

# Estimates of the Genomic Mutation Rate for Detrimental Alleles in *Drosophila melanogaster*

Brian Charlesworth,<sup>1</sup> Helen Borthwick, Carolina Bartolomé and Patricia Pignatelli<sup>2</sup>

*Institute of Cell, Animal and Population Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom*

Manuscript received December 3, 2003

Accepted for publication March 4, 2004

## ABSTRACT

The net rate of mutation to deleterious but nonlethal alleles and the sizes of effects of these mutations are of great significance for many evolutionary questions. Here we describe three replicate experiments in which mutations have been accumulated on chromosome 3 of *Drosophila melanogaster* by means of single-male backcrosses of heterozygotes for a wild-type third chromosome. Egg-to-adult viability was assayed for nonlethal homozygous chromosomes. The rates of decline in mean and increase in variance (DM and DV, respectively) were estimated. Scaled up to the diploid whole genome, the mean DM for homozygous detrimental mutations over the three experiments was between 0.8 and 1.8%. The corresponding DV estimate was  $\sim 0.11\%$ . Overall, the results suggest a lower bound estimate of at least 12% for the diploid per genome mutation rate for detrimental mutations. The upper bound estimates for the mean selection coefficient were between 2 and 10%, depending on the method used. Mutations with selection coefficients of at least a few percent must be the major contributors to the effects detected here and are likely to be caused mostly by transposable element insertions or indels.

**I**N a higher eukaryote such as *Drosophila melanogaster*, homozygous lethal or sterile mutations form a relatively small component of the total set of deleterious mutations, and the most abundant class is that of detrimental mutations with effects of at most a few percent on fitness (CROW 1993). The answers to many questions in evolutionary biology depend on knowledge of the rates of occurrence and sizes of effects of detrimental mutations (CHARLESWORTH and CHARLESWORTH 1998; LYNCH *et al.* 1999). These are, however, very hard to measure, since it is virtually impossible to detect individual mutations with very small effects on components of fitness. It is thus necessary to resort to indirect, statistically based procedures. If a large number of independent lines with a common origin are maintained for many generations in the effective absence of selection, the rates of decline in mean and increase in variance (DM and DV, respectively) for a fitness-related trait can be estimated from the changes with time in the mean and variance of line means, eliminating lines that have been hit by lethal or severely deleterious mutations. A lower bound to the mean number of detrimental mutations arising per generation is provided by  $DM^2/$

DV, and an upper bound to the mean reduction in trait value per mutation by  $DV/DM$  (BATEMAN 1959; MUKAI *et al.* 1972). This method was first applied in the classic experiments of Mukai and his co-workers on viability mutations on chromosome 2 of *D. melanogaster* (MUKAI 1964; MUKAI *et al.* 1972). A number of mutation-accumulation experiments on *D. melanogaster* and other species have since been performed, using a variety of statistical methods to analyze the results (KEIGHTLEY and EYRE-WALKER 1999; LYNCH *et al.* 1999; FRY and HEINSOHN 2002; SHAW *et al.* 2002; GARCÍA-DORADO and GALLEGRO 2003).

Discrepancies among the conclusions of different investigators, especially with respect to the rates of mutational decline in the mean values of fitness components (KEIGHTLEY and EYRE-WALKER 1999; LYNCH *et al.* 1999; FRY and HEINSOHN 2002; SHAW *et al.* 2002), imply that it is important to repeat the *D. melanogaster* mutation-accumulation (MA) experiments. The work described here involves accumulating mutations on chromosome 3 rather than on chromosome 2, for the following reasons. First, the gene content of chromosome 3 is 18% greater than that of chromosome 2 (ADAMS *et al.* 2000), so that the rate at which mutations accumulate should be proportionately higher, increasing the efficiency of the experiment. Second, the *Cy* marker for the cross-over-suppressor chromosome (balancer) used for chromosome 2 is variable in expression on some genetic backgrounds (HOULE *et al.* 1992), which could bias estimates of DM (KEIGHTLEY and EYRE-WALKER 1999). This problem is avoided here by the use of an efficient chromosome 3 balancer, *TM6*, which carries a marker

**Dedicated to the memory of Terami Mukai, whose pioneering paper on mutation accumulation appeared in GENETICS 40 years ago.**

<sup>1</sup>*Corresponding author:* Institute for Cell, Animal and Population Biology, Ashworth Laboratories, King's Bldgs., University of Edinburgh, W. Mains Rd., Edinburgh EH9 3JT, United Kingdom.  
E-mail: brian.charlesworth@ed.ac.uk

<sup>2</sup>*Present address:* Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, United Kingdom.

(*Ubx<sup>P15</sup>*) with reliable expressivity (LINDSLEY and ZIMM 1992).

Our experiments indicate that there is a detectable decline in mean viability associated with the accumulation of detrimental mutations on chromosome 3 of *D. melanogaster* and that there are likely to be at least 0.12 new detrimental mutations arising per diploid genome per generation. The results do not confirm the very high estimates of the genomic deleterious mutation rate reported by Mukai and co-workers (MUKAI 1964; MUKAI *et al.* 1972), but are substantially higher than some more recent estimates (GARCÍA-DORADO 1997; AVILÁ and GARCÍA-DORADO 2002).

## MATERIALS AND METHODS

**Genetic stocks and breeding designs:** MA was conducted using third chromosomes isolated from the inversion-free derivative of the IV wild-derived stock, which has been maintained as a large random-bred population for >25 years (CHARLESWORTH and CHARLESWORTH 1985). Three replicate experiments, initiated at intervals of ~1 year, were performed. For a given experiment, a set of 50 IV third chromosomes was isolated using standard breeding procedures and crossed onto the genetic background of a stock, IS-4. This was constructed from chromosomes originally extracted from a natural population and has reasonably good viability and fertility (CHARLESWORTH *et al.* 1994). It is isogenic for all three major chromosomes (X, second, and third). It is homozygous for a recessive marker, *spa<sup>hol</sup>*, on the dot chromosome 4, which provides a partial guard against contamination. Since its construction in 1990 until the initiation of MA, IS-4 was maintained by single-pair matings. Its isogenicity was initially tested by screening for 11 families of transposable elements, using *in situ* hybridization of element probes to polytene chromosomes (CHARLESWORTH *et al.* 1994). The chromosome 3 balancer system *TM6/Sb* was crossed onto a background of IS-4, both singly and in combination with the second chromosome balancer system *SM1/Pm* (LINDSLEY and ZIMM 1992). For a given experiment, an IV third chromosome that showed high homozygous viability and fertility on the IS-4 background was chosen to initiate MA, thereby ensuring that new mutations arose on chromosomes representative of those found in an equilibrium population. The *IS-4; TM6/Sb; spa<sup>hol</sup>* stock was rechecked for isogenicity by *in situ* hybridization with the most abundant *D. melanogaster* transposable element, *roo*, using methods described previously (CHARLESWORTH *et al.* 1994).

MA was initiated from 75 lines derived from a single IV chromosome 3 stock by mating three *TM6/Sb* females to one *TM6/MA<sub>i</sub>* male, where *MA<sub>i</sub>* denotes the *i*th MA line for the replicate experiment in question. Two replicate crosses of each MA line were made every generation, and one was selected at random for use in the next generation, except for experiment 2, when two separate lineages for each line were maintained after generation 29, to minimize loss of lines due to the accumulation of lethals. All MA lines were reared in 7.5 × 2.5-cm glass vials on Lewis's medium, at 25 ± 1° in a constant-temperature chamber. In this way, the wild-type third chromosomes from the MA lines were sheltered against selection, except for the rare class of mutations with very large heterozygous effects on fitness (MUKAI 1964).

