Evolution, 57(5), 2003, pp. 1022–1030

RAPID FITNESS RECOVERY IN MUTATIONALLY DEGRADED LINES OF *CAENORHABDITIS ELEGANS*

SUZANNE ESTES[1,2] AND MICHAEL LYNCH[3]
[1]*Department of Biology, University of Oregon, Eugene, Oregon 97403*
[3]*Department of Biology, Indiana University, Bloomington, Indiana 47405*
E-mail: mlynch@bio.indiana.edu

Abstract.—Deleterious mutation accumulation has been implicated in many biological phenomena and as a potentially significant threat to human health and the persistence of small populations. The vast majority of mutations with effects on fitness are known to be deleterious in a given environment, and their accumulation results in mean population fitness decline. However, whether populations are capable of recovering from negative effects of prolonged genetic bottlenecks via beneficial or compensatory mutation accumulation has not previously been tested. To address this question, long-term mutation-accumulation lines of the nematode *Caenorhabditis elegans*, previously propagated as single individuals each generation, were maintained in large population sizes under competitive conditions. Fitness assays of these lines and comparison to parallel mutation-accumulation lines and the ancestral control show that, while the process of fitness restoration was incomplete for some lines, full recovery of mean fitness was achieved in fewer than 80 generations. Several lines of evidence indicate that this fitness restoration was at least partially driven by compensatory mutation accumulation rather than a result of a generic form of laboratory adaptation. This surprising result has broad implications for the influence of the mutational process on many issues in evolutionary and conservation biology.

Key words.—*Caenorhabditis elegans*, compensatory mutation, fitness, life-history characters, mutation accumulation, mutation load, natural selection.

Received March 29, 2002. Accepted December 12, 2002.

It is well established that the accumulation of slightly deleterious mutations via genetic drift under conditions of relaxed selection can threaten the health and persistence of small populations (Lynch and Gabriel 1990; Lande 1994; Lynch et al. 1995; Zeyl et al. 2001). Even large populations, such as those of humans, may be at risk of genetic erosion due to high mutation rates coupled with diminished levels of natural selection resulting from modern medicine and public health practices (Muller 1950; Kondrashov 1995; Crow 1997; Eyre-Walker and Keightley 1999; Lynch et al. 1999; Shaw et al. 2002). However, deleterious mutation accumulation (MA) is perhaps most relevant to small captive populations of endangered species living in benign environments (Lande 1995a). Because the descendants of such populations are often targeted for reintroduction into natural habitat, it is important to know whether acquired mutational loads have permanent effects on fitness. Hence, there is a necessity to understand both the process of fitness decline in small populations and the potential for fitness recovery subsequent to the accumulation of genetic deterioration.

Although controversy currently surrounds estimates of deleterious mutational parameters (e.g., rate, magnitude, and distribution of effects) (García-Dorado et al. 1999; Keightley and Eyre-Walker 1999; Lynch et al. 1999; Chavarrías et al. 2001; Fry 2001, Caballero et al. 2002; Fry and Heinsohn 2002), numerous mutation-accumulation studies reveal the potential for spontaneous deleterious mutation to have powerful, negative effects in small populations. Such studies consistently show that the majority of spontaneous mutations having measurable effects on fitness are detrimental (e.g., Lynch et al. 1999). A common assumption in evolutionary theory is that these mutations are not only irreversible at the molecular level, but have essentially irreversible consequences for population fitness. This is due to the presumed low frequency of both back mutations that restore original DNA sequence and other types of beneficial mutations required to facilitate adaptation or mitigate fitness loss due to accumulated mutation load. Although this view is beginning to be challenged by empirical studies of reverse adaptation (reviewed in Teotónio and Rose 2001), it has not been generally expected that ancestral phenotypes may be easily reattained when ancestral alleles are lost by mutation and random genetic drift.

Another related assumption commonly invoked in theoretical investigations is that mutations affect fitness independently of the genetic background in which they arise. In other words, mutational events are categorized as either beneficial or detrimental for fitness regardless of the genomic context or premutational phenotype. However, if context-dependent expression of mutations exists, as suggested by recent empirical studies (e.g., Lenski and Travisano 1994; Burch and Chao 1999; Elena and Lenski 2001; Maisnier-Patin et al. 2002; Rokyta et al. 2002), it could impact the usefulness of predictions made from various evolutionary and conservation-genetic models (Poon and Otto 2000; Whitlock 2000). Although studies of evolving microbial populations have indicated that beneficial mutations conferring adaptation to a new environment occur only rarely (e.g., Elena et al. 1996), recent studies imply that *compensatory* epistatic mutation may be a powerful force in offsetting the damage caused by more commonly occurring deleterious mutations (Elena et al. 1998; Burch and Chao 1999; Moore et al. 2000; Maisnier-Patin 2002; Rokyta et al. 2002). For example, an isolate of the RNA bacteriophage, ϕ-6, in which fitness was severely reduced by a single spontaneous mutation, was

[2] Present address: Department of Zoology, Oregon State University, Corvallis, Oregon 97331; E-mail: estessu@science.oregonstate.edu.

shown to partially or completely recover original levels of competitive ability on a timescale of 100 generations, depending on the population size at which it was maintained (Burch and Chao 1999; Whitlock and Otto 1999). Fitness improvement in the smaller experimental populations was characterized by incremental fitness increases, as opposed to the case in larger populations, which tended to achieve complete recovery to the ancestral fitness state in a single step. From this, the authors inferred that compensatory epistatic mutations of smaller effect arose more frequently than back mutations that precisely restored ancestral DNA sequence. Further, the nonmutated control showed no fitness increase over the course of the experiment. This implies that a greater potential exists for acquired mutations to improve fitness in the context of fixed detrimental mutations (Whitlock and Otto 1999). Similar results were recently obtained from a study of compensatory evolution in independent lines of *Salmonella typhimurium* which had initially reduced fitness due to fixation of an antibiotic resistance mutation. Here it was confirmed through sequence analysis and genetic reconstitution tests that fitness recovery experienced by the lineages was largely a result of the acquisition of compensatory mutations interacting epistatically with the detrimental resistance mutation, rather than a result of reversions or the accumulation of generally beneficial mutations (Maisnier-Patin et al. 2002).

