### Title

Mutation accumulation affects male virility in Drosophila selected for later reproduction.




### Journal

Physiological and Biochemical Zoology, 80(5)

### ISSN

1522-2152



### Authors

Borash, Daniel J
Rose, Michael R
Mueller, Laurence D



### Publication Date

2007-09-01

### DOI

10.1086/520127

# Mutation Accumulation Affects Male Virility in *Drosophila* Selected for Later Reproduction

**Daniel J. Borash***
**Michael R. Rose**†
**Laurence D. Mueller**‡
Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92697-2525

Accepted 3/6/2007; Electronically Published 7/13/2007

## ABSTRACT

An intensive study of longevity, female fecundity, and male reproductive behavior in *Drosophila melanogaster* was undertaken in order to establish whether late-life fitness characters in short-lived populations might be affected by the increase in deleterious alleles due to random genetic drift. We also sought to determine whether selection for late-life fertility could eliminate alleles that produce a decline in later fitness components in short-lived populations, as predicted by the mutation accumulation hypothesis for the evolution of aging. These experiments employed long-lived (O) populations, short-lived (B) populations, and hybrids made from crosses of independent lines from within the O and B populations. No detectable longevity differences were seen between hybrid B males and females and purebred B males and females. Reproduction in aged B purebred females was significantly less than in hybrid females at 3 wk of age only. A full diallel cross of the five replicate B lines showed a steady increase in hybrid male reproductive performance after the first week of adult life, relative to the parental lines. A full diallel cross of the five replicate O lines revealed no significant increase in hybrid O age-specific male reproductive success compared with the purebred O lines when assayed over the first 5 wk of adult life. The results on male reproductive behavior are consistent with the idea that relaxed age-specific selection in the B populations has been accompanied by an increase in deleterious, recessive traits that exhibit age-specific expression. Consequently, we conclude that a mutation accumulation process has been at least partly responsible for the age-specific decline in male B virility relative to that of the O populations.

* Present address: Patro Information K. K., Nakata Mac Toranomon Building, 1-1-10 Atago 3F, Minato-ku, Tokyo 105-0002, Japan; e-mail: drdaninjp@u01.gate01.com.
† E-mail: mrrose@uci.edu.
‡ Corresponding author; e-mail: ldmuelle@uci.edu.

*Physiological and Biochemical Zoology* 80(5):461–472. 2007.  1522-2152/2007/8005-0118$15.00
DOI: 10.1086/520127

## Introduction

The term "senescence" refers to a sustained age-specific decline in an organism's fitness-related characteristics beginning after the onset of reproduction. This decline in fitness-related traits is the result of the decreasing intensity of natural selection after the onset of reproduction (Hamilton 1966; Charlesworth 1980; Rose 1991). Consequently, the rate of senescence is predicted to depend on patterns of natural selection. In particular, if the start of reproduction is postponed, evolutionary theory predicts a postponement of senescence. This prediction has been empirically confirmed in populations of *Drosophila melanogaster* (e.g., Rose and Charlesworth 1980; Luckinbill et al. 1984; Rose 1984; Partridge and Fowler 1992; Leroi et al. 1994). For example, in all these studies, the mean longevity of populations selected for later-life fertility increased for both males and females relative to the control populations, which continued to be cultured using early reproduction.

Two genetic mechanisms have been proposed for the evolution of aging: antagonistic pleiotropy and mutation accumulation. With respect to the first of these, Medawar (1952) and Williams (1957) proposed that alleles benefiting early-life characters at the expense of later fitness will hold a selective advantage, as the force of selection declines with advancing age (see also Charlesworth 1980; Rose 1985). Several studies have suggested that antagonistic pleiotropy is an important factor affecting postponed senescence (e.g., Luckinbill et al. 1984; Rose 1984; Partridge and Fowler 1992).

Under the mutation accumulation hypothesis, senescence is a result of mutation-selection balance. Mutations with deleterious age-specific effects are hypothesized to be constantly generated, and selection becomes less effective at removing them as the age of onset of such allelic effects increases (Medawar 1952; Edney and Gill 1968; Graves and Rose 1989). Promislow and Tatar (1998) reviewed four possible experimental tests of the mutation accumulation theory of senescence. (1) The variance in fitness-related traits should increase with age. (2) Populations that have been selected for late-life fitness should show little response to traits affected by mutation accumulation when reverse selected for high early fitness. (3) Controlled mu-

tation, such as mutation accumulation lines or P-element mutation, should produce recognizable patterns of senescence. (4) Inbreeding depression should increase with age. We agree that methods 1, 2, and 4 are reasonable approaches to studying mutation accumulation. The third method will be contingent on the nature of the mutation and its patterns of pleiotropy. Since these are typically beyond the control of the experimental scientist, this method may have limited utility.

