the ''white'' sublines. For egg viability, however, we scored both sublines.

Before beginning these experiments, we maintained flies from each line in glass vials (*ca.* 50 eggs/vial) with standard media (molasses-agar-yeast-cornmeal-tegosept) at 18°C on a 12:12 L:D cycle. Because we

were unable to test simultaneously all traits in all lines, we grouped the lines into sets of six lines (generally 4 mutant and 2 control lines) and tested the sets sequentially. All measurements were completed within nine months.

The six test lines for each set were selected haphazardly and transferred to 25°C, which appears to be an optimal developmental temperature for *D. melanogaster* (Cohet, 1975; Zamudio *et al.*, 1995) and also was the long-term rearing temperature for the source flies and the balancer stock (Houle *et al.*, 1994*b*). Eggs were collected (*ca*. 50 eggs/vial) and transferred to vials with standard media for development. The six lines were expanded for two generations (25°C) prior to testing.

Measurements of traits

Larval feeding rate. To score feeding rate of third-instar larvae feeding on yeast, we followed procedures in Joshi and Mueller (1988). Eggs were raised (above) at controlled density (25°C). About 70 hr after hatching, larvae from each line were haphazardly selected from several vials and gently transferred to a holding dish with a dollop of yeast paste. A larva was then transferred to another petri dish [on a temperature-controlled plate (25 ± 1°C)] with a thin layer of yeast solution. After 90 sec we began counting the number of sclerite retractions over the next minute (Joshi and Mueller, 1988). Three larvae from each line were tested blind and in haphazard order until 9 larvae per line were measured. We analyzed the mean feeding rate from among all nine larvae.

Larval crawling rate. We measured the speed of third-instar larvae crawling on an agar surface (Pereira *et al.*, 1995). Larvae (70 hr post-hatching) were obtained as described above for larval feeding rate. A larva was gently transferred from the holding dish to another petri dish containing agar (25 ± 1°C). A transparent lid was immediately placed over the dish, and movements of the larva were traced on the lid for the next 20 sec. Each larva was timed three times in close succession. The traces were later transferred to acetate, digitized, and used to calculate speed (cm/min). We scored three larvae per line (haphazardly rotating among lines), until nine larvae per line were scored (always blind). We analyzed the mean speed for each line, based on the averages of the mean speeds for the larvae.

Adult walking speed. To measure adult walking speed, we used an apparatus originally designed to screen for optomotor-deficient mutants (Benzer, 1967), but here we used it to fractionate flies by walking speed. Groups of flies (3 to 6 days post-eclosion) were placed into the apparatus (a vertically oriented frame of six sets of paired tubes), knocked down, given 14 sec to walk vertically (flies are negatively geotropic). (N.B. A more efficient way of measuring speed is described in Gilchrist *et al.* [1996].) Flies that reached the top tube were shifted to tube 2, whereas laggards would remain in tube 1. All flies were again ''knocked down,'' and the above sequence was repeated five times. A fast fly could reach tube 6 by the end of the trial, whereas a very slow (or non-moving) fly would still be in tube 1. The six lines in each set were scored in haphazard order, with five trials per line. At the end of each run, individual flies were given a speed score according to the tube they reached (1 through 6). The walking-speed score for each line is the unweighted average speed (across trials and sexes).

Knock-down temperature. To test the high-temperature performance of adult flies, we measured ''knock-down'' temperature (T_{kd}: Huey *et al.*, 1992; Crill *et al.*, 1996; Gilchrist and Huey, 1999). Groups of adult flies (3 to 6 days post-eclosion) were tossed into a temperature controlled, vertical Weber column (glass tube 1-m long, 7 cm wide). The flies usually (but see below) caught themselves on the internal baffles of the column (see Weber, 1988). The column temperature was initially set at 30°C, and then raised about 1°C/min. Eventually the flies became heat stressed and fell from the column, into collecting vials which were changed at ½-°C intervals. Thus the flies were fractionated by ''knock-down'' temperature. Each line was tested once, and we analyzed the mean knock-down temperature (unweighted average of males and females) for each line. Because all flies had the same acclimation exposure and rapid heating rates, an acclimation effect is unlikely.

Get-a-grip index. In knock-down experiments with wild-type flies, virtually all flies (≫99%) that are tossed into the top of the Weber column are able to catch themselves on the internal baffles. However, while doing the T_{kd} experiments (above), we noted many flies from some MA lines fell directly through the column. Eventually we began scoring the proportion of flies in the remaining unstudied lines (20 MA and 7 control lines) that successfully caught themselves on the baffles. This ''get-a-grip'' index (GAG) is probably a measure of overall coordination or perhaps of reaction time. (Note: Green *et al.*, 1986 used a similar technique to screen flightless mutants. However, they dropped flies into a tube without baffles: flies reaching the walls in their experiment must be able to fly as well as to catch themselves.)