With the present study, we sought whether the general result of fitness recovery would hold true for a more complex organism harboring variable combinations of deleterious mutations, and what genetic mechanism(s) might be responsible for such recovery. Here we employ long-term mutation-accumulation (hereafter, MA) lines exhibiting significantly reduced fitness (25% on average) relative to their ancestral control as revealed by assays of life-history traits (Vassilieva et al. 2000), chemosensory behavior (B. C. Ajie, S. Estes, M. Lynch, and P. C. Phillips, unpubl. ms.), and body size (Azevedo et al. 2002; S. Estes, B. C. Ajie, M. Lynch, and P. C. Phillips, unpubl. ms.). Analyses of length polymorphisms at 31 microsatellite loci (L. M. Frisse, L. L. Vassilieva, M. Lynch, and W. K. Thomas, unpubl. ms.) and of sequence variation in nearly the entire mitochondrial genome (Denver et al. 2000) and in the nuclear genome (D. R. Denver, K. Morris, S. Estes, M. Lynch, and W. K. Thomas, unpubl. ms.) of each line have revealed a substantial number of mutations at the molecular level. For the current experiment, these previously bottlenecked lines were independently expanded to extremely large population sizes to test whether populations that have amassed substantial mutational loads may regain original levels of fitness by selection for new advantageous mutations.

Materials and Methods

Ancestral Mutation-Accumulation Lines

Our experiment was initiated with 74 lines of *C. elegans*, each derived from mutation-accumulation lines that had been independently maintained by single-individual bottlenecks for an average of 240 generations (Vassilieva et al. 2000). These MA lines were themselves derived from a single, wild-type Bristol-N2 individual from the Caenorhabditis Genetics Center (University of Minnesota, St. Paul, MN). The method of transferring single progeny each generation effectively removes natural selection, allowing mutations with mildly deleterious effects to accumulate essentially freely. Because *C. elegans* reproduces by self-fertilization, this procedure also rapidly removes heterozygosity at all other loci (see fig. 4.15 in Hartl and Clark 1997), a particularly important point for the current study.

By the outset of the current experiment, the original MA lines had developed significant mutation loads, with the mean phenotypes relative to control values having declined by approximately 51% for progeny production and 14% for survival to maturity (see fig. 2 in Vassilieva et al. 2000). Our prior indirect estimates of the diploid genomic mutation rate for these two traits, approximately 0.033 and 0.003 per generation, respectively, lead to the prediction that the average MA line at generation 240 was fixed for 4.0 deleterious mutations affecting productivity and for 0.4 mutations affecting survival to maturity. These observations are consistent with average effects of fixed individual mutations in these lines being about 13% relative to the control mean phenotype for productivity and about 35% for survival.

Generation and Maintenance of Large-Population-Size Lines

For the current study, each line remaining after 240 generations of mutation accumulation was separately expanded and maintained at large population sizes by transferring agar chunks containing well over 1000 individuals to fresh plates with a sterilized scalpel every four days (equivalent to approximately one generation) (hereafter referred to as MA-R lines for mutation-recovery). This time period was adequate to insure highly competitive conditions, as population sizes had generally reached several thousands (>5000) of individuals prior to each transfer, with the animals being starved to the extent that cessation of egg laying had occurred. Great care was taken to avoid cross contamination among lines by keeping Petri plates well separated and wrapped in parafilm. All lines were maintained under standard laboratory conditions (at 20° on 60 $\times$ 15 mm Petri plates with NGM agar uniformly seeded with a 80–90 μL suspension of OP50 *Escherichia coli*) (Sulston and Hodgkin 1988).

In the absence of genetic variation, beneficial mutation accumulation provides the only route to fitness recovery. There are three genetic mechanisms by which this can occur: back or reversion mutation, the accumulation of mutations beneficial irrespective of genetic background, or epistatic mutations that specifically ameliorate the phenotypic effects of previously accrued detrimental mutations. To test whether any fitness gains shown by the MA-R lines could be due to a generic form of laboratory adaptation (i.e., due to unconditionally beneficial mutations), 30 lines were also generated from the ancestral (time zero, premutation accumulation) control animals (previously stored cryogenically) and maintained in the same manner as outlined above. Henceforth, these will be referred to as C-R lines, for control-recovery.

Stemming from concerns regarding the potentially selective effects of cryogenic storage, we tested the effect of (short-term) freezing of lines on the two life-history traits in conjunction with this second recovery assay. Two replicates

of each line from the above four treatments (MA, MA-R, Control, and C-R) were frozen at −80° using soft agar freezing solution (Lewis and Flemming 1995). After approximately one week, lines were thawed and taken through two generations of single-individual bottlenecks to avoid maternal and grandmaternal effects, then assayed in parallel with their unfrozen counterparts. The ''unfrozen'' ancestral control was thawed and maintained by single-individual transfer for 10 generations prior to this assay.

Although *C. elegans* self-fertilizes the majority of the time, we noted that most large-population lines were capable of producing a near ''wild-type'' proportion of males (~0.2%) after a few generations. Thus, there existed some opportunity for recombination via outcrossing, such that beneficial mutations arising in different members of the same line could be paired within a descendant genome rather than necessitating their consecutive origination within the same lineage. To the extent that this occurs, more rapid fitness recovery would be possible than under pure selfing.

Fitness Assay

After 10 and 80 generations of large-population-size treatment in the MA-R lines and after 70 generations in the C-R lines, assays of survival to maturity and progeny production were conducted on single individuals in benign environments as in Vassilieva and Lynch (1999). These assays were conducted in parallel with both the ancestral control (which had been frozen prior to the MA experiment) and the contemporary MA lines. In the first assay, 20 ancestral control lines, 74 MA lines, and the 74 corresponding MA-R lines were measured. In the second assay, 30 lines each of the ancestral control, MA, corresponding MA-R, and C-R lines were included. Prior to each assay, all lines were expanded into five replicates (four replicates for the second assay) and maintained by transferring single individuals for two generations. Fitness components were then measured using single, third-generation individuals from each replicate. The same general procedure of line subdivision was applied to thawed control animals (20 animals for the first assay and 30 for the final assay). Survival to maturity was scored as 1 if an animal survived from the first larval stage to produce even a single, surviving offspring, otherwise as 0. Single individuals were transferred to fresh plates daily and progeny production was measured by directly counting the offspring produced over the first four days of life (which covers the majority of the reproductive period).