The MA lines were maintained for 9 or 10 generations before the first assays of their homozygous effects, to allow mutations to accumulate. Assays of homozygous viability at this stage, by the method described below, enabled a check

to be made on the lethal mutation rate (defining lethals as lines with viabilities <10%). The older literature gives the lethal mutation rate for the second chromosome as ~0.5% per generation (CROW and SIMMONS 1983), so that the lethal mutation rate for the third chromosome was expected to be ~0.6% per generation, adjusting for its larger size. In experiment 1, we found that the mutation rate was substantially higher than this, ~2% (see RESULTS). This suggested that there might have been some hybrid dysgenesis, caused by a difference in the transposable element content of the *IS-4; TM6/Sb; spa<sup>hol</sup>* stock from that of the IV stock from which the third chromosomes of the MA lines are derived (BERG and HOWE 1989). We tested for this by examining the fertility at 27.5° of the F<sub>1</sub> reciprocal crosses between the *IS-4; TM6/Sb; spa<sup>hol</sup>* stock and flies carrying the MA third chromosome, but neither direction of the cross gave any evidence for the reduced fertility associated with hybrid dysgenesis (data not shown). This test is not foolproof, since mutations can be induced without hybrid dysgenic fertility in crosses between Q males and M females in the *P*-element system (SIMMONS *et al.* 1980), but we also found no evidence for the high lethal mutation rates that would be expected for autosomal loci in dysgenic crosses (see RESULTS).

We found in pilot experiments that other balancer chromosomes had drastic fitness effects and so decided to use *TM6* for all our viability assays. In addition, *Ubx<sup>P15</sup>* can be reliably scored on flies frozen at -20°, so that counting can be carried out when convenient. All the viability assays reported here were conducted either by one of the authors (B. Charlesworth) or by an undergraduate assistant (Jane Charlesworth), with random cross-checks to ensure comparable accuracy of scoring.

We measured egg-to-adult viabilities of third chromosome homozygotes relative to that of balancer heterozygotes, using the standard method of intercrossing balancer heterozygotes in vial cultures and determining the frequency of *TM6/+* and *+/+* genotypes among the progeny. The viability assays were conducted in four replicate blocks for each MA assay generation of each replicate experiment, with up to five replicate vials of each genotype in each block (an error in record keeping led to only three blocks being set up for generation 22 of experiment 3). Flies were scored at days 17 and 21 after setting up a culture. Males and females were counted separately, but have been pooled for the analyses reported here. Replicate cultures were scored blind. Two mass bottle cultures segregating for *TM6/+* and *+/+* were maintained for each genotype, to provide flies for assays; flies were randomly combined from the two bottles when setting up assays.

The first replicate experiment was assayed at generations 10, 17, 25, and 35. Following the procedure of FRY *et al.* (1999), the other two experiments were split into three subexperiments involving approximately equal numbers of extracted chromosomes. The extractions for the second subexperiment were performed one generation after those for the first one, and the extractions for the third subexperiment were performed one generation later. For the second set of experiments, assays were conducted on subexperiments involving MA generations 9–11, 19–21, 29–31, and 39–41, with the same sets of MA lines (discounting lethal lines and failed lines) being included in subexperiments assayed at successive 10-generation intervals. For the third set, the assays involved MA generations 9–11, 20–22, and 29–31. Totals of 89,322, 236,541, and 189,003 flies were counted in the respective experiments, giving a grand total of 514,866.

To estimate the rate of mutational decline in mean by the Bateman-Mukai method, it is desirable to estimate the mean viability of nonmutated control lines, identical in all other respects to the MA lines. We had originally planned to use a method for freezing fly embryos in liquid nitrogen and then

resurrecting them; this would have provided ideal controls (HOULE *et al.* 1997). After some trials, we had no success in recovering viable larvae and abandoned the method. The alternative method of maintaining a large random-bred stock homozygous for the foundation chromosome as a control population (FRY *et al.* 1999; FRY and HEINSOHN 2002) was not used, since the low starting frequencies of mutations mean that their mutational increase in frequency is initially virtually unopposed by selection (CABALLERO *et al.* 2002), and because of the risk of accidental contamination (HOULE *et al.* 1992).

Instead, we used Mukai's "order method" to identify control lines from among the MA lines (MUKAI 1964). A reanalysis of earlier experiments suggested that the order method provided useful controls for earlier *Drosophila* MA experiments (FRY 2001). The three top-ranking MA lines of a given assay generation were used as the controls for the previous assay generation, on the assumption that these lines carried minimal numbers of deleterious mutations. Only lines measured in all blocks of both assay generations were used for this purpose. For the final assay generation, for which this method is inapplicable, we arbitrarily used the three top-ranking lines from the last block of assays as the controls for the other blocks, omitting the last block from the analyses to avoid bias. Again, only lines present in all blocks were used. However, for reasons given below, the order method probably fails to provide meaningful estimates of the true values of mutation-free genotypes, and so we also estimated rates of decline by regressions of mean viability on generations, without using controls, and by a revision of the order method, which reduces estimation error (see below).

**Statistical analyses:** A given experiment can be viewed as involving a set of genotypes (separate MA lines), assayed over several blocks with several replicates per block. The experiments were not fully balanced, because of unequal numbers of replicates of some genotypes in a given block; occasionally, a genotype was not assayed in all blocks because of crossing failures or errors. Most analyses were carried out using an unbiased estimate of the viability of wild-type (+/+) relative to *TM6/+* heterozygotes: twice the ratio of the number of +/+ to one plus the number of *TM6/+* (HALDANE 1956). Very similar results to those presented below were obtained using Fisher's angular transform of the data, which removes dependence of binomial sampling variance on the mean (SOKAL and ROHLF 1995, pp. 419–422).

In a preliminary analysis, any lines absent in one or more blocks of an assay generation were discarded, and the means for each genotype-block combination were calculated. These provided estimates of block and genotype effects, and the overall distributions of genotypic means for each assay generation were obtained from these estimates. Visual inspection of these distributions allowed lines with unusually low viabilities relative to the rest of the distribution to be picked out; these are candidates for lines containing mutations with major effects on viability. In all cases, these had a mean viability of <50%; in the rest of this article, we treat the remaining lines as "quasi-normals" (MUKAI *et al.* 1972).

Analyses of variance were also conducted using only the lines present in all blocks of an assay generation, on data transformed using the angular transformation of the proportion of wild-type flies. This allowed use of a general linear model (GLM), implemented in the Minitab software package (Release 10; Minitab, State College, PA), to test for genotype, block, and block  $\times$  genotype interactions. Among the 25 ANOVAs conducted in this way, there were frequent examples of highly significant ( $P < 0.01$ ) block and block  $\times$  genotype interactions (see supplemental data at <http://www.genetics.org/supplemental/>). This means that the genotypic components of variance for each assay generation, needed for de-

termining the mutational variance in viability, can be estimated only if block and block  $\times$  genotype interactions are estimated simultaneously. This was done for the full data sets (including lines that are not present in each block), using the following method (KEMPTHORNE 1957, p. 260). First, a one-way components-of-variance ANOVA was applied to a given assay generation, regarding genotype-block combinations as treatments, and treating the variance among replicates within genotype-block combinations as the error component,  $\sigma_c^2$ . This was estimated by equating expected and observed mean squares.

The expectations of mean squares for the full data sets were derived, treating the means for each genotype-block combination as the primary observations, with the model

$$x_{jk} = \mu + g_j + b_k + i_{jk}, \quad (1)$$

where  $x_{jk}$  is the mean over replicates of the combination of line  $j$  with block  $k$ ,  $\mu$  is the overall mean,  $g_j$  is the effect of line  $j$ ,  $b_k$  is the effect of block  $k$ , and  $i_{jk}$  is the interaction effect.

Expectations of the mean squares for genotypes, blocks, and residuals can be written down, treating  $g_j$ ,  $b_k$ , and  $i_{jk}$  as random effects. The variance component corresponding to the latter contains a contribution from the error variance  $\sigma_c^2$  in addition to the true interaction variance  $\sigma_i^2$ ; for a genotype-block combination with  $n_{jk}$  observations, this is equal to  $\sigma_c^2/n_{jk}$ . Using the estimate of  $\sigma_c^2$  from the one-way ANOVA, its contribution to the expected mean squares can be removed; equating the observed and expected mean squares derived from model (1) generates a set of three linear equations, which yield estimates of the components of variance for genotypes, blocks, and interaction: these are  $\sigma_g^2$ ,  $\sigma_b^2$ , and  $\sigma_i^2$ , respectively.

Only the first of these is of interest from the point of view of estimating mutational parameters. For a given assay generation,  $t$ , of an experiment, the mutational rate of increase in variance, DV, was estimated from  $\sigma_g^2(t)/t$ , where  $\sigma_g^2(t)$  is the estimate of genotypic variance for generation  $t$ . The alternative method of estimating DV from the regression of  $\sigma_g^2$  on time was not used, since the individual estimates of  $\sigma_g^2$  were very noisy and no clear trend with time could be observed (see DISCUSSION).