Egg viability. We scored egg viability in both yellow and white extractions (sublines) from each line (above). To obtain parental flies, we transferred about 100 flies to a freshly yeasted bottle and allowed them to lay for about two hours, such that development took place at uncontrolled but low density (25°C, 12:12 L:D). Freshly eclosed parental flies were then tossed over into a fresh, yeasted bottle and generally maintained for two to three days, but occasionally as long as nine days (but tossed over every two to three days). We then tossed over parental flies into a new bottle (without media) and allowed them to lay for about two hours on an agar plate (with one drop of acetic acid, a sprinkling of yeast). We scored egg viability as the percentage of eggs that hatched over 48 hr. We typically scored 6 lines in a given set (5 MA, 1 control),

usually with three replicate plates for each subline ($\bar{x}$ = 2.8, range = 2 to 7). The number of eggs per plate averaged 340 (range 11 to 2,110).

To determine whether the observed low egg viability rates in some of the MA lines (see RESULTS) might reflect parental or zygotic effects, we made reciprocal crosses between each of the four MA lines with the highest egg mortality rates with one of the control lines (614y). An individual virgin female (MA or control) plus two males (both either control or MA) were placed in a yeasted vial for 24 hr. We counted the number of eggs produced by a given female over 24 hr and the proportion of eggs that hatched. Low egg production would indicate failure of the males either to mate or to stimulate egg production (or lack of female receptivity or fecundity), whereas low egg viability would indicate a problem with fertilization or early development. A paternal (maternal) effect on viability would be indicated if low viability occurred when the male (female) parent was from a MA line; and a zygotic effect would be suggested if viabilities were instead close to control levels.

Competitive fitness. We used fitness scores that had previously been determined in a competitive assay (Sved, 1971) in which treatment lines were competed against a marker stock (for methods, see Houle *et al.*, 1992). These scores assume a uniform level of viability selection across all lines. Fitness scores were available for 40 of the MA and 8 of the control lines studied herein.

The fitness index used here (Houle *et al.*, 1992) is a composite and is influenced by male competitive ability for mates as well as by female fecundity, larval survival, and larval competitive ability. The relative importance of each component is unknown. However, isogenic males have severely reduced ''virility'' (mating propensity and fertility, Brittnacher, 1981), suggesting that male virility might be a major contributor to fitness here (see also Mueller and Ayala, 1981), as well as fecundity and viability (Kondrashov and Houle, 1994).

Statistical analyses

Trait values were generally not normally distributed. Accordingly, we natural log transformed most traits, but angular-transformed three traits (egg viability, GAG, and fitness). We used these transformations prior to all tests (except non-parametric ones).

For three traits (larval feeding rate, egg viability, and fitness), we had multiple measurements for each line, and so we assessed heterogeneity of lines within treatment via a mixed-model ANOVA (treatment as a fixed effect, line nested in treatment as a random effect). For all traits, we compared means of control *vs.* MA lines via a *t*-test with a Welch modification, as variances were usually different between control and mutation lines (F-tests, see Table 1).

Because control lines were contaminated (INTRODUCTION), we computed correlation matrices (among traits) separately for the MA and the control lines. For the subset of lines for which GAG was measured, we recomputed these correlation matrices. Because data were non-normal, we used Spearman *rho* tests.

To look for multivariate relationships between performance trait scores and competitive fitness, we used analysis of covariance to assess fitness as a function of mutation treatment (MA *vs.* control) with the various performance trait scores as covariates. We ran a separate analysis for the subset of lines in which GAG had been measured.

In several cases we computed sets of pair-wise correlations (*e.g.,* between traits and fitness), which increases the probability of a Type I error. The standard correction is to use a Holm (or sequential Bonferroni) adjustment (Holm, 1979; Rice, 1989). However, this adjustment has recently been criticized on mathematical and logical grounds (Moran, 2003). Because this is obviously an unresolved issue, we present both raw *P*-values and corrected significance levels (Benjamini and Hochberg, 1995).

RESULTS

Heterogeneity among lines

For three traits (egg viability, larval feeding rate, and fitness), we were able to assess between-line heterogeneity via a mixed-model ANOVA (see METHODS AND MATERIALS). For all three, treatment (MA *versus* control) was highly significant (Ps $\ll 0.001$, data not shown), and the between-line variation (within treatments) was also highly significant (Ps < 0.001, data not shown). Thus the lines vary significantly, at least in these traits.