The first test has been applied to age-specific fecundity in *Drosophila* (Rose and Charlesworth 1980, 1981; Engström et al. 1989; Tatar et al. 1996), age-specific mortality (Hughes and Charlesworth 1994; Promislow et al. 1996), and male virility (Kosuda 1985; Hughes 1995). While increased genetic variance with age is a commonly reported finding, it is not universally seen in these studies. Moreover, Promislow and Tatar (1998) suggest that certain experimental and statistical techniques used in this research might have biased the results toward inferring increases in variance (Shaw et al. 1999).

The second test of mutation accumulation—reverse selection—to our knowledge has been used twice before (Service et al. 1988; Graves et al. 1992; Teotónio and Rose 2000). These studies provided evidence for both antagonistic pleiotropy and mutation accumulation. Method 4 has been used with *Drosophila* female fecundity (Mueller 1987) and age-specific survival (Charlesworth and Hughes 1996). Both of these studies found evidence of inbreeding depression at later ages and thus mutation accumulation. In this study, we will also examine inbreeding depression as a function of age.

A corollary of the mutation accumulation hypothesis is that in populations where the force of selection is reduced to 0 early in life, senescence should be accelerated. In such cases, alleles with effects confined to late ages, after the cessation of reproduction, would be free to drift to higher frequency unimpeded by natural selection at earlier ages. We propose testing this idea with a collection of five independent populations, called B's, that had been maintained in the laboratory for 416 generations at the start of these experiments (Rose 1984; Rose et al. 2004). The B's are cultured on a discrete 2-wk generation cycle. Effectively, any fitness-related trait expressed after about day 5 of adult life is neutral except to the extent that these traits depend on alleles having pleiotropic effects on early-life fitness. Consequently, we expect some deleterious late-acting alleles already present in the B populations to have increased in frequency over this period due to drift. However, it is unlikely that the identical deleterious alleles will have increased in all five replicate populations, due to the random nature of genetic drift. While it is possible that new mutations may have been created and raised to high frequency over this period, we think that possibility unlikely due to the large size of the populations ($N_e \sim 1{,}000$) and the relatively short period of time.

The critical test of inbreeding depression, and thus mutation accumulation, is to compare late-life fitness in the five replicate B populations with the late-life fitness of hybrids between replicate B populations. We expect the hybrids to show restored fitness because they should be heterozygous at many of the fitness-depressing loci, with directional dominance in favor of

| | | FEMALES | | | | |
|---|---|---|---|---|---|---|
| | | B1 | B2 | B3 | B4 | B5 |
| MALES | B1 | P | H | H | H | H |
| | B2 | H | P | H | H | H |
| | B3 | H | H | P | H | H |
| | B4 | H | H | H | P | H |
| | B5 | H | H | H | H | P |

Figure 1. Diallel cross of the five replicate B populations. Each shaded cell with a P indicates a purebred or parental population ($B_i$ female × $B_i$ male). Each unshaded cell with an H indicates a hybrid of two of the parental populations ($B_i$ female × $B_j$ male, $i \neq j$). Each $F_1$ set of populations was initiated by at least 400 individuals of each gender comprising the parents of the crossed populations. The diallel cross of the O populations was performed in the same manner.

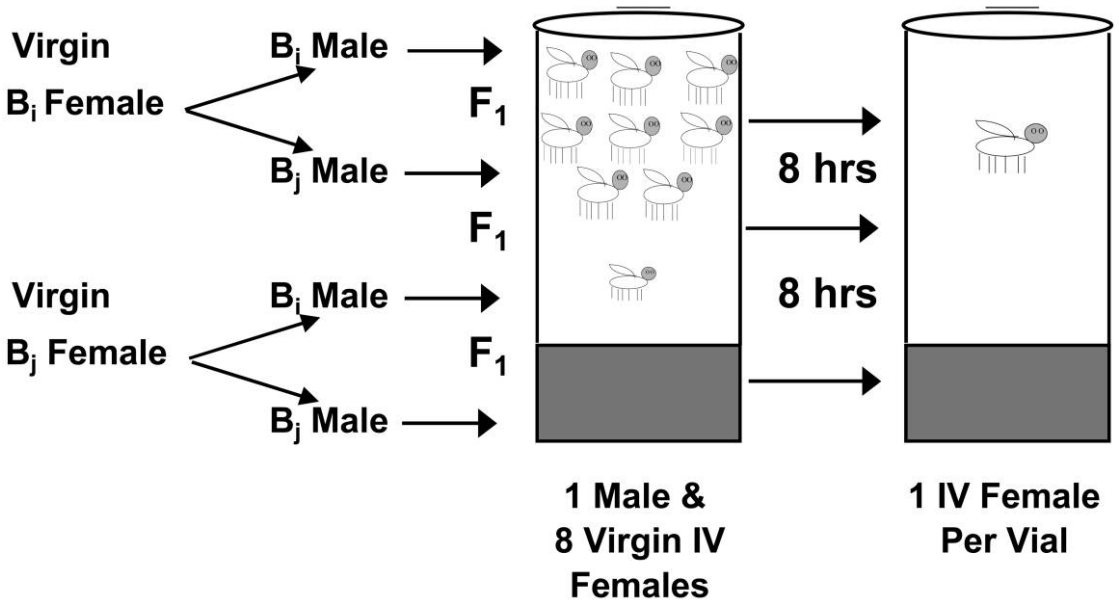


Figure 2. Design of the male virility assays. Virgin females from one of the B replicates were mated to males from the same replicate (purebred) and to males from another replicate (hybrid). The $F_1$ progeny from these matings were then assayed every week; one of the $F_1$ males was placed with eight virgin IV females for 8 h to mate. After those 8 h passed, the male was discarded, and each IV female was placed in her own vial, where she could lay eggs for 3–4 d. The presence of even a single larvae indicated that the male successfully inseminated that female. Assays of the O populations were carried out in the same manner.