To estimate the mutation rate by the Bateman-Mukai procedure, we also need the estimate of DM, the rate of mutational decline in mean. For data analyzed by the order method (MUKAI 1964), DM for generation  $t$  was estimated from the difference,  $\Delta M$ , between the overall genotypic mean for generation  $t$  and the genotypic mean of the controls (generated by the methods described above):  $DM(t) = \Delta M(t)/t$ .

For data analyzed by the regression method, DM was estimated from the regression of the mean values of genotype-block combinations for each block on  $t$ . In this way, environmental differences associated with blocks are included in the error around the regression line. The regression coefficients for the overall and control means,  $b_M$  and  $b_C$ , respectively, were obtained by the standard least-squares procedure for unweighted regression. For experiments 2 and 3, generation times for each subexperiment (9–11; 19–21, etc.) were assigned as the middle values of each set of assay generations (*i.e.*, as 10, 20, etc.).

Using the Bateman-Mukai formulas (MUKAI *et al.* 1972), the mutational parameters  $u^*$  and  $s^*$  are given by

$$u^* = u/(1 + C^2) = DM^2/DV \quad (2a)$$

$$s^* = \bar{s}(1 + C^2) = DV/DM, \quad (2b)$$

where  $u$  is the mean number of new mutations per haploid third chromosome per generation,  $\bar{s}$  is the mean reduction in viability caused by a mutation, and  $C$  is the coefficient of

variation of the distribution of the viability effects of mutations (weighted by their probability of occurrence).

To provide overall estimates of the mutational statistics for each value of  $t$  for a given experiment, and to obtain sampling variances and confidence intervals, the following bootstrap procedures were implemented (EFRON and TIBSHIRANI 1993). Each of the three replicate experiments was treated as a single entity for this purpose. The set of genotypes assayed in the last generation of the chosen experiment was used for the start of the bootstrap procedure, after removing any genotypes that had not been assayed in previous generations. An equivalent number of genotypes was chosen from among these by sampling randomly with replacement. The one-way and two-way ANOVA procedures described above were then used to obtain estimates of  $\sigma_g^2$  and DV for this generation. Control lines were identified from the last block as the best-performing set of three lines in terms of their genotypic means, omitting any lines that were not present in all four blocks. Using only the remaining blocks, DM was calculated from the difference between the overall genotypic mean and the genotypic mean of the controls.

For the preceding assay generation, the bootstrapped set of genotypes for the final assay generation was used; additional genotypes were then sampled from the set of genotypes in the real data set for this generation, so that the total number of bootstrapped lines for this assay generation was the same as the number in the real data set. In experiments 1 and 3, controls for this penultimate assay generation were assigned as the best-performing three lines for the final assay generation, in terms of their genotypic means (the unweighted means over blocks of the means of genotype-block combinations). Only lines present in all blocks of both generations were used for this purpose. For experiment 2, the small numbers of surviving genotypes in the final assay generations forced us to pool generations 40 and 41 into one set assayed simultaneously. This meant that separate controls could not be identified for generations 30 and 31; to avoid complications, generations 29–31 were treated in the same way as the final assay generation as far as the identification of controls was concerned; *i.e.*, only the first three blocks were used to estimate DM by the order method.

A similar procedure was applied to earlier assay generations, such that genotypes sampled from all later generations were used to produce the bootstrapped set for a given generation, plus an additional set sampled from the real data set for this generation. In this way, a set of bootstrapped values of DM and DV for each assay generation of a given experiment was generated. By resampling a large number of times (usually 1000 times, but sometimes 300 times because of failure to obtain sufficient numbers of controls), the distributions of these values were obtained.

A second round of bootstrapping was then carried out to obtain a single weighted mean estimate of DM and DV for a given bootstrap replicate of the experiment by weighting the estimates for each assay generation by the inverse of their variances obtained from the first bootstrap round. The values for a given bootstrap replicate were used to generate an order method estimate of  $u^*$  and  $s^*$  from Equations 2a and 2b. Similarly, the bootstrapped regression coefficients of block means on generation were used to obtain alternative estimates of DM,  $u^*$ , and  $s^*$ . By repeating this many times, the distributions of the bootstrap estimates of the parameters were obtained.

Bootstrapping over blocks and replicates was not implemented for the results reported here, since the small numbers of blocks and replicates within blocks mean that there is a high probability that only a single block or replicate is included in a sample, distorting the resulting bootstrap estimates. This

should not greatly influence the distributional properties of the genotypic mean and variance for a given generation, since the equations of estimation show that these are influenced much more heavily by genotypic effects than by block and environmental effects.

It should be noted that the above estimators involve equating observed values to their expectations. This introduces a bias into the estimates of  $u^*$  and  $s^*$ , since the expectation of a ratio of two quantities differs from the ratio of their expectations (which is what is desired). A first-order correction can be obtained from the Taylor expansions of  $u^*$  and  $s^*$ ; the bootstrapped means and variances of DM and DV then yield the respective coefficients of variation,  $C_{DM}$  and  $C_{DV}$ . The adjusted estimates are obtained by multiplying the estimate of  $u^*$  by  $(1 - C_{DM}^2 - C_{DV}^2)$ , and the estimate of  $s^*$  by  $(1 - C_{DV}^2)$ . These corrections were applied to both the regression and the order method estimates.

**Revising the order method:** This method of applying the order method yields rather noisy estimates of DM. It assumes that the control lines are mutation free, but these are selected as the set of top-performing lines in a given generation,  $t$ , with no guarantee that they are in fact mutation free, unless such large numbers of replicate measurements are made on each MA line that extremely accurate estimates of their genotypic values are obtained. In general, the mean value of the controls in this generation deviates from that of mutation-free individuals by  $-\bar{s}_c \bar{n}_c$ , where  $\bar{n}_c$  is the mean number of mutations among the controls, and  $\bar{s}_c$  is the mean selection coefficient associated with them. The assumption that the controls are mutation free implies that DM is underestimated, by an amount that is hard to determine. This almost certainly explains much of the discrepancy between the regression-based estimates of DM and the order method results (see supplementary data at <http://www.genetics.org/supplemental/>). A method of examining this problem is described in the APPENDIX, together with a revision of the order method.

## RESULTS

**Lethal mutations:** The numbers of chromosomes with new lethal mutations that arose in each assay period for each experiment are shown in Table 1, together with the corresponding total numbers of independent third chromosomes assayed. Contingency tests failed to show any heterogeneity among generations within experiments in the frequencies of lethal chromosomes. The mean of the estimates of the lethal mutation rate over all three experiments is 0.015. This is close to what would be expected from the estimate of 0.01 for the second chromosome obtained in recent experiments (FRY *et al.* 1999), given the larger size of chromosome 3 (ADAMS *et al.* 2000). This rate is about twice that reported in experiments done 20 years or more earlier, including those of Mukai (CROW and SIMMONS 1983). The (nonsignificant) difference in lethal mutation rates between experiment 1 and the others is not reflected in the mutation rates for detrimental mutations (see below).

**Estimates of mutational parameters:** The observed distributions of genotypic values for each MA generation are available in the supplementary data at <http://www.genetics.org/supplemental/>. There is clear evi-

**TABLE 1**  
Lethal mutations detected in the mutation-accumulation experiments

Generation	Total chromosomes	Lethal chromosomes	Mutation rate (%)
Experiment 1			
10	52	11	2.38
17	37	2	0.79
25	27	7	3.75
35	17	2	1.25
Total	133	22	2.17
Experiment 2			
9-11	69	5	0.75
19-21	58	5	0.90
29-31	49	9	2.09
39-41	45	2	0.40
Total	221	21	1.00
Experiment 3			
9-11	77	7	0.95
20-22	58	8	1.35
29-31	47	6	1.59
Total	182	21	1.38

Only lethal mutations that arose since the previous assay generation are included. The lethal mutation rate was calculated by equating the observed frequency of nonlethal chromosomes to the first term of a Poisson distribution, whose mean is the product of the mutation rate and the number of generations since the last assay generation. For experiments 2 and 3, where sets of three different generations are pooled, this number is equated to the generation number of the middle member of each set. The total mutation rates for each set were calculated using the mean number of generations over the set.

dence from many generations of each experiment for significant genetic variance, indicating the existence of mutational variance in viability. Both block effects and genotype  $\times$  block interactions are also sporadically significant, with variance components that are of similar magnitude to the genetic components (supplemental data).