Potential for Line Contamination

Although great care was taken to avoid cross-contamination of individual MA-R lines, accidental gene flow among lines might provide a route to fitness recovery. To test this possibility, D. R. Denver (Indiana University, Bloomington, IN) sequenced DNA from portions of three haphazardly chosen MA-R line genomes found to harbor unique mutations in a previous survey of mitochondrial and nuclear sequence variation (Denver et al. 2000; D. R. Denver, K. Morris, S. Estes, M. Lynch, and W. K. Thomas, unpubl. ms.). All examined lines continued to be fixed for their ancestral markers, as confirmed by sequences from both strands of DNA. Although this result does not exclude the possibility of contamination in the remaining lines, it indicates that gene flow among MA-R lines was unlikely, providing further credence to the argument that any observed recovery must be due to selection for new mutations.

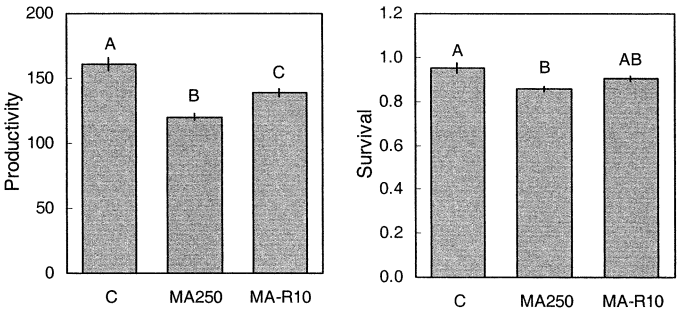

FIG. 1. Mean progeny production and survival to maturity for MA-R lines after 10 generations of large-population size (MA-R10), the ancestral control (C), and contemporaneous MA lines (MA250). Error bars represent one standard error. For each bar graph, groups of means labeled with different letters indicate a significant difference at the $\alpha = 0.05$ level (Tukey HSD; Zar 1999).

Statistical Analysis

To test for differences between pairs of treatment group means, least-squares contrasts (Tukey's HSD for all pairwise comparisons; Zar 1999) were performed using JMP 4.0 (SAS Institute 2000) on the data for each life-history trait and each of the two assays. Taking the paired nature of the MA and MA-R lines into account, results for these groups were further analyzed using a two-factor ANOVA with treatment (MA versus MA-R) as a fixed effect and line as a random effect for each trait and assay using SPSS 11.0 (SPSS, Inc. 2001). For a better comparison of results from the first and second assays, we also analyzed the subset of MA and MA-R line pairs from the first assay that were included in the second assay (reported as ''Assay 1 subset''). Finally, a one-way ANOVA was performed separately for both traits in each assay of each treatment to partition the total phenotypic variance into within- and among-line components. With our experimental design, the among-line variance component provides an estimate of the total genetic variance present for each trait, whereas the within-line variance estimates the environmental variance.

RESULTS

Fitness Assays

After 10 generations of large-population-size treatment, fitness of the MA-R lines was assessed in parallel with MA generation 250 (maintained by single-individual bottlenecks since the beginning of the recovery experiment) and the ancestral control. Despite this short period of time, mean fitness of the MA-R lines had rebounded substantially, approximately 11% for progeny production and 5% for survival to maturity. The increase was statistically significant for progeny production (Fig. 1). Further, comparing the MA and MA-R groups, the treatment-by-line interaction terms were significant for both traits (Tables 1, 2), indicating that individual

TABLE 1. Two-factor ANOVA for progeny production in MA and MA-R lines using entire dataset from the first fitness assay (Assay 1), the subset of 30 of these lines that went on to be included in the second assay (Assay 1 subset), and the 30 lines reassayed after 80 generations of large-population-size treatment (Assay 2).

Source	SS	df	MS	*F*	*P*
		Assay 1			
Treatment	39605.306	1	39605.306	5.596	0.021
Line	3304068.52	73	45261.213	6.394	<0.001
Treatment × Line	516705.337	73	7078.155	2.785	<0.001
Error	1501789.050	591	2541.098		
		Assay 1 (subset)			
Treatment	25613.280	1	25613.280	3.161	0.086
Line	1038636.947	29	35815.067	12.610	<0.001
Treatment × Line	235019.320	29	8104.114	2.853	<0.001
Error	681676.000	240	2840.317		
		Assay 2			
Treatment	155161.166	1	155161.166	13.255	0.001
Line	587453.163	29	20257.006	1.624	0.094
Treatment × Line	357793.705	29	12337.714	3.563	<0.001
Error	533326.917	154	3463.162		

MA lines responded differently to the large-population-size treatment. Among-line variance components for both the control and MA treatments (see results for ''Assay 1'' in Table 3) were compatible with previous results from the long-term mutation-accumulation experiment (see fig. 3 in Vassilieva et al. 2000). The estimated variance among MA-R lines for both characters tended to decrease relative to the MA estimate, although these reductions are not significant (Table 3).

At this point, to determine the degree to which fitness gains shown by the MA-R lines might be due to the accumulation of mutations conferring adaptation to the lab environment (i.e., beneficial in any genetic background) rather than to compensatory mutations specific to the accumulated mutation load, 30 C-R lines generated from thawed, ancestral control (premutation accumulation) animals were initiated and maintained at large population sizes in the same manner as the MA-R lines. These lines had been through 70 generations of such treatment at the time of a second (generation 80) assay of the MA-R lines. If the ancestral control strain (Bristol N2) is at or near a fitness peak with respect to the laboratory environment, little or no fitness improvement would be expected for the C-R lines. This is because there would be little or no opportunity for fitness improvement in the context of deleterious mutations.

TABLE 2. Two-factor ANOVA for survival to maturity in MA and MA-R lines using entire dataset from the first fitness assay (Assay 1), the subset of 30 of these lines that were included in the second assay (Assay 1 subset), and the 30 lines assayed after 80 generations of large-population-size treatment (Assay 2).