Overall effects of MA

Descriptive statistics for MA and control lines, and the results of univariate *t*-tests and F-tests comparing MA and control lines for the various traits are presented in Table 1. MA lines had highly significantly lower scores in most (larval feeding rate, adult knockdown temperature, GAG, egg viability, fitness) but not all (larval crawling rate, adult walking speed) traits. *P*-values for the five significant traits remain significant (Ps < 0.01) after adjustment by a multiple-comparisons test (Benjamini and Hochberg, 1995).

Variances among lines are also significantly greater (or marginally so) for MA than for control lines in four traits (Table 1, except larval feeding, adult walking speeds, and egg viability). *P*-values remain significant (Ps < 0.05) for these four traits following adjustment for multiple comparisons. If the control lines are valid controls (see INTRODUCTION), then MA would clearly have deleterious effects on most performance traits as well as on egg viability and on fitness.

Effects of MA on egg viability

We scored egg viability for both ''white'' and ''yellow'' extractions and also made representative crosses with control lines (see METHODS AND MATERIALS). Average egg viabilities were positively correlated between extractions (''yellow'' *vs.* ''white'' extractions,

TABLE 1. *Mean larval and adult performance (or fitness) for control (C) and mutation accumulation (M) lines.*

				t-test(Welch)[b]				F test[b]		
Trait	Trt	$\bar{x}$[a]	95% CI[a]	N	t	df	*P*[c]	F	df	*P*[c]
Larval performance										
Feeding (bites/min)	C	174.5	172.9, 176.2	13	5.61	20.40	0.000***	0.92	[12, 43]	0.921ns
	M	154.08	153.657, 154.514	44						
Crawling (cm/min)	C	1.15	1.116, 1.177	13	0.59	14.64	0.563ns	2.78	[12, 43]	0.014*
	M	1.1	1.103, 1.113	44						
Adult Performance										
T_{KD} (°C)	C	38.55	38.408, 38.700	13	3.91	14.56	0.001**	2.87	[12, 43]	0.011*
	M	37.30	37.336, 37.382	44						
GAG (proportion)	C	0.997	0.9956, 0.9981	7	8.95	23.57	0.000***	0.05	[6, 19]	0.002*
	M	0.824	0.8124, 0.8355	20						
Walking index	C	2.96	2.803, 3.128	13	0.61	16.71	0.551ns	1.59	[12, 43]	0.261ns
	M	2.75	2.718, 2.785	44						
Egg viability	C	0.93	0.9308, 0.9398	13	9.44	34.13	0.000***	0.35	[12, 43]	0.052ns
	M	0.77	0.7699, 0.7772	44						
Fitness										
W	C	0.539	0.5180, 0.5593	8	7.90	32.53	0.000***	0.12	[7, 39]	0.008*
	M	0.17	0.1654, 0.1812	40						

[a] The mean and 95% confidence intervals were back-transformed from the original transformation (see MATERIALS AND METHODS). N = number of lines, Trt = treatment.

[b] The *t*-test with Welch modification compares the transformed means of C *versus* MA lines for a given trait, and the F test compares variances of the transformed means among lines.

[c] *P*-values in the table are uncorrected for multiple comparisons, but significance levels following correction (see MATERIALS AND METHODS) are indicated next to the *P*-values (ns = $P > 0.05$, * = <0.05, ** = 0.01, *** = <0.001).

MA plus control lines combined; $r = 0.596$, $P < 0.001$): thus egg viability is repeatable within lines (see above). Moreover, viability was significantly lower in MA than in controls ($P < 0.001$, Table 1).

Because some of the MA lines had very low viabilities, we wanted to determine whether this reflected dysfunction of males, of females, or of both (see METHODS AND MATERIALS). For an exploratory test, we reciprocally crossed four sublines having very low viabilities (range 51.8% to 62.2%) with a single control stock (614y [viability = 87.0%]). (Note: Two of the sublines were the white and yellow extractions from line 147 and hence should be genetically similar.) Whether males or females appeared more dysfunctional with respect to egg viability varied unpredictably among crosses (Table 2): in the two crosses involving line 147, a male from a MA line depressed egg viability, but the opposite was the case in the other two crosses (lines 24, 130). In all cases, crosses with the control line enhanced viability relative to crosses within-MA lines (Table 2), suggesting a general zygotic effect.