Table 2 shows estimates of the mutational parameters (DM, DV,  $u^*$ , and  $s^*$ ) for the quasi-normal lines, derived from the primary data by the standard order method with bootstrapping as described in MATERIALS AND METHODS, as well as the results for the revised order method. To obtain these statistics, the data were normalized by dividing by the mean genotypic values of the controls for the earliest generations of the experiments (generation 10 for experiment 1, and the means over generations 9, 10, and 11 for experiments 2 and 3). This allows mutational effects on viability to be measured as fractions of the means for genotypes that are putatively free of new mutations.

The bootstrap confidence intervals for the experiment-wide estimates of the standard order method val-

**TABLE 2**  
Mutational statistics from order method analyses for quasi-normal lines

	Experiment 1		Experiment 2		Experiment 3	
	Standard order method statistics	Revised estimate	Standard order method statistics	Revised estimate	Standard order method statistics	Revised estimate
Control mean	0.886		0.896		0.830	
DM ( $\times 10^3$ )	1.58 (-0.46/3.70)	1.69 (-0.35/3.79)	1.43 (0.52/2.32)	1.63 (1.07/2.12)	1.67 (0.35/2.99)	2.16 (1.48/2.77)
DV ( $\times 10^4$ )	4.83 (2.40/7.72)		1.88 (1.14/2.59)		2.59 (1.59/3.54)	
$u^*$ ( $\times 10^2$ )	0.40 (-0.34/2.02)	0.51 (-0.31/2.22)	1.00 (-0.05/2.43)	1.31 (0.87/1.70)	1.00 (-0.24/2.91)	1.69 (1.20/2.11)
$s^*$	0.370 (-1.16/1.22)	0.230 (-0.889/1.06)	0.114 (0.058/0.273)	0.108 (0.095/0.122)	0.157 (0.045/0.397)	0.113 (0.099/0.124)

The control means are genotypic values for the first set of assay generations for each experiment. The other statistics are based on normalizing the data with respect to the relevant control means. DM and DV are the estimates of the rates of decline in mean and increase in variance, respectively, weighting the values from each assay generation of an experiment by the inverses of their bootstrap variances. The means of the statistics over either 300 or 1000 bootstrap replicates, using the corrections for bias in the revised order method estimate of DM, are displayed. The values in parentheses are the corresponding bootstrap 95% confidence intervals.  $u^*$  and  $s^*$  are the estimates of  $u/(1 + C^2)$  and  $\bar{s}(1 + C^2)$ , with corrections for bias. See text for further explanation.

**TABLE 3**  
**Mutational statistics from order method analyses for all nonlethal lines**

	Experiment 1	Experiment 2	Experiment 3
DM ( $\times 10^3$ )	3.10 (0.38/5.97)	3.50 (2.02/4.94)	2.01 (0.32/3.45)
DV ( $\times 10^4$ )	6.48 (3.44/9.52)	8.54 (5.63/11.2)	2.21 (0.53/3.85)
$u^*$ ( $\times 10^2$ )	1.27 (-0.49/4.37)	1.34 (0.46/2.33)	1.45 (-0.68/4.52)
$s^*$	0.414 (0.041/0.732)	0.238 (0.168/0.342)	0.112 (0.010/0.326)

See Table 2 for explanation. Only standard order method statistics are presented.

ues of DM are wide and even overlap zero for experiment 1. The mean of DM across the three experiments, weighting each estimate by the inverse of its bootstrap variance, is 0.15%, with standard error 0.04%. DV is quite well estimated for each experiment, with a weighted mean across experiments of  $0.021 \pm 0.003\%$ . The confidence intervals on the individual estimates of the mutation rate parameter  $u^*$  are wide; the weighted mean across experiments is  $0.78 \pm 0.49\%$ . The selection coefficient parameter  $s^*$  is poorly estimated, especially for experiment 1. Its weighted mean across experiments 2 and 3 is  $16 \pm 38\%$ .

The revised order method for obtaining estimates, described in the APPENDIX, gives more reliable estimates of DM,  $u^*$ , and  $s^*$  (Table 2). The weighted means (with standard error) of these statistics across experiments are  $0.17 \pm 0.025\%$ ,  $1.4 \pm 0.14\%$ , and  $11 \pm 0.5\%$ , respectively. This method might, however, be biased by non-normality of the distribution of line means. This was examined by calculating values of the skewness and kurtosis of the line mean distributions for each generation. The means and standard errors of skewness across generations were  $-0.219 \pm 0.379$ ,  $-0.063 \pm 0.208$ , and  $-0.157 \pm 0.205$  for experiments 1, 2, and 3, respectively; the corresponding values for kurtosis were  $0.337 \pm 1.044$ ,  $0.203 \pm 0.412$ , and  $0.146 \pm 0.415$ . There is thus no firm evidence for nonnormality in these distributions. The tighter confidence intervals on the estimates obtained by this method suggest that they are to be preferred to the results from the standard order method.

Table 3 shows the analogous results from analyses including all the nonlethal lines. As expected, the estimates of DM and DV are mostly larger than those for the quasi-normal lines (weighted means of  $0.27 \pm 0.06\%$  for the uncorrected estimate of DM and  $0.039 \pm 0.008\%$  for DV), with a weighted mean of  $u^*$  of  $1.3 \pm 0.51\%$ . The weighted mean of  $s^*$  is  $24 \pm 5.6\%$ . The revised order method values are not shown here, since the distributions of genotypic means are far from normal (supplemental data at <http://www.genetics.org/supplemental/>).

Additional information on the effects of mutations on viability can be obtained by examining the linear regressions of block means on generation,  $b_M$ , as de-

scribed in MATERIALS AND METHODS (Table 4). In all cases, these are negative. For the quasi-normals in experiments 2 and 3 they are highly significant by  $t$ -tests ( $P < 0.01$ ), but not for experiment 1, although they are all of similar magnitude. The mean value of  $b_M$  across experiments, weighted by the inverses of the sampling variances of the regression estimates, is  $-0.47 \pm 0.065\%$ . When all nonlethal lines are included, the regressions are larger, except for experiment 3, and even the experiment 1 regression is significant at the 5% level. The corresponding estimates of  $u^*$  are larger than those from the order method, with a weighted mean of  $0.99 \pm 1.6\%$  for the quasi-normal lines and  $4.4 \pm 1.3\%$  for all the nonlethal lines. The corresponding  $s^*$  estimates are not very reliable, with pathologically high negative values for experiment 1. The mean  $s^*$  for the quasi-normals of experiments 2 and 3 is  $2.9 \pm 0.82\%$ , and for all nonlethals it is  $10.2 \pm 2.0\%$ .

If detrimental mutations affecting viability have predominantly negative effects, the considerations outlined in MATERIALS AND METHODS predict that the mean genotypic values of the controls should decline with assay generation. Unfortunately, these are based on numbers much smaller than the overall means, and the relevant regression values ( $b_c$ ) have large sampling errors, but that for experiment 2 is highly significant ( $P < 0.01$ ). The mean value of  $b_c$  across the experiments, weighted inversely by sampling variances, is  $-0.39 \pm 0.13\%$ .

## DISCUSSION

As discussed in several recent publications (KEIGHTLEY and EYRE-WALKER 1999; LYNCH *et al.* 1999; FRY and HEINSOHN 2002), different mutation-accumulation experiments on *D. melanogaster* have yielded widely varying estimates of mutational parameters. The previous experiments involved either chromosome 2, using balancer crosses similar to those conducted here, sib-mated lines in which mutations accumulated over the whole genome (GARCÍA-DORADO 1997; CHAVARRÍAS *et al.* 2001; AVILÁ and GARCÍA-DORADO 2002), or an outbred population in which selection was supposedly suspended and genome-wide mutations were allowed to

accumulate (SHABALINA *et al.* 1997). The different experiments also used different types of controls, putatively free of mutations, for the purpose of calibrating the DM values, as well as different measures of viability or fitness (FRY 2001). This makes it somewhat difficult to interpret the results.