Source	SS	df	MS	*F*	*P*
		Assay 1			
Treatment	0.164	1	0.164	1.091	0.300
Line	27.034	73	0.370	2.472	<0.001
Treatment × Line	10.936	73	0.150	2.609	<0.001
Error	34.000	592	0.057		
		Assay 1 (subset)			
Treatment	0.120	1	0.120	0.946	0.339
Line	6.747	29	0.233	1.833	0.054
Treatment × Line	3.680	29	0.127	2.307	<0.001
Error	13.200	240	0.055		
		Assay 2			
Treatment	0.881	1	0.881	5.253	0.029
Line	4.697	29	0.162	0.906	0.604
Treatment × Line	5.184	29	0.179	7.865	<0.001
Error	3.500	154	0.023		

In this second assay, 30 randomly chosen pairs of MA-R and MA lines were surveyed for fitness in parallel with the ancestral control and with the C-R lines. At the time of this assay, the MA lines had reached 280 generations on average. Our results indicate that the MA-R lines had fully recovered on average for both fitness-related characters, whereas the C-R lines showed no significant fitness gains compared to the ancestral control (Fig. 2, Table 3). Again, focusing on only the MA and MA-R lines, the treatment-by-line interaction terms were significant for both traits, signifying differential response of MA lines to the large-population size treatment. This is evident in the pattern of change in mean progeny production for the subset of 30 individual lines across generations of large-population-size treatment (Fig. 3). After 80 generations of large-population-size treatment, estimates of among-line variance for the MA-R lines show further (nonsignificant) reductions, whereas those for the other treatments are again similar to those reported previously (Vassilieva et al. 2000).

Although results from the final assay demonstrated that the effects of freezing and thawing on the fitness of control and MA lines are not statistically significant (data not shown), a trend did exist such that the MA lines revived from −80° storage (a common practice in nematological studies) performed slightly better in fitness assays than their unfrozen counterparts. We view this pattern as consistent with a scenario of fitness recovery due to selection for newly arisen beneficial or compensatory mutations, as (like our MA-R and C-R treatments) freezing and subsequent thawing involve the expansion of individual lines to the extremely large population sizes that promote such change (Lewis and Fleming 1995). Nevertheless, because of the statistical nonsignificance, we pooled frozen and unfrozen treatments of each of

TABLE 3. One-way ANOVA for each trait and population-size-treatment for the first fitness assay (Assay I), the subset of 30 of these lines that were included in the second assay (Assay 1 subset), and the 30 lines assayed after 80 generations of large-population-size treatment (Assay 2).

Treatment	Progeny production			Survival to maturity		
	Mean	Within-line variance	Between-line variance	Mean	Within-line variance	Between-line variance
			Assay I			
Control	160.58 (13.86)	3375 (520)	3168 (1220)	0.952 (0.024)	0.043 (0.006)	0.003 (0.004)
MA250	120.01 (8.06)	2163 (165)	4908 (840)	0.863 (0.027)	0.068 (0.005)	0.050 (0.010)
MA-R10	138.72 (8.41)	3046 (250)	4552 (855)	0.908 (0.025)	0.047 (0.004)	0.037 (0.007)
			Assay I (subset)			
MA250	134.61 (12.24)	2783 (359)	3790 (1143)	0.893 (0.040)	0.063 (0.008)	0.033 (0.012)
MA-R10	149.84 (12.50)	2826 (384)	3500 (1130)	0.926 (0.033)	0.052 (0.007)	0.018 (0.008)
			Assay 2			
Control	189.34 (5.55)	2195 (400)	720 (390)	0.983 (0.011)	0.016 (0.003)	0.000 (0.002)
MA-280	138.06 (12.53)	2238 (345)	4109 (1235)	0.886 (0.053)	0.027 (0.004)	0.076 (0.021)
MA-R80	182.94 (12.01)	4708 (722)	3179 (1163)	0.974 (0.016)	0.023 (0.004)	0.002 (0.002)
C-R70	197.94 (7.36)	6400 (1168)	0.000 (804)	0.991 (0.008)	0.008 (0.002)	0.000 (0.001)

the four groups. This causes a slight upward bias in MA line means, making our inference of recovery more conservative.

Only two MA lines went extinct between generations 240 and 280, so between-MA line selection resulting in loss of MA and MA-R line pairs could not have accounted for a substantial portion of the increase in fitness of the MA-R lines observed in the second assay. Sampling 30 of the 72 available line pairs to include in the second assay may have resulted in a similar bias if, by chance, line pairs exhibiting a greater difference between MA and MA-R means were sampled. For a more direct comparison of the results of the first and second assays, means and variance components (Table 3), as well as ANOVA results (Tables 1, 2), are reported for the subset of 30 lines included in both assays (i.e., compare ''Assay 1 subset'' and ''Assay 2''). Note that when focusing upon this subset of 30 lines, the difference between MA and MA-R treatments for productivity becomes nonsignificant in the first assay (''Assay 1 subset'' in Table 1). The most important point of this comparison, however, is that the mean fitness of MA-R lines exhibits an increase between 10 and 80 generations of large-population-size treatment even when only the subset of 30 lines is focused upon (also see Fig. 3). This indicates that the general result of fitness recovery is not a spurious consequence of sampling a subset of lines for the second assay.

Nevertheless, to gain further insight into the magnitude of the potential bias caused by this sampling, we can compare the amount of recovery shown by the full set of MA-R lines after 10 generations to that shown by the subset of these lines that were reassayed after 80 generations of large-population-size treatment. The subset of 30 lines produced 15.3 more offspring on average than the full complement of 74 lines (calculated from values in Table 3). The difference between the means of the MA280 and MA-R80 treatments is 44.9 (Table 3). From this, a corrected estimate of recovery in progeny production can be calculated as: 44.9 − 15.3, or 29.6 additional progeny per generation. Doing the same for survival to maturity, we estimate the total amount of recovery for this trait after 80 generations to be 0.055, rather than 0.088.

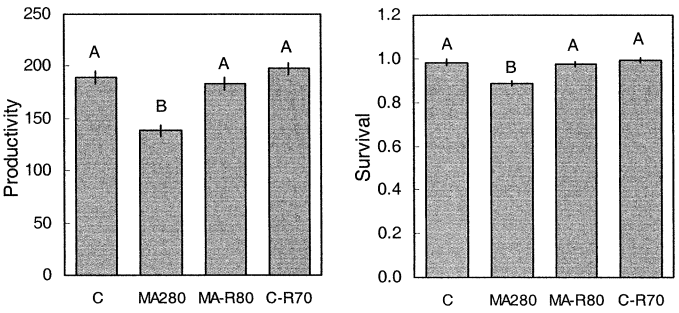

FIG. 2. Mean progeny production and survival to maturity for MA-R lines after 80 generations of large population size (MA-R80), the ancestral control (C), contemporaneous MA lines (MA280), and the expanded control after 70 generations under large population conditions (C-R70). Error bars represent one standard error. For each bar graph, groups of means labeled with different letters indicate a significant difference at the $\alpha = 0.05$ level (Tukey HSD; Zar 1999).