more fit alleles (Wright 1977, pp. 26–27; Rose 1985). We also expect that the relative difference between hybrid and parental populations will increase with age. There are two reasons for this expectation. (i) Late-acting deleterious alleles may not increase in the B populations if they have pleiotropic effects on early-life fitness components. It seems reasonable that the chance of such pleiotropic effects would decrease with age. (ii) The equilibrium frequency of later-acting deleterious alleles in the source population is expected to be greater at later ages because the strength of selection against these alleles decreases with age (Charlesworth and Hughes 1996). A second important control will be the examination of inbreeding depression in populations selected for late reproduction. The O populations are cultured on discrete 10-wk generations. Hence, there continues to be strong selection for high fitness in late life. This should keep late-acting deleterious alleles at low frequencies, unlike the B populations.

Important fitness characteristics, such as (i) longevity, (ii) male virility, and (iii) female fecundity have been shown to display a general decline in *Drosophila* populations as they age (Rose 1984; Kosuda 1985; Service 1993). The differentiation of some *Drosophila* laboratory populations for longevity and fecundity make them a critical system with which to investigate whether mutation accumulation has had an effect on senescence (Rose et al. 2004). Here, we extend the study of such populations to the use of hybrid populations assayed for male virility.

## Material and Methods

In 1980, five early-cultured replicate populations ($B_{1–5}$) and five late-cultured replicate populations ($O_{1–5}$) were derived from a long-established laboratory population called IV (Rose 1984). The B populations have been maintained on discrete 2-wk generation cycles in the same manner as the IV populations: at low larval densities and with 14 d to develop and reproduce. The O populations were maintained in an identical manner to the B populations for the first 2 wk following egg collection; afterward, they were provided with fresh food three times a week for up to eight additional weeks. Before egg collection, they were placed in Plexiglas cages and provided with banana food plates with live yeast in order to stimulate oviposition. The O populations eventually had their adult phase extended to a 10-wk life cycle. At the start of these experiments, the B populations had undergone over 400 generations of selection, and the O populations had undergone over 100 generations of selection (Rose et al. 2004).

### *Experimental Design*

*All Experiments.* Before crosses, virgin males and females were collected from each population required for that particular assay. Eggs from the stock populations were collected at normal density (60–90 eggs/8-dram vial) and allowed to proceed through their development. At the first sign of pupal darkening,