TABLE 2. *Egg production (over 24 hr) and percent of eggs from crosses involving MA lines.*

Line	Cross	N crosses[a]	Number of eggs $\bar{x} \pm$ SE	Percent viable[b] $\bar{x} \pm$ SE
24 w	♂MA♀MA	3	—	62.2
	♂MA♀C	18	14.9 ± 3.47	83.8 ± 3.85
	♂C♀MA	13	38.8 ± 2.76	69.5 ± 9.86
130 w	♂MA♀MA	3	—	62.2
	♂MA♀C	19	13.3 ± 4.14	90.7 ± 2.65
	♂C♀MA	19	25.2 ± 4.10	83.1 ± 4.30
147 w	♂MA♀MA	3	—	51.8
	♂MA♀C	17	10.2 ± 2.83	70.4 ± 5.83
	♂C♀MA	20	36.0 ± 3.55	90.8 ± 3.55
147 y	♂MA♀MA	3	—	62.2
	♂MA♀C	16	12.7 ± 4.00	71.9 ± 3.95
	♂C♀MA	20	30.9 ± 3.45	84.4 ± 5.57

[a] Each cross involved a single female with two males. Crosses are within MA lines or are reciprocal crosses between a MA line and a control line (614y).

[b] The average percent viability for the control line was 87.0%.

We also monitored egg production (over 24 hr) from these crosses (Table 2). Egg production was cut by half or more if the male was from a MA line (Wilcoxon-Mann-Whitney test stratified by line, $P \ll 0.001$).

Paired correlations among performance (life-history) traits

Table 3 is a correlation matrix for the various larval and adult traits (except fitness) for all lines for which complete data sets were available, whereas Table 4 is a matrix for the subset of lines in which GAG was also measured. In both tables, values above the diagonal represent correlations for MA lines, whereas

TABLE 3. *Matrix of Spearman* rho *correlation coefficients for line means for various traits among all lines.*

	Larval		Adult		
	Feeding	Crawling	T_{KD}	Walking	Egg viability
Larval					
Feeding	—	−0.123[a,b]	0.061	−0.185	−0.172
Crawling	−0.412	—	**−0.343***[,l]	**−0.291**§[,l]	0.157
Adult					
T_{KD}	0.166	**−0.645***[,b]	—	**0.332***	0.050
Walking	−0.093	−0.451	0.465	—	0.015
Egg viability	−0.137	−0.104	0.113	0.319	—

[a] Coefficients for mutants are shown above the diagonal (N = 44), controls are below (N = 13).
[b] Uncorrected *P*-values are §: $P < 0.10$, *: $P < 0.05$. No correlation is significant following correction for multiple comparisons.

those below the diagonal represent correlations for control lines.

Signs of correlation coefficients between traits were inconsistent, indicating that MA does not similarly affect performance in all traits. In fact, only nine of 20 coefficients for the performance or life-history traits were positive in sign for the full data set (Table 3), and only 14 of 30 were positive for the reduced data set (GAG lines only, Table 4). Patterns of inter-trait correlations (Table 3) were, however, very similar between control and MA lines. In fact, correlation coefficients of MA lines are significantly correlated with those from control lines (Spearman $rho = 0.770$, $P < 0.02$, data from Table 3).

Few performance traits were strongly correlated in the full data set (Table 3). For the MA lines, adult knock-down temperature was positively correlated with adult walking speed but negatively correlated with larval crawling speed. For the control subset, for which power is obviously limited, adult knock-down temperature was again negatively correlated with larval crawling speed. However, significance of these three correlations disappears if *P*-values are corrected for multiple tests.

For the reduced data set (GAG lines only, Table 4), power is very limited. The only significant correlation was a negative one between larval crawling speed and larval feeding rate, but this significance disappears following correction for multiple comparisons.

Correlations with fitness

Bivariate correlations generally suggest only weak (at best) associations between the various performance traits and competitive fitness. In the full data set (Table 5, left), fitness was not significantly correlated with any trait in either the MA or the control lines even in the absence of correction for multiple comparisons. In the reduced data set (Table 5, right) following correction, fitness was positively correlated only with GAG ($rho = 0.66$, $P < 0.01$) and with egg viability ($rho = 0.516$, $P < 0.05$); and only the correlation with GAG remained after correction ($P = 0.03$). Neither correlation was significant in the control lines. Note that lack of correlation of fitness with GAG in the control lines undoubtedly reflects the fact that virtually 100% of control flies caught themselves, such that no between-line variation exists (Table 1).