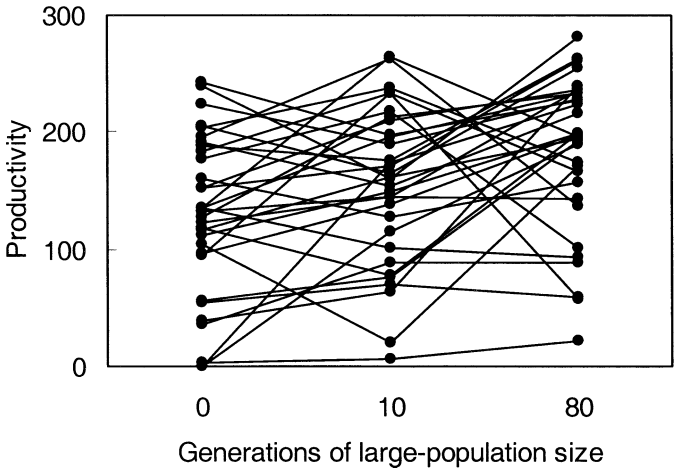

FIG. 3. Trajectories of mean progeny production of individual MA lines across successive generations of large-population-size treatment. Values for generation zero are equivalent to the MA250 line means.

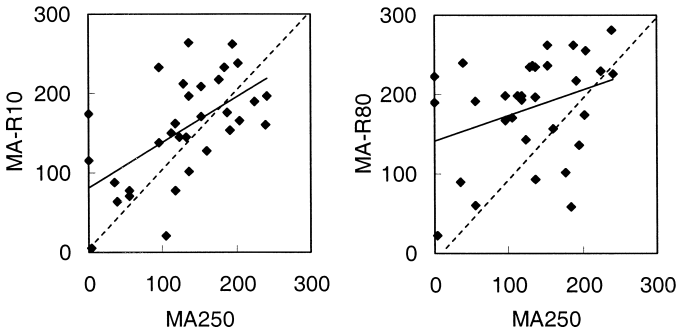

FIG. 4. Regression of line means for progeny production of MA-R lines after 10 (left) and 80 (right) generations of large-population-size treatment on corresponding generation-250 MA lines (MA250). For the two respective regressions, slope (SE) = 0.55 (0.14) and $P < 0.001$; slope (SE) = 0.32 (0.17) and $P = 0.062$. The slope of the dashed line represents a 1:1 correspondence.

It is not appropriate to correct the above computations for the effects of additional mutations accumulated in the MA lines during the generations of single-individual bottlenecking experienced between the first and second assays. This is because the MA lines exhibited no additional decline in mean phenotype during this period of time (Figs. 1, 2). The downward trend in average MA line fitness (see fig. 2 in Vassilieva et al. 2000) continues, although results from immediately adjacent assays are sometimes not statistically significantly different. In any case, because no additional decline in the mean trait values for the MA lines was observed between generations 240 and 280, we do not correct the difference in MA-R80 and MA280 means from the second assay for the effects of mutation accumulation.

Pattern of Fitness Recovery

Under a scenario of fitness recovery due to compensatory mutation accumulation, the slope of the regression of recovery (MA-R) lines on parental MA lines is expected to be less than one, as the lines showing high initial fitness are expected to show little or no improvement, whereas the least fit lines have the maximum opportunity for improvement. In addition, the slope of this regression is expected to decline in subsequent assays, as the lines most loaded with mutations exhibit the greatest recovery response. Focusing on only the 30 pairs of MA and MA-R lines that were included in the second assay, both patterns are apparent in the data (Fig. 4).

DISCUSSION

Although recent theory suggests that reverse mutation could slow fitness decline and time to population extinction (Schultz and Lynch 1997; Lande 1998), it is an empirical issue as to whether such mutations are actually frequent enough to retard or halt these processes. We show that when returned to a population-genetic environment that is conducive to efficient natural selection, mutationally degraded lines are capable of recovering original levels of mean fitness at a rate that is at least three times that of mutational degradation in the absence of selection, although there is variation in response among individual lines. The possible underlying mechanisms driving this recovery are the accumulation of: (1) back (reversion) mutations; (2) unconditionally beneficial mutations, i.e., those that increase fitness irrespective of genetic background; (3) compensatory epistatic mutations; and (4) selection for segregating mutational variance within mutation-accumulation lines.

Although any mechanism of fitness recovery involving the accumulation of new mutations would be an important result, several lines of evidence suggest that fitness recovery observed in the MA-R lines was largely due to compensatory mutation accumulation. First, due to the long-term maintenance of the ancestral MA lines by single-progeny descent and the low estimated genomic deleterious mutation rates, essentially all of the initial lines were expected to be devoid of mutations segregating for fitness. Applying the genotypic transition-probability matrix (Ewens 1979) reveals that the mean number of generations that a newly arisen mutation remains segregating in a selfing population is no greater than three (including the initial generation of appearance) for an allele with additive effects on fitness and no greater than four for a recessive allele. Thus, with a diploid genomic mutation rate for productivity of about 0.033 per generation, the average number of segregating mutations expected within a line is only about 0.1; or in other words, about 10% of the lines are expected to have contained a single segregating locus for productivity, with the remaining lines being completely devoid of variation for this trait. With a genomic mutation rate of about 0.003 for survival to maturity, only 1% of the initial lines are expected to have contained a single segregating mutation for this trait. Given these initial frequencies and our previous estimates of the homozygous effects of mutations in these lines (Vassilieva et al. 2000), the maximum average improvement in fitness resulting from the complete fixation of such mutations would be on the order of 1% or less, which is more than an order of magnitude less than what was actually observed. Although additional segregating alleles might be maintained in the MA lines by overdominance, we have no evidence for the presence of such polymorphisms and, in any event, these cannot contribute to selection response.

Second, the lack of overall fitness increase for the C-R lines provides both corroborating evidence for the rarity of segregating mutations in our base populations as well as support for the argument that the improvement in fitness in the MA-R lines was at least not entirely a simple consequence of mutationally driven adaptation to the lab environment. The lack of fitness increase in the C-R lines is perhaps not surprising given that the Bristol-N2 strain of *C. elegans* was maintained in laboratory culture for several years before a frozen stock was established (Brenner 1974; Hodgkin and Doniach 1997). As a consequence of such long-term treatment, the N2 strain may be nearly maximally adapted to common laboratory conditions. If this had not been the case, to demonstrate support for the compensatory mutation hypothesis, we would have been required to show that MA-R lines increased in fitness compared to their parental MA lines by an amount beyond that of any fitness increase exhibited by the expanded control (C-R) lines.