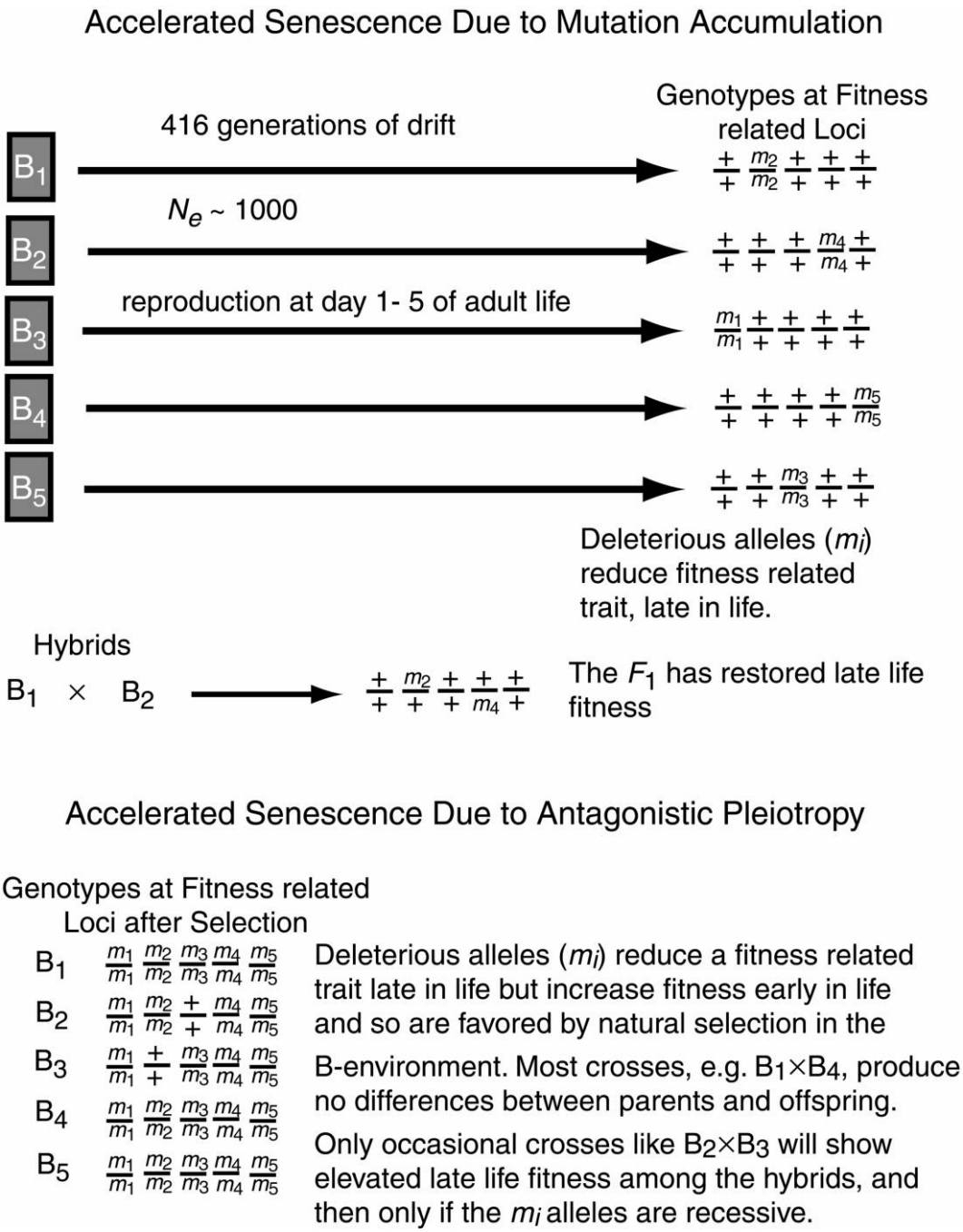


Figure 3. Expected results from hybrid crosses under two theories of late-life fitness. The expected results, if mutation accumulation was the most important determinant of late-life fitness, are shown on the top. Five independent populations are maintained on a 2-wk generation cycle for 416 generations at an effective population size of about 1,000. This culture regime removes the action of selection on traits expressed after about day 5 of adult life. We assume that at the start of the experiment, the populations were polymorphic for deleterious late-acting recessive alleles at multiple loci. In this case, we have shown five such alleles—$m_1$, $m_2$, $m_3$, $m_4$, and $m_5$—at five different loci. One of these deleterious alleles becomes fixed by genetic drift in each of the five populations. When late-life fitness is examined in the five B populations at the end of the experiment, they show a depression. When the hybrids are made from any two B populations, they show elevated late-life fitness because they are now heterozygous for the dominant wild-type allele at the relevant loci. If antagonistic pleiotropy is the major mechanism of evolution, then we expect most replicate populations to have fixed the same set of alleles due to the early beneficial effects on fitness (see bottom of figure). When hybrids are created, the decline in late-life fitness is usually unaffected because the genotype of the hybrids is the same as each parent.

checks for eclosing adults began. At 7-h intervals, emerging adults were removed from vials using $CO_2$ anesthesia, with females and males collected and sequestered separately. Single-sex vials were checked for contamination, and any "females only" vials containing one or more males were discarded. Once we had at least 400 flies of each gender, for each separate cross population, the flies were placed in their appropriate cages to mate and supplied banana food plates with liberal amounts of live yeast to stimulate female oviposition.

*Female Fecundity and Adult Longevity Assays.* A complete diallel cross using all 25 possible combinations of the B populations was performed with a minimum of 400 of each sex contributing to the progeny ($F_1$). Twenty hybrid and five purebred populations were created (Fig. 1). The $F_1$ eggs that resulted from these matings were collected and reared at 60–80 eggs/vial for 14 d.

Longevity assays on the 20 hybrid populations and five purebred populations employed four pairs of approximately 4–5-d-old adults placed into vials using $CO_2$ anesthesia. Twenty replicate vials were set up for each of the 20 hybrid B populations and five purebred B populations. Vials were checked daily for mortality, with transfers to fresh food occurring every other day. Dead flies were removed during these transfers. No longevity assays were performed on the O populations (see "Discussion").

Fecundity was measured on adults approximately 4–5 d old by placing one pair per vial on a yeasted charcoal-sucrose medium for 24 h, after which they were transferred back to banana food vials and the eggs were counted. This age is the typical reproductive age for the B populations and was designated "week 0." During the time between each assay, each pair was housed together in vials, with transfers to fresh banana food occurring every other day. At weekly intervals, up until what would have been 4 wk past the normal reproduction period for the B populations, the pair was again placed on charcoal medium and left to lay eggs for 24 h. During this experiment, dead females were removed and dead males were replaced with backup stock males that were the same age as the female in the vial. Each tested population began with 20 females. By the fourth week there were, on average, nine females left alive. No fecundity assays were performed on the O populations (see "Discussion").