**Estimates of DM:** An attempt to bring at least part of this rather confusing plethora of results under a common framework has been made by FRY (2001), in which he reanalyzed the earlier experiments of Mukai and co-workers (MUKAI 1964; MUKAI *et al.* 1972), using the same viability measure and order method employed here and taking into account some criticisms of these experiments (KEIGHTLEY and EYRE-WALKER 1999). He also reanalyzed the experiments of FRY *et al.* (1999), employing the order method instead of a random-bred control population started from the progenitor chromosome 2 used in the MA lines, to avoid the problem that some mutations must have accumulated in this population, due to the ineffectiveness of selection on rare alleles (CABALLERO *et al.* 2002).

This seemingly goes some way to reconciling the different results on DM. Mukai's 1964 experiments gave a mean value of DM of 0.42% for viability effects of chromosome 2 quasi-normal lines with 0.48% for the 1972 experiments. These are similar to regression-based estimates that do not use controls. Fry's experiments gave a mean value of 0.32%. Ohnishi's experiment (OHNISHI 1977), which lacked a control, gave an initial rate of decline of mean viability of 0.52% and an overall rate of decline of 0.25%, with a significant quadratic component (FRY 2001). The weighted mean bootstrapped standard and revised order method estimates of DM for quasi-normals in the present experiments (0.15 and 0.17%, respectively) are much smaller than these estimates, whereas the weighted mean regression estimate is closer ( $0.51 \pm 0.74\%$ ). The expectation from the relative sizes of the euchromatin of chromosomes 2 and 3 is that DM should be 1.18 times greater for chromosome 3 (ADAMS *et al.* 2000).

As mentioned in MATERIALS AND METHODS, the order method probably seriously underestimates DM, due to the presence of mutations in the control lines. The difference between control and overall means may thus seriously underestimate the difference between mutation-free genotypes and the overall mean of a mutation-accumulation generation. This problem is worse for later generations of mutation experiments, when it becomes increasingly unlikely that a line is mutation free; *e.g.*, with a mutation rate of 0.05, the fraction of mutation-free lines at generation 40 is only 13.5%. The problem is less severe for generations 20 and earlier, where the probability of selecting a mutation-free control is high if lines are assayed with complete accuracy. However, the low heritability of line means in our experiments (<40% in the earlier generations) means that

TABLE 4  
Mutational statistics from regression analyses

	Experiment 1		Experiment 2		Experiment 3	
	Mutational statistics for the quasi-normal lines	All nonlethals	Mutational statistics for the quasi-normal lines	All nonlethals	Mutational statistics for the quasi-normal lines	All nonlethals
$b_M (\times 10^3)$	$-2.08 \pm 1.77$	$-4.88 \pm 1.87$	$-5.68 \pm 0.857$	$-7.65 \pm 0.873$	$-4.01 \pm 1.20$	$-3.79 \pm 1.44$
$b_C (\times 10^3)$	$-4.12 \pm 5.19$		$-5.33 \pm 1.62$		$-0.730 \pm 2.45$	
DM ( $\times 10^3$ )	1.06 (-2.15/4.48)	2.86 (0.10/9.72)	6.66 (5.16/8.25)	8.42 (6.42/10.5)	3.75 (1.60/6.43)	3.27 (0.89/6.35)
$u^*$ ( $\times 10^2$ )	0.25 (-0.88/3.45)	2.86 (-1.57/12.2)	24.1 (10.9/44.6)	8.12 (4.68/12.4)	5.89 (-0.34/16.7)	11.2 (-2.01/37.8)
$s^*$	-1.27 (-3.56/-0.181)	0.205 (-0.022/0.764)	0.023 (0.010/0.035)	0.100 (0.069/0.133)	0.064 (0.021/0.165)	0.064 (-0.003/0.215)

$b_M$  and  $b_C$  are the regression coefficients on assay generation (with standard errors) of the genotypic means and control means for each generation, respectively. The estimates of DM,  $u^*$ , and  $s^*$  (after normalizing the data by the mean control genotypic values) are the means over 1000 bootstrap replicates, with corrections for bias.

misclassification due to nongenetic effects is still possible.

If this interpretation is correct, using only the earlier MA generations should give higher estimates of DM by the order method. This was tested by repeating the analyses using only generations 10 and 17 for experiment 1, 9–21 for experiment 2, and 9–22 for experiment 3. The bootstrapped mean DM estimates for the revised order method applied to quasi-normal lines were 0.15, 0.35, and 0.24% for experiments 1–3, respectively, with a weighted mean of  $0.24 \pm 0.047\%$ . This lends some support to the hypothesis that the inclusion of mutation-carrying genotypes in the controls causes the discrepancy between our regression-based estimates of DM and the order method results. The order method may therefore be too conservative. Fry's reanalysis used mostly generations close to 20 and earlier (FRY 2001), so that there should be less bias, consistent with the generally high DM estimates that he obtained. The negative values of the regressions of control means on generation in the present experiments, significant in the case of experiment 2 (Table 2), are consistent with this interpretation and with the idea that new mutations predominantly reduce fitness (CABALLERO and KEIGHTLEY 1998; KEIGHTLEY and LYNCH 2003). The approach described in the supplemental data (<http://www.genetics.org/supplemental/>) shows that estimates of DM from the order and regression methods can be reconciled if the former are biased by the presence of mutations in the controls.

This suggests that the regression of mean viability on generation of MA may be the most reliable method of estimating the mutational parameters if the mutation rate is too high, despite the disadvantage of not having a control. In the present case, the disadvantage of the lack of a control is at least partly offset by the fact that each generation of mutation accumulation was assayed over several blocks at widely separated times. Environmental fluctuations will therefore be averaged over blocks, unless there is a systematic trend with time over the duration of the whole experiment. The possibility of such an environmental trend was examined using the data on the block means for each experiment. Since the successive blocks were assayed at intervals of more than one generation (2 weeks), any systematic deterioration of the environment or change in the balancer chromosome, leading to an apparent decline in viability, would also lead to a decline in block means over the intervals in question. The trend of block means within MA generations was within the limits of sampling error, although of a similar magnitude to the regressions of block means on MA generation (supplemental data at <http://www.genetics.org/supplemental/>).

The possibility of a purely environmental cause of the regressions of means on time is thus not excluded, but it would require a remarkable coincidence for all three experiments to show a similar change in mean over MA generations for purely environmental reasons, generat-

ing similar regressions of means on time. In addition, the balancer stock used for these experiments (*TM6/Sb*) was maintained for many years in the laboratory before the start of these experiments (LINDSLEY and ZIMM 1992), so that continuing evolutionary change in the balancer (KEIGHTLEY and EYRE-WALKER 1999) is unlikely to produce the observed decline in viability.

It therefore seems reasonable to accept the regression coefficients as providing the least biased estimators of DM, although the very high value for experiment 2 might be questioned. Bootstrapping all three experiments simultaneously yields a joint mean bootstrap regression estimate for DM for quasi-normals of 0.40% (95% confidence interval  $\pm 0.14\%$ ). This is consistent with the DM for the second chromosome reported by FRY and HEINSOHN (2002), which corresponds to a value of 0.30% for the third chromosome 3.

The results from the different balancer experiments thus agree in suggesting that there is a real mutational decline in homozygous viability caused by detrimental mutations in the quasi-normal lines, in the range 0.2–0.5% for an individual autosome and 1–2% for the genome as a whole. This compares well with the estimate of a 2% rate of decline in competitive fitness for an outbred population sheltered from selection, estimated under competitive conditions (SHABALINA *et al.* 1997), taking into account the partial recessivity of detrimental mutations (GARCÍA-DORADO and CABALLERO 2000). The only experiments that gave a much lower estimate for DM are those involving the sib-mated "Madrid lines," which consistently yield a genome-wide DM that is substantially <1%, even for a competitive measure of net fitness (CHAVARRÍAS *et al.* 2001; AVILÁ and GARCÍA-DORADO 2002). The reason for this disagreement with the results of the other experiments is unclear, although it may reflect the fact that the measurements of chromosome 2 viability and fitness under competitive conditions were done after >200 generations of mutation accumulation, when many of the original lines had been lost (CHAVARRÍAS *et al.* 2001), thus biasing the experiments against lines with low fitness.