We examined also the multivariate relationship between the performance scores (covariates) and fitness for the MA and control lines (Table 6). We excluded interactions between mutation treatment and performance variables as no interaction was significant. In the full data set, no performance score was significantly correlated with fitness (Table 6, left). However, the MA lines had significantly lower residual fitness than did the controls (ANCOVA: $F[41,1]: 6.491$, $P = 0.015$), consistent with the above univariate analysis. In the reduced data set (Table 6, right), GAG, egg

TABLE 4. *Matrix of Spearman* rho *correlation coefficients for line means among GAG lines.*

	Larval		Adult			
	Feeding	Crawling	T_{KD}	GAG	Walking	Egg viability
Larval						
Feeding	—	**−0.559***[a,b]	−0.027	0.274	0.359	−0.265
Crawling	−0.400	—	−0.128	−0.124	−0.364	0.165
Adult						
T_{KD}	0.051	0.564	—	0.193	0.377	0.214
GAG	**−0.872**§[,b]	0.667	0.026	—	0.215	0.091
Walking	−0.600	0.800	0.667	0.564	—	−0.068
Egg viability	0.300	0.600	0.564	−0.154	0.500	—

[a] Mutants are shown above the diagonal (N = 20), controls below (N = 7).
[b] Uncorrected significance levels are §: $P < 0.10$, *: $P < 0.05$. No correlation remains significant after correction for multiple comparisons.

TABLE 5. *Matrix of Spearman* rho *correlations for fitness and various performance traits in all lines (left) and in GAG lines only (right).*

	Fitness (all lines)		Fitness (GAG lines only)	
	Control[a]	Mutant[a]	Control[b]	Mutant[b]
Larval				
Feeding	−0.262	−0.187	−0.100	−0.074
Crawling	0.262	0.020	0.300	0.133
Adult				
T_{KD}	−0.299	0.044	−0.205	0.300
GAG	—	—	0.051	**0.664***[,c]
Walking	0.143	−0.013	0.300	0.065
Egg viability	0.667	0.256	0.600	**0.516***[,c]

[a] For controls, N = 8, for mutants, N = 40. No correlation is significant even before correcting for multiple comparisons.
[b] For controls, N = 5; for mutants, N = 16.
[c] Only GAG (and only in Mutant lines) remains significant following correction for multiple comparisons ($P = 0.03$).

viability, and walking speed were all significantly correlated with fitness, and the difference in residual fitness between the MA and the Control lines was no longer significant. The model for the reduced data set explained over 85% of the variation.

The differences in patterns with and without GAG (Table 6) could reflect the impact of information on GAG sores, or simply be an artifact of the fact that GAG and non-GAG lines were measured at different times. Consequently, we reran the basic linear model (without GAG scores) but added a factor for GAG *vs.* non-GAG lines. This factor was not significant ($P = 0.33$). Thus the impact of GAG appears real.

DISCUSSION

MA and performance

One motivation for our study (INTRODUCTION) was to determine whether MA negatively influenced multiple performance traits, which would suggest that MA had deleterious effects on overall organismal vitality. Such an influence would be suggested if MA lines had consistently lower performances than did control lines. Performance of MA lines was indeed lower on average than that of controls for all traits, and was significantly so for all traits except larval crawling and adult walking speed (Table 1). In some cases the reductions in performance were striking. For example, on average only about 82% of the MA flies are able to catch themselves when dropped into a Weber column (GAG in Table 1), whereas essentially 100% of the control flies (and many other lines of flies we have tested; Gilchrist and Huey, unpublished data) are similarly able to catch themselves in this apparatus. Thus GAG seems an especially sensitive—and simple—index.

Egg viability of MA lines (unweighted $\bar{x} = 77\%$) was also significantly lower than that of control lines ($\bar{x} = 93\%$, Table 1). Mukai *et al.* (1972) found reduced egg-to-adult viability in their MA lines, suggesting that the reductions observed here may be real and not an artifact of contaminated control lines. Exploratory crosses between a few MA and one control lines suggest that either males or females may be the dysfunctional sex with regards to egg viability (Table 2). However, because viability rates always increased in reciprocal crosses of MA flies with the control line, a zygotic effect is implicated. Interestingly, female fecundity is markedly reduced if a control female mated with a MA male, relative to that for the reciprocal cross. Whether this reduction reflects inadequate male semen (Chapman *et al.*, 1995; Markow, 1996) or rather a behavioral response by females to mutant males is unknown.