Third, our analyses of molecular markers that had previously appeared endogenously within specific MA lines provide no evidence for contamination among lines. While this

does not entirely rule out the possibility of contamination in other MA-R lines, given the physical nature of our experimental protocols, it is unlikely that any contamination occurred. Moreover, it is significant that the three MA-R lines checked for contamination did exhibit moderate to very large fitness increases meaning that the results from these lines can be reasonably extrapolated to the other, unchecked, lines. For example, after 80 generations of large-population-size treatment, these particular lines exhibited increases in average progeny production of 223.5, 37.8 and 13.8 compared with their parental MA lines. With an average improvement of 91.7, these values are even more extreme than the average increase shown by all MA-R lines of 29.6 (corrected for sampling of lines in Assay 2, see Results) or 44.9 (uncorrected, see Table 3) more progeny produced than the average MA line. This indicates that fitness recovery was indeed possible in the absence of accidental gene flow among MA-R lines.

It is also of note that a number of MA lines had acquired severe morphological aberrations (see inset in *Science* 2000, Vol. 289) and/or extreme body sizes prior to the initiation of the current experiment (Azevedo et al. 2002; S. Estes, B. C. Ajie, M. Lynch, and P. C. Phillips, unpubl. ms.). Surprisingly, although these phenotypes were retained by the MA lines throughout the course of the recovery experiment, the morphological abnormalities disappeared in the MA-R lines and the body-size abnormalities generally showed reversion to the ancestral phenotype prior to the final fitness assay. Specifically, one MA line with both an uncoordinated (Unc) phenotype in hermaphrodites as well as a high incidence of males (Him) lost both of these features after only 10 generations of large-population-size treatment. And, of the lines with morphological defects included in the second recovery assay, one exhibited a long body (Lon) phenotype, while the other expressed ''social'' rather than solitary foraging even under low density conditions with ample food (e.g., Sokolowski 2002). Both of these lines had lost their mutant phenotypes by the time of the second assay. (As the latter phenotypes are difficult to discern when extremely large populations of worms are present on a Petri dish, we do not know precisely at what point during the large-population-size treatment these losses occurred.)

Lastly, as the per nucleotide rate of mutation for nuclear DNA in this *C. elegans* strain is approximately 4.4×10^{-8} ($\pm 9.8 \times 10^{-9}$) per generation (D.R. Denver, K. Morris, S. Estes, M. Lynch, and W.K. Thomas, unpubl. ms.) and the maximum population size obtained by an MA-R line prior to each passage being fewer than 10,000 individuals, back mutations (even if one assumes that all mutations in the MA lines were single-nucleotide changes) could not account for the total amount of fitness increase shown by the MA-R lines after 80 generations, i.e., because $(4.4 \times 10^{-8} \times 10{,}000 \times 80) < 1$. More specifically, for each previously accumulated mutation, this exercise yields a probability of only about 0.036 that a single individual per line will obtain a back mutation. Further, only a small fraction of such mutations are likely to reach fixation. The asymptotic probability is equivalent to twice the selection coefficient when population size is large.

Having ruled out both the accumulation of back and/or generally beneficial mutations as well as selection acting on segregating genetic variation within lines as the sole drivers of fitness recovery, compensatory mutation appears to be the only viable explanation for the observed results. Given that the observed fitness restoration did in fact result from compensatory mutation accumulation, our findings suggest that a surprisingly high fraction of deleterious mutations can be compensated and that, following the same logic as above, there are potentially a large number of ways to ameliorate the negative effects of a given mutation in this system. These observations support the conclusion that the recovery of the MA lines was at least partly a consequence of the accumulation of new mutations specifically advantageous in the various mutant backgrounds. This contention is further consistent with the observation that the magnitude of line-specific fitness improvement was inversely proportional to ancestral-line fitness (Fig. 4). The overall fitness increase observed in the MA-R lines was dominated by a few lines that expressed particularly low performance as MA lines, whereas the lines with higher initial levels of fitness mainly showed less or no improvement under large-population conditions. Such a trend provides still additional evidence that recovery was not an artifact of generic adaptation to the laboratory setting, as this form of adaptation would be expected to influence all lines equally. Further, this result is congruent with the idea that the probability of a population acquiring a compensatory mutation increases with its distance from a phenotypic optimum (Wagner and Gabriel 1990; Whitlock 2000).

The potential for compensatory evolution to drive recovery of fitness depressed as a result of major-effect mutations, such as those conferring drug resistance, has precedent in studies of experimental viral and bacterial evolution (e.g., Burch and Chao 1999; Levin et al. 2000; Moore et al. 2000; Rokyta et al. 2002; Maisnier-Patin et al. 2002). However, such a result has not been shown for a complex metazoan exhibiting low fitness as a result of numerous and variable combinations of spontaneously accumulated detrimental mutations. According to Fisher's geometric model of adaptation, mutations of a given size are less likely to be beneficial for the fitness of a complex than a simple organism. This is because the probability that a mutation will be favorable becomes exceedingly small as the number of interacting components of an organism's genotype or phenotype increases (Fisher 1930, chapter 2; Orr 2000; Poon and Otto 2000). In combination with results from viruses and bacterial species, our findings suggest that neither the increased complexity of the *C. elegans* genome (e.g., increased genome size, functional redundancy, and more highly evolved regulatory processes), nor its increased morphological complexity, entirely preclude fitness compensation. This indicates that complexity does not, in this case, necessarily hinder the genomic response to mutational damage, although we can make no statements about the source of the departure of our results from predictions of recent models (e.g., large average effects of beneficial mutations in *C. elegans* compared to microbes and/or increased selection intensity in our experiment; Orr 2000).