**Estimates of DV:** There is substantial variation in the estimates of DV among different experiments reported in the literature, possibly associated with differences in larval densities among experiments (FRY and HEINSOHN 2002). For second chromosome viability experiments, the range of DV for quasi-normals is 0.006–0.022% (FRY and HEINSOHN 2002, Table 8), with a mean of 0.017%. This scales to a diploid genome-wide mean value of 0.077%, compared with a joint bootstrap mean across all three of our experiments of 0.11% (95% confidence interval 0.084–0.14%).

The present estimate is, however, well within the range for the second chromosome when corrected for the relative sizes of the autosomes. The number of pairs of parents per vial was four, intermediate between the high- and low-density experiments analyzed by FRY and

HEINSOHN (2002). However, the highly inbred nature of the genetic background meant that the productivity per vial was very low, with a mean of  $\sim 100$  flies per vial across all three experiments, compared with  $\sim 198$ – $300$  flies in the low-density treatments of FRY and HEINSOHN (2002), which yielded low values of DV. It is thus not clear whether density really has a systematic effect on DV, especially as the high-density experiment of CHAVARRÍAS *et al.* (2001) yielded a low value of DV. The overall genome-wide value of DV, based on the present value and a simple mean of the results summarized in Table 8 of FRY and HEINSOHN (2002), is  $\sim 0.1\%$ .

The method of estimating DV used here uses the assumption that its value is zero for the initial generation of each experiment and is similar to forcing a regression through the origin. As mentioned in MATERIALS AND METHODS, the regression-based estimates of the rate of change of variance for the quasi-normal lines are very noisy. However, with the exception of experiment 1, the mean bootstrapped weighted regression coefficients of genetic variance on MA generation (using bootstrap variances of the individual generation genetic variances as weights) are similar to the third chromosome DV values obtained here: 0.017, 0.022, and 0.025% for experiments 1–3, respectively. The confidence intervals on these are, however, wide and overlap zero in the case of experiments 1 and 3. It is possible that the DV value for experiment 1 may be artifactually inflated, since the intercept of the regression of genetic variance is significantly positive in this case (0.82% with 95% confidence interval 0.32–1.4%). Estimates of the mutation rate that use the DV value for experiment 1 may therefore be too small, but the relatively high variance of this estimate means that it does not greatly affect the pooled estimate given above. The source of such an artifact is unclear; it is unlikely to be due to common environmental effects of the cultures in which the parents of the flies used for assays of a given genotype were raised, since the genotypic variance components depend on means across blocks, with independent parents in different blocks. Common parental environment effects are therefore absorbed into genotype  $\times$  block interactions.

**Estimates of the mutation rate:** We denote the whole diploid genome Bateman-Mukai estimator (equivalent to Equation 2a) by  $U^*$ . Using the mean of DM for quasi-normals over all experiments of FRY and HEINSOHN (2002) with their low-density estimate of DV, and scaling up to the diploid euchromatic genome by dividing by the proportion of the euchromatin represented by chromosome 2, we obtain  $U^* = 56\%$ , comparable with the classical Mukai estimates (MUKAI 1964; MUKAI *et al.* 1972). On the other hand, their high-density estimate of DV yields  $U^* = 9\%$ . Bootstrapping over our three experiments simultaneously gives a regression estimate of  $U^*$  of 50% (95% confidence interval 28–79%). A more stable estimate is obtained by using all nonlethal

lines: with these,  $U^*$  is estimated from the regressions as 29% (95% confidence interval 14–54%). The true mutation rates for all nonlethals and quasi-normal lines are likely to be very close to each other, since the rates of occurrence per generation of mutations that create low viability but nonlethal lines (as estimated from the frequencies of low-viability lines in the final MA generations of the experiments) are low: 0.48, 0.69, and 0.32% for the respective experiments, with a mean value of 0.50%. The estimate from the nonlethal lines is thus probably to be preferred, given its tighter confidence interval.

Much lower values are obtained from the order method applied to all generations (Tables 2 and 3). Estimation of the diploid genome-wide  $U^*$ , using the ratio of the square of the weighted mean across experiments of DM from the standard order method to the weighted mean estimate of DV, yields a  $U^* = 4.2\%$  (standard error 2.3%) for quasi-normals; the revised method gives a value of  $6.3 \pm 2.0\%$ . These should be less biased than the corresponding means of the  $U^*$  estimates. Higher values are obtained using only the first two assay generations of each experiment (see RESULTS); the ratio of the weighted mean of the corresponding estimates of DM to the weighted mean estimate of DV gives an estimate of  $U^*$  over the three experiments of  $12 \pm 5.5\%$  for these generations. Given the concerns about the biases in the order method estimates, it seems reasonable to accept a  $U^*$  of 12–30% as an estimate for the diploid genome of *D. melanogaster*, on the basis of these results and those of FRY (2001) and FRY and HEINSOHN (2002), although a higher value is not excluded if the present regression estimates are taken at face value. The true diploid mutation rate to detrimental alleles,  $U$ , will be much higher than this, if there is a wide distribution of fitness effects of mutations (MUKAI *et al.* 1972; GARCÍA-DORADO and GALLEGO 2003), especially since only viability mutations have been accounted for in these experiments. For example, with an exponential distribution of mutational effects, the true value of  $U$  would be twice our estimate, and the mean value of the selection coefficient would be one-half our estimate (MUKAI *et al.* 1972).

There is thus still considerable uncertainty concerning the value of the per genome deleterious mutation rate,  $U$ , for detrimental alleles in *D. melanogaster*, although it seems clear that suggestions that it is of the order of only 2–3% (CABALLERO and KEIGHTLEY 1998; AVILÁ and GARCÍA-DORADO 2002) are inconsistent with the more recent reappraisals of the data on balancer-based experiments, as well as with the present study. Estimates based on amino acid sequence comparisons between species indicate that  $U$  cannot be  $< 6$ – $8\%$  for *D. melanogaster* (KEIGHTLEY and EYRE-WALKER 2000). A recent analysis of selective constraints on noncoding sequences near genes, using a comparison of *D. melanogaster* and *D. simulans*, suggests that there are at least

40% more nucleotide sites of functional significance than amino acid coding sites, so that  $U$  from single-nucleotide changes is at least 8–11% (HALLIGAN *et al.* 2004). The rate of insertion/deletions in *D. melanogaster* is  $\sim 25\%$  of that for nucleotide changes (JENSEN *et al.* 2002), so that mutations arising from this source may increase this estimate of  $U$  to 10–14%. These estimates are, however, subject to considerable uncertainty, due to the difficulty of calibrating the molecular clock and assigning generation times for *Drosophila* (KEIGHTLEY and EYRE-WALKER 1999).

**Estimates of mutational effects:** The values of  $s^*$  for the quasi-normal lines in our experiment are poorly estimated by the regression method (Table 4), except for experiment 2, which yields a value of  $\sim 2\%$ ; these are in any case upper bounds to the mean selection coefficients. Much higher values (of the order of 10%) are obtained from the revised order method (Table 2), but these are likely to be even more upwardly biased. In any case, it seems unlikely that mutations with effects of less than a few percent on viability can contribute much to our DM and DV estimates. As has been pointed out before (KEIGHTLEY and EYRE-WALKER 1999), mutations with small effects on fitness, such as are likely to be caused by most amino acid substitutions or nucleotide substitutions in regulatory sequences, will not contribute significantly to DM or DV in mutation-accumulation experiments. For instance, if mutations with selection coefficients of 0.15% arise at a rate of 0.5 per genome per generation, DM is only 0.075% and DV is only  $1.1 \times 10^{-6}$ ,  $\sim 1\%$  of the value we observe. This conclusion has been validated in a *Caenorhabditis elegans* mutagenesis experiment, by comparing observed numbers of mutations to the number of amino acid mutations estimated to have occurred (DAVIES *et al.* 1999). The only alternative is that there is a very high mutation rate, much higher than that implied by the sequence comparisons, with a large variance in selection coefficients among mutations.