The large reciprocal crossing effects we observed suggest that traditional designs using balancers for inferring variation in viability may have missed an important kind of mutational effect. For example, Mukai *et al.* (1972) crossed *Cy*/+ males and females, and scored the relative proportions of +/+ and *Cy*/+ flies to infer the viability of the + chromosome. Parental effects of the + chromosome will affect both offspring genotypes, leaving relative viability unchanged. This suggests that using balancers to estimate viability effects misses an important cause of variation in viability.

The among-line variances of MA lines were generally larger than those of control lines (Table 1). Similar patterns are seen in life history traits (Houle *et al.*, 1994*b*), and reflect a mutational contribution to variance.

TABLE 6. *Linear model of mutation treatment (Control vs. MA) on fitness (ln transformed) on with the performance scores as covariates, with separate models were done for all of the lines (left) and for those in which GAG was measured (right).*

	All lines[a]				GAG lines only[b]			
	Value	SE	*t* value	*P*	Value	SE	*t* value	*P*
(Intercept)	3.998	9.452	0.423	0.675	−4.547	11.216	−0.405	0.692
Trt	**−0.208**	**0.082**	**−2.548**	**0.015**	0.054	0.070	0.777	0.451
Feeding	−0.544	0.489	−1.112	0.273	−0.868	0.517	−1.677	0.117
Crawling	0.015	0.368	0.040	0.968	−0.080	0.418	−0.192	0.850
T_{KD}	−0.250	2.599	−0.096	0.924	2.079	3.032	0.686	0.505
Walking	0.017	0.120	0.139	0.890	**−0.318**	**0.135**	**−2.357**	**0.035**
Egg viab.	0.248	0.332	0.748	0.459	**0.585**	**0.264**	**2.213**	**0.045**
GAG	—	—	—	—	**1.270**	**0.222**	**5.721**	**<0.001**

[a] Multiple R^2: 0.337, $F_{[6,41]} = 3.47$, $P = 0.007$.
[b] Multiple R^2: 0.858, $F_{[7,13]} = 11.20$, $P < 0.001$.

Mutation accumulation thus appears detrimental to diverse measures of performance, except the rates of larval crawling and of adult walking. This supports the widely held notion that most mutations are deleterious. However, two caveats must be considered. First, because some control lines were contaminated (Houle *et al.*, 1994*a*, *b*), the lower performance of the MA lines might alternatively reflect the beneficial effects of outbreeding on control-line performance. Second, even the control stocks used here were inbred and weren't especially vigorous (A. Kondrashov, personal communication). So whether the patterns observed here apply to outbred, vigorous flies remains to be determined. An unambiguous resolution of the question of whether MA reduces performance will, therefore, await further studies using different protocols.

Correlations among traits

A second motivation for our study was to determine whether performance scores of the MA lines were correlated for different traits. If so, then most new (nonlethal) mutations would appear to affect diverse performance traits in similar ways, perhaps by reducing the overall vitality of the flies (Houle *et al.*, 1994*b*). Indeed, life history traits showed such strong and positive covariances among traits (Houle *et al.*, 1994*b*).

In both control and MA lines, correlations between performance traits were generally weak and insignificant even before we corrected for multiple comparisons (Tables 3 and 4), suggesting that mutations do not reduce all performance traits via a generalized vitality effect. This pattern differs from that seen for life history traits, where correlations among lines in life history traits (measured at generation 44) were generally positive and strong (Houle *et al.*, 1994*b*). Thus new mutations seem to have more consistent effects on life history traits than on performance traits.

Mutation accumulation and fitness

Is the relative performance of a line correlated with its relative (competitive) fitness? Among the univariate correlations (Table 5) of the MA lines, only one performance trait (GAG) was significantly correlated with fitness following correction for multiple comparisons (Table 5). In the multivariate comparisons for the full data set (Table 6), none of the performance trait scores was significantly related to fitness. Even so, the significant residual difference between the MA and control lines suggests the influence of an unmeasured factor. For the GAG subset of lines, however, GAG, egg viability, and walking speed all are significantly correlated with fitness; and the residual difference between the MA and control lines was not significant. Overall, GAG seems to be the most sensitive indicator of fitness in the MA lines (Table 6).

CONCLUSIONS

Our results suggest that mutation accumulation in *D. melanogaster* has broad effects on performance traits as it does on life history traits (Houle *et al.*, 1992, 1994*b*; Kondrashov and Houle, 1994). However, patterns involving performance traits differ somewhat from those involving life history traits (fecundity, longevity, productivity, male mating ability). For example, although mutation accumulation significantly depresses all life history traits (Houle *et al.*, 1992, 1994*b*), it significantly depressed only some performance traits (Table 1). There are several possible explanations for this. Perhaps traits are simply influenced only by few loci, such that their mutational target is small. Alternatively, perhaps the values of our performance traits are not maximized by directional selection, but rather are under stabilizing selection. To the extent that this is true, we do not expect a bias in the direction of mutational effects.