Given the theoretically expected threat of deleterious-mutation accumulation to small natural or captive-bred populations, the results of our study are encouraging. Although the immediate causes of recent species extinctions may have

been mainly nongenetic (e.g., habitat destruction, introduction of exotic species, and exploitation by humans; Hedrick and Miller 1992), genetic factors including deleterious-mutation accumulation are likely to be a significant threat to small populations living in protected environments, especially as current management practices, including manipulation of environments and/or equalizing family or lineage contributions, inadvertently act to relax natural selection (Hedrick and Miller 1992; Lande 1995b; Lynch and O'Hely 2001; Ford 2002). Our study indicates that compensatory or beneficial mutations may arise frequently enough that, provided an *appropriate* selection regime is in place, large population size can lead to substantial recovery from accumulated mutational damage on a time scale relevant to conservation biology. Nevertheless, before an overly optimistic picture is painted of the potential for fitness compensation, it should be noted that not all MA-R lines showed high levels of fitness at the time of the final assay and that some of these lines actually showed substantial fitness *reductions* compared to their parental MA lines (Fig. 4). The former may indicate the existence of a class of mutations whose effects cannot be compensated. The latter is potentially a result of genotype-by-environment interaction such that beneficial mutations accumulated during the large-population-size treatment phase proved detrimental in the low-density assay environment. However, the more parsimonious explanation is that this pattern is simply due to environmental noise. In any case, this indicates that although recovery of ancestral phenotypes may be possible on average, one should be extremely cautious in making predictions regarding the trajectories of individual populations.

Finally, recent results from our ongoing mutation-accumulation experiment suggest the possibility that the fitness recovery observed in the current experiment may have been greatly facilitated by positive pleiotropic mutational effects (S. Estes, B. C. Ajie, M. Lynch, and P. C. Phillips, unpubl. ms.). We find evidence for extensive and highly positive correlations among deleterious mutations influencing life-history and morphological characters, a not entirely unprecedented result (e.g., Yoshimaru and Mukai 1985; Houle et al. 1994, but see Fernández and López-Fanjul 1996). If such synergistic pleiotropy is a general feature of both deleterious and beneficial mutations, it could have aided in the process of fitness recovery. This is because a compensatory mutation producing positive pleiotropic effects on multiple fitness components will enjoy a much greater selective advantage than an allele acting to compensate only single traits (e.g., Poon and Otto 2000).

The outcome of this experiment illustrates the powerful efficiency with which evolution can contend with the cumulative effects of deleterious mutation. Due to their apparent rarity, the incidence and properties of compensatory and beneficial mutations are virtually always obscured by the sea of deleterious mutations generated in standard mutation-accumulation experiments in which selection is largely relaxed. Nevertheless, numerous evolutionary genetic models clearly indicate the disproportionate role that beneficial and/or compensatory mutations can play in defining the genetic features of natural populations, including those under conservation-genetic management (Wagner and Gabriel 1990; Schultz and Lynch 1997; Lande 1998; Whitlock and Otto 1999; Bataillon 2000; Poon and Otto 2000; Whitlock 2000). To bring the theory closer to reality, there is a clear need for further empirical work on mutationally driven fitness compensation, the beneficial mutation rate (Shaw et al. 2002), the spectrum of fitness effects of such mutations, as well as the effect of environmental context on these parameters (e.g., Bjorkman et al. 2000).

Acknowledgments

We thank B. Ajie and S. Wong for invaluable laboratory assistance, D. Denver for conducting the sequence analyses, P. Phillips for helpful comments and discussion, C. Baer, J. Colbourne, and K. Lacy for comments on an earlier draft, L. Vassilieva for assistance with the first fitness assay, reviewer, A. Kondrashov, and an anonymous reviewer for insightful comments on the manuscript. This work was supported by National Institutes of Health grant RO1-GM36827 to ML and by training grants from the National Science Foundation (DBI-9413223) and the U.S. Public Health Service (GM-07413) to SE.

Literature Cited

Azevedo, R. B. R., P. D. Keightley, C. Laurén-Määttä, L. L. Vassilieva, M. Lynch, and A. M. Leroi. 2002. Spontaneous mutational variation for body size in *Caenorhabditis elegans*. Genetics 162:755–765.

Bataillon, T. 2000. Estimation of spontaneous genome-wide mutation rate parameters: whither beneficial mutations? Heredity 84:497–501.

Bjorkman, J., I. Nagaev, O. G. Berg, D. Hughes, and D. I. Andersson. 2000. Effects of environment on compensatory mutations to ameliorate costs of antibiotic resistance. Science 287: 1479–1481.

Brenner, S. 1974. The genetics of *Caenorhabditis elegans*. Genetics 77:71–94.

Burch, C. L., and L. Chao. 1999. Evolution by small steps and rugged landscapes in the RNA virus ϕ-6. Genetics 151:921–927.

Caballero, A., E. Cusi, C. García, and A. García-Dorado. 2002. Accumulation of deleterious mutations: additional *Drosophila melanogaster* estimates and simulation of the effects of selection. Evolution 56:1150–1159.

Chavarrías, D., C. López-Fanjul, and A. García-Dorado. 2001. The rate of mutation and the homozygous and heterozygous mutational effects for competitive viability: a long-term experiment with *Drosophila melanogaster*. Genetics 158:681–694.

Crow, J. F. 1997. The high spontaneous mutation rate: is it a health risk? Proc. Natl. Acad. Sci. USA 94:8380–8386.

Denver, D. R., K. Morris, M. Lynch, L. L. Vassilieva, and W. K. Thomas. 2000. High direct estimate of the mutation rate in the mitochondrial genome of *C. elegans*. Science 289:2342–2344.

Elena, S. F., and R. E. Lenski. 2001. Epistasis between new mutations and genetic background and a test of genetic canalization. Evolution 55:1746–1752.

Elena, S. F., V. S. Cooper, and R. E. Lenski. 1996. Punctuated evolution caused by rare beneficial mutations. Science 272: 1802–1804.

Elena, S. F., M. Davila, I. S. Novella, J. J. Holland, E. Domingo, and A. Moya. 1998. Evolutionary dynamics of fitness recovery from the effects of Muller's ratchet. Evolution 52:309–314.

Ewens, W. J. 1979. Mathematical population genetics. Springer, Heidelberg, Germany.

Eyre-Walker, A., and P. D. Keightley. 1999. High genomic deleterious mutation rates in hominids. Nature 397:344–347.

Fernández, J., and C. López-Fanjul. 1996. Spontaneous mutational variances and covariances for fitness related traits in *Drosophila melanogaster*. Genetics 143:829–837.