This suggests that transposable element (TE) insertions and insertion/deletion mutations, which are more likely than point mutations to severely disrupt the function of coding and regulatory sequences, may contribute many of the deleterious mutations detected by *D. melanogaster* MA experiments (KEIGHTLEY and EYRE-WALKER 1999). Excluding unusually mobile elements, TE insertions typically occur at a rate of  $\sim 20\%$  per diploid genome per generation in *D. melanogaster* (MASIDE *et al.* 2000). Very active TEs, such as *P* elements, may contribute at least as much again, even in nondysgenic stocks like those used here (EGGLESTON *et al.* 1988; FRY and NUZHIDIN 2003); this is likely to vary according to genetic background (NUZHIDIN *et al.* 1998), perhaps accounting for the discrepancies between different experiments. Many TE insertions will, however, be into intergenic sequences and will not contribute to large viability re-

ductions. Nevertheless, direct estimates of the homozygous viability effects of random P insertions suggest mean  $s$  values of several percent (EANES *et al.* 1988; MACKAY *et al.* 1992; LYMAN *et al.* 1996). There is thus little difficulty in accounting for a large fraction of the mutational effects detected in MA experiments in terms of transposable element activity, rather than point mutations. The best hope for progress on the contribution of single-nucleotide changes and small insertions/deletions (indels) to the genomic deleterious mutation rate is to make more detailed molecular comparisons between related species, which may resolve the question of the extent of selective constraints on different components of the genome.

Finally, reexamination of the properties of the likely fitness effects of heterozygous mutations using the revised estimates of mutational parameters suggests that previous conclusions (CHARLESWORTH and HUGHES 2000) remain essentially unchanged (see supplemental data at <http://www.genetics.org/supplemental/>).

**Epistatic effects of new mutations:** The multigeneration nature of our MA experiments allows a test of the extent to which mutational effects at different loci interact, as measured by departure from linearity of the regression of viability on MA generation (MUKAI 1969). This is of especial interest in relation to the mutational-deterministic theory of the advantage of sex and recombination (KONDRASHOV 1988, 1993), which requires synergistic epistasis among deleterious mutations with respect to the logarithm of fitness, implying significantly negative quadratic coefficients of the regression of log fitness on generation of mutation accumulation. Following logarithmic transformations of the viability measurements, quadratic regressions of block means on MA generation were fitted for each experiment. No evidence for any significant quadratic terms was obtained; the quadratic terms were at least 10-fold smaller in magnitude than the linear terms in all three experiments and were in any case positive in experiments 2 and 3 (values of  $0.004 \pm 0.012\%$  and  $0.031 \pm 0.032\%$ , respectively). This does not rule out the mutational-deterministic theory, since very small epistatic effects can have evolutionarily significant consequences (CHARLESWORTH 1990), but clearly fails to provide positive support for it. As pointed out previously, the high level of synergism estimated by MUKAI (1969) cannot easily be reconciled with evidence from measurements of inbreeding load (CHARLESWORTH 1998). Currently, there is only weak evidence for widespread synergistic effects of detrimental mutations (RIVERO *et al.* 2003; SZAFRANIEC *et al.* 2003).

We thank Jane Charlesworth for assistance in counting flies and Helen Cowan for media preparation. We also thank Peter Keightley, Michael Simmons, and two referees for their helpful comments on the manuscript. This research was supported by grants from the Biotechnology and Biological Sciences Research Council of the United Kingdom and the Royal Society.

## LITERATURE CITED

- ADAMS, M. D., S. E. CELNIKER, R. A. HOLT, C. A. EVANS, J. D. GOCAYNE *et al.*, 2000 The genome sequence of *Drosophila melanogaster*. *Science* **287**: 2185–2195.
- AVILÁ, V., and A. GARCÍA-DORADO, 2002 The effects of spontaneous mutation on competitive fitness in *Drosophila melanogaster*. *J. Evol. Biol.* **15**: 561–566.
- BATEMAN, A. J., 1959 The viability of near-normal irradiated chromosomes. *Int. J. Radiat. Biol.* **1**: 170–180.
- BECKER, W. A., 1976 *Manual of Quantitative Genetics*, Ed. 4. Academic Enterprises, Pullman, WA.
- BERG, D. E., and M. M. HOWE, 1989 *Mobile DNA*. American Society of Microbiology, Washington, DC.
- CABALLERO, A., and P. D. KEIGHTLEY, 1998 Inferences on genome-wide deleterious mutation rates in inbred populations of *Drosophila* and mice. *Genetica* **102/103**: 229–239.
- CABALLERO, A., E. CUSI, C. GARCÍA and A. GARCÍA-DORADO, 2002 Accumulation of deleterious mutations: additional *Drosophila melanogaster* estimates and a simulation of the effects of selection. *Evolution* **56**: 1150–1159.
- CHARLESWORTH, B., 1990 Mutation-selection balance and the evolutionary advantage of sex and recombination. *Genet. Res.* **55**: 199–221.
- CHARLESWORTH, B., 1998 The effect of synergistic epistasis on the inbreeding load. *Genet. Res.* **71**: 85–89.
- CHARLESWORTH, B., and D. CHARLESWORTH, 1985 Genetic variation in recombination in *Drosophila*. I. Responses to selection and preliminary genetic analysis. *Heredity* **54**: 71–84.
- CHARLESWORTH, B., and D. CHARLESWORTH, 1998 Some evolutionary consequences of deleterious mutations. *Genetica* **102/103**: 3–19.
- CHARLESWORTH, B., and K. A. HUGHES, 2000 The maintenance of genetic variation in life-history traits, pp. 369–392 in *Evolutionary Genetics from Molecules to Morphology*, edited by R. S. SINGH and C. B. KRIMBAS. Cambridge University Press, Cambridge, UK.
- CHARLESWORTH, B., P. JARNE and S. ASSIMACOPOULOS, 1994 The distribution of transposable elements within and between chromosomes in a population of *Drosophila melanogaster*. III. Element abundances in heterochromatin. *Genet. Res.* **64**: 183–197.
- CHAVARRÍAS, D., C. LOPEZ FANJUL and A. GARCÍA-DORADO, 2001 The rate of mutation and the homozygous and heterozygous competitive effects for competitive viability: a long term experiment with *Drosophila melanogaster*. *Genetics* **158**: 681–693.
- CROW, J. F., 1993 Mutation, mean fitness, and genetic load. *Oxf. Surv. Evol. Biol.* **9**: 3–42.
- 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. Academic Press, London.
- DAVIES, E. K., A. D. PETERS and P. D. KEIGHTLEY, 1999 High frequency of cryptic deleterious mutations in *Caenorhabditis elegans*. *Science* **285**: 1748–1751.
- EANES, W. F., C. WESLEY, J. HEY and D. HOULE, 1988 Fitness consequences of P element insertion in *Drosophila melanogaster*. *Genet. Res.* **52**: 17–26.
- EFRON, B., and R. J. TIBSHIRANI, 1993 *An Introduction to the Bootstrap*. Chapman & Hall/CRC Press, Boca Raton, FL.
- EGGLESTON, W. B., D. JOHNSON-SCHLITZ and W. R. ENGELS, 1988 P-M hybrid dysgenesis does not mobilize other transposable elements in *D. melanogaster*. *Nature* **331**: 368–370.
- FALCONER, D. S., and T. F. C. MACKAY, 1996 *An Introduction to Quantitative Genetics*, Ed 4. Longman, London.
- FRY, J. D., 2001 Rapid mutational declines of viability in *Drosophila*. *Genet. Res.* **77**: 53–60.
- FRY, J. D., and S. L. HEINSOHN, 2002 Environment dependence of mutational parameters for *Drosophila melanogaster*. *Genetics* **161**: 1155–1167.
- FRY, J. D., and S. V. NUZHIDIN, 2003 Dominance of mutations affecting viability in *Drosophila melanogaster*. *Genetics* **163**: 1357–1364.
- FRY, J. D., P. D. KEIGHTLEY, S. L. HEINSOHN and S. V. NUZHIDIN, 1999 New estimates of the rates and effects of mildly deleterious mutation in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **96**: 574–579.
- GARCÍA-DORADO, A., 1997 The rate and effects distribution of viability mutations in *Drosophila*: minimum distance estimation. *Evolution* **51**: 1130–1139.
- GARCÍA-DORADO, A., and A. CABALLERO, 2000 On the average coefficient of dominance of deleterious spontaneous mutations. *Genetics* **155**: 1991–2001.
- GARCÍA-DORADO, A., and A. GALLEGO, 2003 Comparing analysis methods for mutation-accumulation data: a simulation study. *Genetics* **164**: 807–819.
- HALDANE, J. B. S., 1956 The estimation of viabilities. *J. Genet.* **54**: 294–296.
- HALLIGAN, D. L., A. EYRE-WALKER, P. ANDOLFATTO and P. D. KEIGHTLEY, 2004 Patterns of evolutionary constraints in intronic and intergenic DNA of *Drosophila*. *Genome Res.* **14**: 273–279.
- HOULE, D., D. K. HOFFMASTER, S. ASSIMACOPOULOS and B. CHARLESWORTH, 1992 The genomic mutation rate for fitness in *Drosophila*. *Nature* **359**: 58–60. [erratum: *Nature* **371**: 358 (1994)].
- HOULE, D., A. S. KONDRASHOV, L. Y. YAMPOLSKY, S. CALDWELL and P. L. STEPONKUS, 1997 The effect of cryopreservation on the lethal mutation rate in *Drosophila melanogaster*. *Genet. Res.* **69**: 209–213.
- JENSEN, M. A., B. CHARLESWORTH and M. KREITMAN, 2002 Patterns of genetic variation at a chromosome 4 locus of *Drosophila melanogaster* and *D. simulans*. *Genetics* **160**: 493–507.
- KEIGHTLEY, P. D., and A. EYRE-WALKER, 1999 Terumi Mukai and the riddle of deleterious mutation rates. *Genetics* **153**: 515–523.
- KEIGHTLEY, P. D., and A. EYRE-WALKER, 2000 Deleterious mutations and the evolution of sex. *Science* **290**: 331–333.
- KEIGHTLEY, P. D., and M. LYNCH, 2003 Toward a realistic model of mutations affecting fitness. *Evolution* **57**: 683–685.
- KEMPTHORNE, O., 1957 *An Introduction to Genetic Statistics*. John Wiley & Sons, New York.
- KONDRASHOV, A. S., 1988 Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**: 435–440.
- KONDRASHOV, A. S., 1993 A classification of hypotheses on the advantage of amphimixis. *J. Hered.* **84**: 372–387.
- LINDSLEY, D. L., and G. G. ZIMM, 1992 *The Genome of Drosophila melanogaster*. Academic Press, San Diego.
- LYMAN, R. F., F. LAWRENCE, S. V. NUZHIDIN and T. F. C. MACKAY, 1996 Effects of single P element insertions on bristle number and viability in *Drosophila melanogaster*. *Genetics* **143**: 277–292.
- LYNCH, M., J. BLANCHARD, D. HOULE, T. KIBOTA, S. SCHULTZ *et al.*, 1999 Perspective: spontaneous deleterious mutation. *Evolution* **52**: 645–663.
- MACKAY, T. F. C., R. F. LYMAN and M. S. JACKSON, 1992 Effects of P element mutations on quantitative traits in *Drosophila melanogaster*. *Genetics* **130**: 315–332.
- MASIDE, X., S. ASSIMACOPOULOS and B. CHARLESWORTH, 2000 Rates of movement of transposable elements on the second chromosome of *Drosophila melanogaster*. *Genet. Res.* **75**: 275–284.
- 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., 1969 The genetic structure of natural populations of *Drosophila melanogaster*. VII. Synergistic interactions of spontaneous mutant polygenes affecting viability. *Genetics* **61**: 749–761.
- 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.
- NUZHIDIN, S. V., E. G. PASYUKOVA, E. A. MOROZOVA and A. J. FLAVELL, 1998 Quantitative genetic analysis of copia retrotransposon activity in inbred *Drosophila melanogaster* lines. *Genetics* **150**: 755–766.
- OHNISHI, O., 1977 Spontaneous and ethyl methanesulfonate-induced mutations controlling viability in *Drosophila melanogaster*. II. Homozygous effects of polygenic mutations. *Genetics* **87**: 529–545.
- RIVERO, A., F. BALLOUX and S. A. WEST, 2003 Testing for epistasis between deleterious mutations in a parasitoid wasp. *Evolution* **57**: 1698–1703.
- SHABALINA, S. A., L. Y. YAMPOLSKY and A. S. KONDRASHOV, 1997 Rapid decline of fitness in panmictic populations of *Drosophila melanogaster* maintained under relaxed selection. *Proc. Natl. Acad. Sci. USA* **94**: 13034–13039.