Although variation among MA lines in life history traits is closely correlated with competitive fitness (Houle *et al.*, 1994*b*), variation in most performance traits is independent of fitness. In fact, the only performance trait showing a strong link with fitness is the get-a-grip index (GAG), which is probably a measure of coordination. This implies that the links between performance and competitive fitness (Bartholomew, 1964; Huey and Stevenson, 1979; Arnold, 1983) are not as strong as those between life history traits and competitive fitness (Houle *et al.*, 1992, 1994*b*; Kondrashov and Houle, 1994). Perhaps that is not surprising: the effect of performance on fitness will often be relatively indirect and mediated via its impact on the acquisition of energy and survival. Moreover, a weak correlation between performance and fitness would also occur if performance traits start at an intermediate optimum value. In any case, fitness components such as viability and productivity don't always have large effects on competitive fitness (Sved, 1971; Haymer and Hartl, 1983).

ACKNOWLEDGMENTS

We thank B. Charlesworth for giving us access to these lines and M. Lynch for suggesting that we do this study. Research was funded by a SGER grant to RBH, by DEB-9981598 to RBH, and by DEB-9981555/0242313 to GWG. We thank J. Kingsolver for the opportunity to participate in this symposium; and we thank anonymous reviewer and especially L. Mueller for comments and suggestions.

REFERENCES

Arnold, S. J. 1983. Morphology, performance and fitness. Am. Zool. 23:347–361.

Bartholomew, G. A. 1964. The roles of physiology and behaviour in the maintenance of homeostasis in the desert environment. Symp. Soc. exp. Biol. 18:7–29.

Benjamini, Y. and Y. Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. R. Statist. Soc. B. 57:289–300.

Benzer, S. 1967. Behavioral mutants of *Drosophila* isolated by countercurrent distribution. Proc. Nat. Acad. Sci. U.S.A. 58:1112–1119.

Borash, D. J., H. Teotonio, M. R. Rose, and L. D. Mueller. 2000. Density-dependent natural selection in *Drosophila*: Correlations