Fisher, R. A. 1930. The genetical theory of natural selection. Oxford Univ. Press, Oxford, U.K.
Ford, M. 2002. Selection in captivity during supportive breeding may reduce fitness in the wild. Conserv. Biol. 16:815–825.
Fry, J. D. 2001. Rapid mutational declines of viability in *Drosophila*. Genet. Res. 77:53–60.
Fry, J. D., and S. L. Heinsohn. 2002. Environmental dependence of mutational parameters for viability in *Drosophila melanogaster*. Genetics 161:1155–1167.
García-Dorado, A., C. López-Fanjul, and A. Caballero. 1999. Properties of spontaneous mutation affecting quantitative traits. Genet. Res. 74:341–350.
Hartl, D. L., and A. G. Clark. 1997. Principles of population genetics, 3rd ed. Sinauer, Sunderland, MA.
Hedrick, P. W., and P. S. Miller. 1992. Conservation genetics: techniques and fundamentals. Ecol. Appl. 2:30–46.
Hodgkin, J., and T. Doniach. 1997. Natural variation and copulatory plug formation in *Caenorhabditis elegans*. Genetics 146: 149–164.
Houle, D., K. A. Hughes, D. K. Hoffmaster, J. Ihara, S. Assimacopoulos, D. Canada, and B. Charlesworth. 1994. The effects of spontaneous mutation on quantitative traits. I. Variances and covariances of life history traits. Genetics 138:773–785.
Keightley, P. D., and A. Eyre-Walker. 1999. Terumi Mukai and the riddle of deleterious mutation rates. Genetics 153:515–523.
Kondrashov, A. S. 1995. Contamination of the genome by very slightly deleterious mutations: why have we not died 100 times over? J. Theor. Biol. 175:583–594.
Lande, R. 1994. Risk of population extinction from fixation of new deleterious mutations. Evolution 48:1460–1469.
———. 1995a. Mutation and conservation. Conserv. Biol. 9: 782–791.
———. 1995b. Breeding plans for small populations based on the dynamics of quantitative genetic variance, Pp. 318–340 *in* J. D. Ballou, M. Gilpin, and T. J. Foose, eds. Management for survival and recovery: analytical methods and strategies in small population conservation. Columbia Univ. Press, New York.
———. 1998. Risk of population extinction from fixation of deleterious and reverse mutations. Genetica 102/103:21–27.
Lenski, R. E., and M. Travisano. 1994. Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. Proc. Natl. Acad. Sci. USA 91:6808–6814.
Levin, B. R., V. Perrot, and N. Walker. 2000. Compensatory mutations, antibiotic resistance and the population genetics of adaptive evolution in bacteria. Genetics 154:985–997.
Lewis, J. A., and J. T. Fleming. 1995. Basic culture methods. Pp. 3–29 *in* H. F. Epstein and D. C. Shakes, eds. *Caenorhabditis elegans*: modern biological analysis of an organism. Academic Press, San Diego, CA.
Lynch, M., J. Blanchard, D. Houle, T. Kibota, S. Schultz, L. L. Vassilieva, and J. Willis. 1999. Spontaneous deleterious mutation. Evolution 53:645–663.
Lynch, M., J. Conery, and R. Burger. 1995. Mutation accumulation and the extinction of small populations. Am. Nat. 146:489–518.
Lynch, M., and W. Gabriel. 1990. Mutation load and the survival of small populations. Evolution 44:1725–1737.
Lynch, M., and M. O'Hely. 2001. Supplementation and the genetic fitness of natural populations. Conserv. Gen. 2:363–368.
Maisnier-Patin, S., O. G. Berg, L. Liljas, and D. I. Andersson. 2002. Compensatory adaptation to the deleterious effects of antibiotic resistance in *Salmonella typhimurium*. Mol. Microbiol. 46: 355–366.
Moore, F. B. G., D. E. Rozen, and R. E. Lenski. 2000. Pervasive compensatory adaptation in *Escherichia coli*. Proc. R. Soc. Lond. B 267:515–522.
Muller, H. J. 1950. Our load of mutations. Am. J. Hum. Gen. 2: 111–176.
Orr, H. A. 2000. Adaptation and the cost of complexity. Evolution 54:13–20.
Poon, A., and S. P. Otto. 2000. Compensating for our load of mutations: freezing the meltdown of small populations. Evolution 54:1467–1479.
Rokyta, D., M. R. Badgett, I. J. Molineaux, and J. J. Bull. 2002. Experimental genomic evolution: extensive compensation for loss of DNA ligase activity in a virus. Mol. Biol. Evol. 19: 230–238.
SAS Institute. 2000. JMP 4.0 statistics and graphics guide. SAS Institute, Cary, NC.
Schultz, S. T., and M. Lynch. 1997. Mutation and extinction: the role of variable mutational effects, synergistic epistasis, beneficial mutations, and degree of outcrossing. Evolution 51: 1363–1371.
Shaw, F. H., C. J. Geyer, and R. G. Shaw. 2002. A comprehensive model of mutations affecting fitness and inferences for *Arabidopsis thaliana*. Evolution 56:453–63.
Sokolowski, M. B. 2002. Social eating for stress. Nature 419: 893–894.
SPSS. 2001. SPSS base user's guide. SPSS, Inc., Chicago, IL.
Sulston, J., and J. Hodjkin. 1988. Methods. Pp. 587–606 *in* W. B. Wood, ed. The nematode *Caenorhabditis elegans*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Teotónio, H., and M. R. Rose. 2001. Reverse evolution. Evolution 55:653–660.
Vassilieva, L. L., and M. Lynch. 1999. The rate of spontaneous mutation for life-history traits in *Caenorhabditis elegans*. Genetics 151:119–129.
Vassilieva, L. L., A. M. Hook, and M. Lynch. 2000. The fitness effects of spontaneous mutations in *Caenorhabditis elegans*. Evolution 54:1234–1247.
Wagner, G. P., and W. Gabriel. 1990. Quantitative variation in finite parthenogenetic populations: what stops Muller's ratchet in the absence of recombination? Evolution 44:715–731.
Whitlock, M. 2000. Fixation of new alleles and extinction of small populations: drift load, beneficial alleles, and sexual selection. Evolution 54:1855–1861.
Whitlock, M., and S. P. Otto. 1999. The panda and the phage: compensatory mutation and the persistence of small populations. Trends. Ecol. Evol. 14:295–296.
Yoshimaru, H., and T. Mukai. 1985. Relationships between the polygenes affecting the rate of development and viability in *Drosophila melanogaster*. Jpn. J. Genet. 60:307–334.
Zar, J. H. 1999. Biostatistical analysis, 4th ed. Prentice Hall, Upper Saddle River, NJ.
Zeyl, C., M. Mizesko and, J. A. G. M. de Visser . 2001. Mutational meltdown in laboratory yeast populations. Evolution 55: 909–917.

Corresponding Editor: C. López-Fanjul