*Male Mating Success Assays.* The complete diallel cross (5 × 5) was sectioned into separately executed blocks due to the size of the experiments (Fig. 1). For any block, the complete diallel crosses from the populations examined were assayed together (i.e., $B_i$ female × $B_j$ male, $B_j$ female × $B_i$ male, $B_i$ female × $B_i$ male, and $B_j$ female × $B_j$ male). All crosses $i = j$ were termed "purebred crosses" and were, in effect, the normal mating conditions for the respective B or O populations. All crosses where $i \neq j$ were termed "hybrid crosses." We did not attempt any between-treatment population crosses (B × O) for these experiments (but see Hutchinson and Rose 1991).

The $F_1$ generation was collected and housed in incubators (24L : 0D, 25°C) for 2 wk before being placed in cages as adults. The adults were maintained in these cages throughout the experiment and were given fresh food plates three times per week. Assays were performed once per week, with flies being directly sampled from those cages but not returned to the cages afterward.

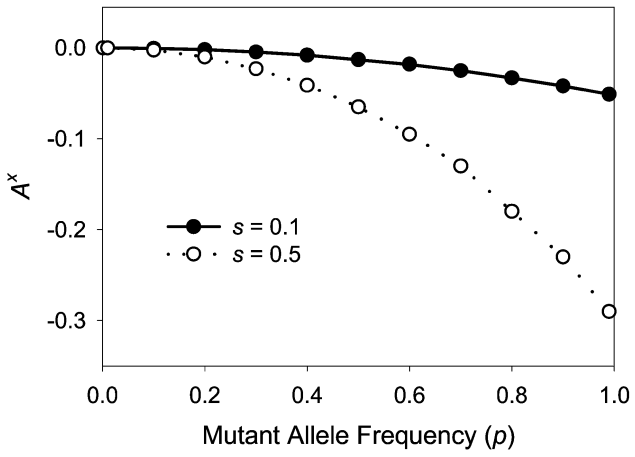


Figure 4. $A^x$ as a function of allele frequency. This curve was computed assuming that the frequency of the two deleterious alleles at the major polymorphism, $m_1$ and $m_2$, were equal to $p$ in each population. The alternative mutant allele at the minor polymorphism in each population was kept at a constant low frequency of 0.001. Fitness was computed as described in the text with $s = 0.1$ and 0.5. The value of $A^x$ was standardized at each allele frequency by the mean of the parental and hybrid populations.

Approximately 3–4 d before the onset of the experiment, a large cohort of virgin IV females was collected to be used as the female mating partners to measure male mating ability. These virgin females were also housed in vials and checked for male contamination at least three times after being housed in vials. A vial that had even one male, or displayed the presence of a single larva, was discarded before the experiment.

Approximately 20–25 males from each of the experimental populations cages were obtained, and from this each male was placed into a vial with eight virgin IV females and allowed to mate for 8 h. Eight hours was determined to be an adequate length of time to assay the males, as male fertility returns, for *Drosophila melanogaster*, after 2 to 3 h since the last mating, and males may often mate with two to three females before exhausting their supply of sperm and accessory fluid (Lefevre and Jonsson 1962; Service 1993). After the 8 h of mating, the vial was emptied onto a gas plate (using $CO_2$ anesthesia), the male was discarded, and each female was placed into a fresh vial to lay eggs (Fig. 2). If a vial was found to have a dead male or more than one male, that vial was discarded, and a replacement vial of one male from the same population and eight virgin IV females was substituted. At least 20 replicate mating vials for each population were used each of the 5 wk of the

Table 1: B-fecundity ANOVA

| Source | Sum of Squares | df | Mean Square | *F* | *p* |
|---|---|---|---|---|---|
| Between females: | | | | | |
| Population | 77.8 | 24 | 3.24 | 2.40 | .001 |
| Error | 269 | 199 | 1.35 | | |
| Within females: | | | | | |
| Age | 377 | 4 | 94.3 | 115 | <.001 |
| Age × population | 187 | 96 | 1.95 | 2.37 | <.001 |
| Error | 796 | .822 | | | |

Note. Repeated-measures ANOVA was used to test the effects of age and source population on fecundity. Only females with complete sets of measurements were used.

assay. The females were left to lay eggs for at least 4 d, then the females were removed, and the vials were carefully observed for the presence of larvae. The presence of even one larva in a vial served to indicate that the female had been inseminated by the male. Each male's mating ability was determined by the number of females he successfully inseminated over the 8-h mating period.

This procedure was employed once a week for 5 wk, which included the normal egg laying period for the B populations, in order to ascertain the age-specific decline in male mating ability. The experiments involving the B populations were performed from fall 1995 until summer 1996. The O populations were assayed in an identical manner but were assayed from spring 1997 until spring 1998. At the time of the experiments, the B populations had undergone approximately 500 generations of selection for early-life fertility, while the O populations had undergone approximately 100 generations of selection for delayed reproduction. In the course of these experiments 3,636 B males and 3,296 O males were assayed for virility.