between feeding rate, development time and viability. J. Evol. Biol. 13:181–187.
Brittnacher, J. G. 1981. Genetic variation and genetic load due to the male reproductive component of fitness in *Drosophila*. Genetics 97:719–730.
Chapman, T., L. F. Liddle, J. M. Kalb, M. F. Wolfner, and L. Partridge. 1995. Cost of mating in *Drosophila melanogaster* females is mediated by male accessory gland products. Nature 373:241–244.
Charlesworth, B. 1994. *Evolution in age-structured populations.* 2nd ed. Cambridge University Press, Cambridge, U.K.
Cohet, Y. 1975. Epigenetic influences on the lifespan of the *Drosophila*: Existence of an optimal growth temperature for adult longevity. Exp. Geront. 10:181–184.
Crill, W. D., R. B. Huey, and G. W. Gilchrist. 1996. Within- and between-generation effects of temperature on the morphology and physiology of *Drosophila melanogaster*. Evolution 50: 1205–1218.
Dobzhansky, T., B. Spassky, and N. Spassky. 1952. A comparative study of mutation rates in two ecologically diverse species of *Drosophila*. Genetics 37:650–664.
Fellowes, M. D. E., A. R. Kraaijeveld, and H. C. J. Godfray. 1999. Association between feeding rate and parasitoid resistance in *Drosophila melanogaster*. Evolution 53:1302–1305.
Gilchrist, G. W. 1996. Quantitative genetic analysis of the thermal sensitivity of locomotory performance curve of *Aphidius ervi*. Evolution 50:1560–1572.
Gilchrist, G. W. and R. B. Huey. 1999. The direct response of *Drosophila melanogaster* to selection on knockdown temperature. Heredity 83:15–29.
Gilchrist, G. W., R. B. Huey, and L. Partridge. 1997. Thermal sensitivity of *Drosophila melanogaster*: Evolutionary responses of adults and eggs to laboratory natural selection at different temperatures. Physiol. Zool. 70:403–414.
Green, C. C., J. C. Sparrow, and E. Ball. 1986. Flight testing columns. Dros. Inf. Ser. 63:141.
Haymer, D. S. and D. L. Hartl. 1983. The experimental assessment of fitness in *Drosophila*. Genetics 104:343–352.
Hoffmann, A. A., H. Dagher, M. Hercus, and D. Berrigan. 1997. Comparing different measures of heat resistance in selected lines of *Drosophila melanogaster*. J. Insect Physiol. 43:393–405.
Holm, S. 1979. A simple sequentially rejective multiple test procedure. Scand. J. Stat. 6:65–70.
Houle, D., D. K. Hoffmaster, S. Assimacopoulos, and B. Charlesworth. 1992. The genomic mutation rate for fitness in *Drosophila*. Nature 359:58–60.
Houle, D., D. K. Hoffmaster, S. Assimacopoulos, and B. Charlesworth. 1994*a*. The genomic mutation rate for fitness in *Drosophila*. Nature 371:358.
Houle, D., K. A. Hughes, D. K. Hoffmaster, J. Ihara, S. Assimacopoulos, D. Canada, and B. Charlesworth. 1994*b*. The effects of spontaneous mutation on quantitative traits. I. Variances and covariances of life history traits. Genetics 138:773–785.
Huey, R. B., W. D. Crill, J. G. Kingsolver, and K. E. Weber. 1992. A method for rapid measurement of heat or cold resistance of small insects. Funct. Ecol. 6:489–494.
Huey, R. B. and R. D. Stevenson. 1979. Integrating thermal physiology and ecology of ectotherms: A discussion of approaches. Am. Zool. 19:357–366.
Jayne, B. C. and A. F. Bennett. 1990. Selection on locomotor performance capacity in a natural population of garter snakes. Evolution 44:1204–1229.
Joshi, A. and L. D. Mueller. 1988. Evolution of higher feeding rate in *Drosophila* due to density-dependent natural selection. Evolution 42:1090–1093.
Joshi, A. and L. D. Mueller. 1996. Density-dependent natural selection in *Drosophila*: Trade-offs between larval food acquisition and utilization. Evol. Ecol. 10:463–474.
Kondrashov, A. S. and D. Houle. 1994. Genotype-environment interactions and the estimation of the genomic mutation rate in *Drosophila melanogaster*. Proc. Roy. Soc. London (Ser. B: Biol. Sci.) 258:221–227.
Markow, T. A. 1996. Evolution of *Drosophila* mating systems. Evol. Biol. 29:73–106.
Medawar, P. B. 1952. *An unsolved problem in biology*. H. K. Lewis, London.
Moran, M. D. 2003. Arguments for rejecting the sequential Bonferroni in ecological studies. Oikos 100:403–405.
Mueller, L. D. and F. J. Ayala. 1981. Fitness and density-dependent population growth in *Drosophila melanogaster*. Genetics 97: 667–677.
Mueller, L. D. and A. Joshi. 2000. *Stability in model populations*: Monographs in Population Biology No. 31. Princeton University Press, Princeton, New Jersey.
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., 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.
Partridge, L. 1991. Variation in mating behaviour and reproductive success of male *Drosophila*. Proc. Int. Symp. Biol. Control Fruit Flies:204–213.
Pereira, H. S., D. E. MacDonald, A. J. Hilliker, and M. B. Sokolowski. 1995. Chaser (Csr), a new gene affecting larval foraging behavior in *Drosophila melanogaster*. Genetics 141:263–270.
Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution 43: 223–225.
Roff, D. A. 1992. *The evolution of life histories.* Chapman and Hall, New York.
Schmitt, J., S. A. Dudley, and M. Pigliucci. 1999. Manipulative approaches to testing adaptive plasticity: Phytochrome-mediated shade-avoidance responses in plants. Am. Nat. 154:S43–S54.
Simmons, M. J. and J. G. Crow. 1977. Mutations affecting fitness in *Drosophila* populations. Annu. Rev. Genet. 11:49–78.
Sinervo, B., P. Doughty, R. B. Huey, and K. Zamudio. 1992. Allometric engineering: A causal analysis of natural selection on offspring size. Science 258:1927–1939.
Stearns, S. C. 1992. *The evolution of life histories.* Oxford University Press, Oxford, U.K.
Sved, J. A. 1971. An estimate of heterosis in *Drosophila melanogaster*. Genet. Res., Camb. 18:97–105.
Wallace, B. 1956. Studies on irradiated populations of *Drosophila melanogaster*. J. Genet. 54:280–293.
Weber, K. E. 1988. An apparatus for measurement of resistance to gas-phase agents. Dros. Inf. Serv. 67:91–93.
Zamudio, K. R., R. B. Huey, and W. D. Crill. 1995. Bigger isn't always better: Body size, developmental and parental temperature, and male territorial success in *Drosophila melanogaster*. Anim. Behav. 49:671–677.