- SHAW, R. G., C. J. GEYER and F. H. SHAW, 2002 A comprehensive model of mutations affecting fitness and inferences for *Arabidopsis thaliana*. *Evolution* **56**: 453–463.
- SIMMONS, M. J., N. A. JOHNSON, T. M. FAHEY, S. M. NELLETT and J. D. RAYMOND, 1980 High mutability in male hybrids of *Drosophila melanogaster*. *Genetics* **96**: 479–490.
- SOKAL, R. R., and F. J. ROHLF, 1995 *Biometry*. W. H. Freeman, San Francisco.
- SZAFRANIEC, K., D. M. WLOCH, P. SLIWA, R. H. BORTS and R. KORONA, 2003 Small fitness effects and weak genetic interactions between deleterious mutations in heterozygous mutations of the yeast *Saccharomyces cerevisiae*. *Genet. Res.* **82**: 19–31.

Communicating editor: M. SIMMONS

## APPENDIX

New mutations arise independently of preexisting ones, so that the expected genotypic value of a control genotype identified in generation  $t$  when measured in an earlier generation  $t'$  will be  $(t - t')$ DM greater than the expected value for the same genotypes in generation  $t$ , unless selection is so rigorous that only mutation-free lines are included in the controls. If line means are distributed with standard deviation  $\sigma_g(t)$ , their expected genotypic value is equal to the overall genotypic mean  $\mu(t)$ , plus the product of  $\sigma_g$  and the selection differential  $i(t)$  for generation  $t$ . With a normal distribution of the line means,  $i$  depends only on the proportion of controls among all lines assayed (FALCONER and MACKAY 1996).

The expected genotypic value of the controls in generation  $t'$ ,  $\mu_c(t')$ , is thus

$$\mu(t) + (t - t')\text{DM} + i(t)h(t)\sigma_g(t) = \mu(t') + i(t)h(t)\sigma_g(t),$$

where  $h^2$  is the proportion of variance in line means due to genetic effects (FALCONER and MACKAY 1996). This implies that

$$\mu_c(t') - \mu(t') = -\bar{s}_c(t')\bar{n}_c(t') + t'\text{DM} = i(t)h(t)\sigma_g(t). \quad (\text{A1})$$

This shows how estimates of DM by the order method are biased downward by the presence of the deleterious mutations in the control lines. In addition, it suggests

that, if we ignore this source of bias, we can obtain an alternative lower-bound estimate of DM from

$$\sum_t (i(t)\sigma_g(t)h(t)w(t')/t') / \sum_t w(t'), \quad (\text{A2})$$

where  $w(t')$  is the bootstrap weight for assay generation  $t'$  used in the order method estimate of DM.

To carry out this procedure using the assumption of normality, the value of  $i$  for each generation is obtained from the proportion of lines used as controls, using tables of  $i$  based on extreme value theory (BECKER 1976). We also need to estimate the genetic variance for each assay generation, as well as the nongenetic component of the variance of line means. Since the prediction equations rely on the underlying true values of the genetic and environmental variance components, and the parameter estimates for individual generations are very noisy, the genetic variance for generation  $t$  was equated to the product of  $t$  and the mean bootstrap estimate of DV for an experiment. Similarly, the mean value of the residual variance of genotype/block means over all assay generations was used to estimate the nongenetic variance term in  $h$ , denoted by  $\sigma_r^2$ . In cases where block four was used to identify controls for a given assay generation, this residual variance component was divided by three (the number of blocks over which means were taken to estimate genotypic values); when a later generation was used for this purpose, it was divided by four.

A first-order correction for the bias involved in equating the expectation of the observed value of  $h\sigma_g$  to its true value was also applied, using the bootstrap means and variances of DV and the residual variance. The correction term is

$$c = (\sigma_g^2 + \sigma_r^2)^{-\frac{3}{2}} \left( 1 - \frac{3}{4}(\sigma_g^2 + \sigma_r^2)^{-1} \right). \quad (\text{A3})$$

Confidence intervals on the revised estimates of DM and the other mutational parameters were obtained by bootstrapping as before.