### *Quantifying Mutation Accumulation*

In this study we have created conditions in the B populations for mutation accumulation. Such an accumulation will depend on several factors: (1) the elimination of selection in late life, (2) the finite population size, and (3) the existence of late-acting deleterious alleles for the life-history characters we examine. The first two factors are part of the experimental design; the third factor constitutes the biological hypothesis of mutation accumulation. The dynamic aspects of the process of mutation accumulation in the B populations are shown in Figure 3. This process presumes multiple loci affecting the trait of interest. It further assumes that some existing deleterious alleles will rise to high frequency in each population and others will not. The particular alleles that rise to high frequency in one population are also likely to be different from one population to the next.

We assume that most deleterious alleles that rise to high frequency by drift will be recessive or partially recessive. This conclusion follows from the simple population genetic considerations that suggest that in the ancestral population, recessive deleterious alleles will be at a much higher equilibrium frequency than dominant alleles. Consequently, in the simplest case, which involves fixation of the deleterious alleles, the parental populations will show a depression in these late-life characters that will be elevated in the $F_1$ hybrids (Fig. 3). If, on the other hand, fitness characters early and late in life are determined wholly by alleles with antagonistic effects (Fig. 3), we expect to see little difference between the hybrids and parents for their late-life fitness characters.

However, there is nothing about the design of this experiment that guarantees or assumes that these deleterious alleles will be fixed after the 416 generations of drift. For neutral alleles at an initial frequency of $p$ and with an effective population size of $N$, it will take on average $-4N(1-p)\ln(1-p)/p$ generations to fix the allele (assuming it is fixed, which will occur

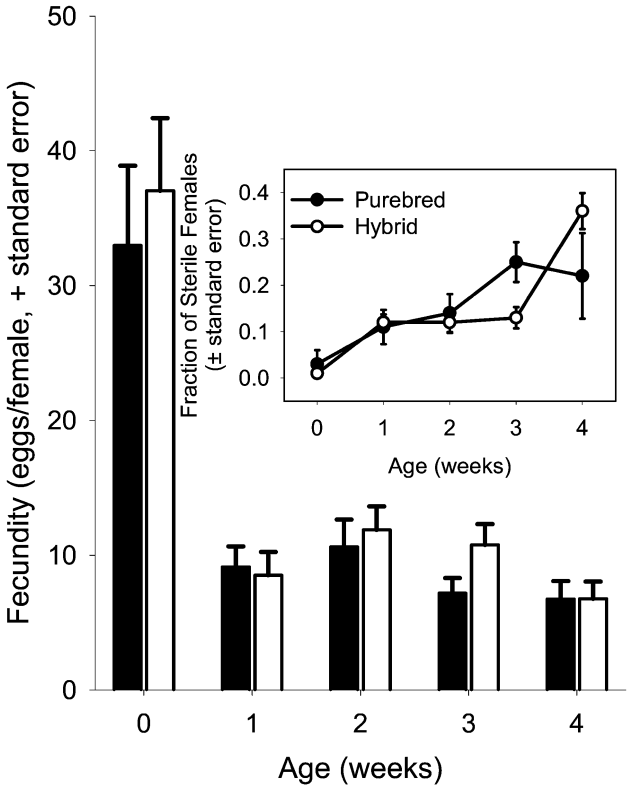


Figure 5. Age-specific decline in female fecundity for both hybrid and purebred B females reveals that until 3 wk after the normal reproductive period for the B populations, there is no indication of mutation accumulation affecting these populations. After 3 wk, the hybrids display greater fecundity than the purebred B populations. This difference disappears by the fourth week due to small sample size, but it indicates that the hybrid collapse in female fecundity occurs later than the purebred collapse. The inset graph denotes the increase in B purebred female sterility by week 3, with the B hybrids reaching the same level of female sterility by week 4.

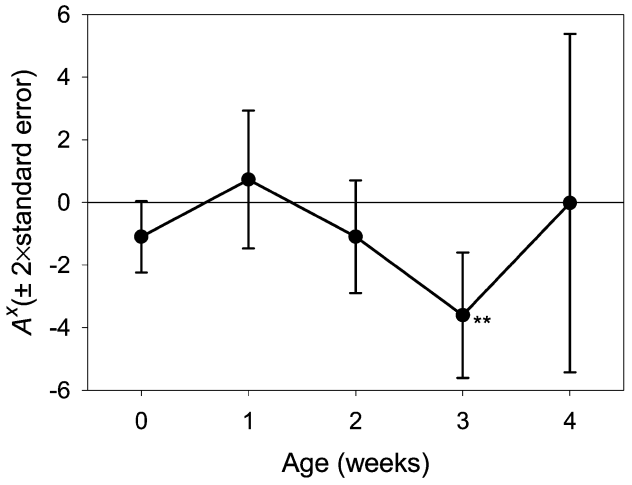


Figure 6. $A^x$ statistic for female fecundity in the B populations and their hybrids. Significant departures from additivity, as determined from the ANOVA, are indicated by asterisks (two asterisks = $p < 0.01$).

with probability $p$; Ewens 1979, p. 77). In the B populations, $N_e \cong 1{,}000$, and if $p = 0.05$, the average time to fixation would be 3,898 generations.

While 416 generations is not sufficiently long for most initially rare neutral alleles to be fixed, there may be some that have risen to sufficiently high frequencies that late-life fitness will begin to suffer. For instance, using the stationary distribution of neutral alleles, we can calculate the chance of finding neutral alleles in certain frequency ranges (Crow and Kimura 1970, p. 383). In the B populations, 4%–9% of the neutral alleles are expected to be at a frequency of 0.4 or greater (assuming $N_e = 1{,}000$ and the initial frequencies are between 0.01 and 0.1). At frequencies above 0.4, there will be sufficient numbers of deleterious homozygotes so that late-life fitness characters will begin to be affected. So we need to address the problem of what experimental results to expect if deleterious alleles have risen to high frequency but are still polymorphic at the time these experiments were conducted.

Our basic indicator of mutation accumulation (Fig. 3) is some degree of hybrid vigor in the $F_1$ population. Suppose at age $x$ we measure a phenotype (e.g., either fecundity or virility) in the $i$th and $j$th parental population to get $P_i^x$ and $P_j^x$, respectively. In the hybrid of these two populations, the phenotype measure is $P_{ij}^x$ (and $P_{ji}^x$ in the reciprocal cross). The mean parental phenotype minus the mean hybrid is just

$$\alpha_{ij}^x = \frac{1}{2}(P_i^x + P_j^x) - \frac{1}{2}(P_{ij}^x + P_{ji}^x). \tag{1}$$

If there is some degree of heterosis, then this statistic will be negative. If the alleles are completely additive, then equation (1) will be 0. For the measurements of fecundity, the same set of parental phenotypes was used with all hybrid measurements. If we then sum up the $\alpha_{ij}^x$ statistic over all parental and hybrid populations in our experimental design (Fig. 1), we have a measure of mutation accumulation

$$A^x = \frac{1}{2}\left(4\sum_{i=1}^{5} P_i^x - \sum_{i=1}^{5}\sum_{j=1, i\neq j}^{5} P_{ij}^x\right). \tag{2}$$

The variance of this statistic is

$$\mathrm{Var}(A^x) = \frac{1}{4}\left[16\sum_i \mathrm{Var}(P_i^x) + \sum_i \sum_{i\neq j} \mathrm{Var}(P_{ij}^x)\right]. \tag{3}$$

For the virility experiment, the parental phenotypes are measured multiple times. In these cases, the statistic $A^x$ is still the sum of the $\alpha_{ij}^x$ for each reciprocal cross, but the variance calculations differ. For female fecundity, $P_i^*$ and $P_{ij}^*$ are estimated from the mean of nine to 20 females. The variance of these means is calculated in the standard way. For male virility, $P_i^*$ and $P_{ij}^*$ are the fraction of 160 females successfully mated. We use binomial variance estimates for virility, that is, $P_i^*(1 - P_i^*)/160$.

An important component of this study is to compare the relative levels of dominance at different ages using the statistic in equation (2). A problem encountered in such comparisons

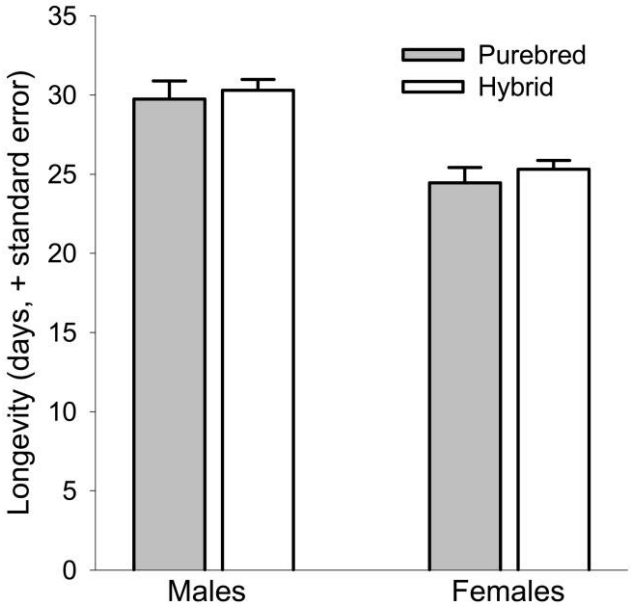


Figure 7. Mean longevity of all five purebred populations and all 20 hybrid populations for both sexes are shown with the standard errors around the means. Standard errors were computed by treating each population mean as a single observation. While males lived significantly longer than females, hybrid superiority within the same gender adult was not detected.

Table 2: B-virility ANOVA

| Source | Sum of Squares | df | Mean Square | *F* | *p* |
|---|---|---|---|---|---|
| Block | .680 | 4 | .170 | 2.15 | .072 |
| Age | 93.6 | 4 | 23.4 | 296 | <.001 |
| Population | 15.0 | 24 | .624 | 7.89 | <.001 |
| Population × age | 8.60 | 96 | .090 | 1.13 | .18 |
| Error | 277 | 3,507 | .079 | | |

Note. Population refers to all 25 hybrid and parental populations.

is that the mean of the phenotype measure may decrease substantially with increasing age even as the relative levels of heterosis increase. Because the phenotypes we are measuring are fitness components, the important contrast between hybrid and parental populations is simply the relative differences in fecundity and virility, not their absolute values. Accordingly, we have standardized each $A^x$ statistic by dividing it by the mean of all 25 parental and hybrid populations. Because this mean is based on a very large number of observations, we treat it as a constant for the purposes of computing the variance.

*Dynamics of the* $A^x$ *Statistic.* Consider two loci with two alleles. At the first locus, the fitnesses of the three genotypes $AA$, $Am_1$, and $m_1m_1$ are 1, 1, and $1 - s$, respectively. At the second locus, the fitnesses of the three genotypes $BB$, $Bm_2$, and $m_2m_2$ are 1, 1, and $1 - s$, respectively. We can further assume that there are two populations. In the first population there is a polymorphism at the $A$ locus with a substantial number of $m_1$ alleles. This is called the major polymorphism. The second locus is referred to as a minor polymorphism because the $m_2$ allele is present only at very low frequencies. In the second population, the first locus is the minor polymorphism, and the second locus is the major polymorphism. We assume that if the $A$ statistic is calculated for these two populations, then the sample of flies from each parental population will contain all three genotypes at each locus. Similarly, the hybrids formed by crossing these two populations will not all be heterozygotes at both loci. If the genotypes are assumed to be in Hardy-Weinberg and linkage equilibrium, we can compute the value of $A^x$ as a function of the frequency of the $m_1$ and $m_2$ alleles (Fig. 4). The value of the $A^x$ statistic shows a steady decline with increasing frequency of the recessive deleterious alleles at the major polymorphic locus. The ability to detect significant departures from 0 will clearly depend on the frequency of the deleterious allele and the magnitude of their effects. Because $A^x$ decreases only slowly at low allele frequencies, it is probable that this technique will be unable to detect mutation accumulation unless there has been some substantial allele frequency differentiation.

*Statistical Analysis.* A *t*-test was performed on the longevity assay, using male and female data separately. The effects of age and population on fecundity and virility were evaluated with an ANOVA. Because the mean and variance of female fecundity decline with age, the fecundity data were transformed with the natural logarithm of 1 + fecundity. Fecundity was measured on the same females multiple times. Thus, a repeated-measures ANOVA design was used with the fecundity data. The virility data used independent samples of males, and this does not require a repeated-measures design. However, there were block effects, because groups of populations were measured at different time periods. The raw data from the virility experiment were expressed as a fraction the eight females fertilized. These data were subject to an arcsin-square root transformation.

The ANOVA on fecundity and virility can also be used to statistically test for departures from additivity. For each age, we examined the null hypothesis that four times the sum of the five parental phenotypes (either fecundity or virility) equals the sum of all 20 hybrid phenotypes. This hypothesis is related to our estimates of $A^x$, because $A^x = 0$ with complete additivity and $A^x$ will be less than 0 with heterosis. Twice the standard error of the $A^x$ statistic can be used as an approximate 95% confidence interval.

## Results

### *Female Fecundity*

ANOVA revealed significant effects of age and source population on female fecundity (Table 1). As expected, female fe-

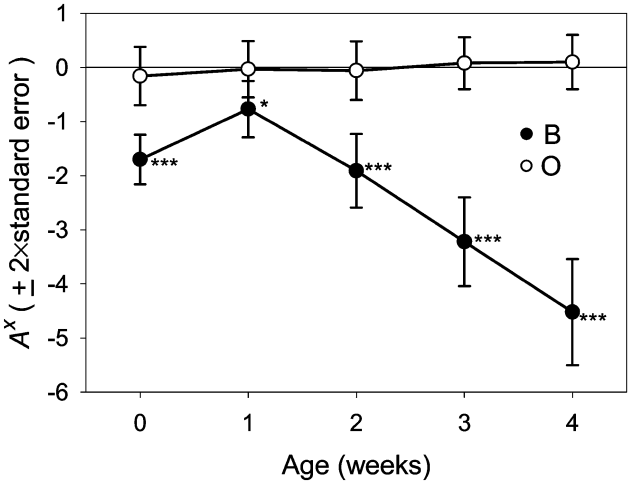


Figure 8. Hybrid vigor as revealed through the $A^x$ value. For each week, the summed means of both purebred population had subtracted from it the weighted averaged parental population's mean male mating success. The parental populations were the two replicate populations comprising the reciprocal cross to create the hybrids. While the O populations showed no heterosis, the B populations show a high degree of heterosis starting with their normal reproductive period (week 0), which decreases for week 1, and begins a steady rise throughout the remainder of the experiment. Significant departures from additivity, as determined from the ANOVA, are indicated by asterisks (one asterisk = $p < 0.05$, three asterisks = $p < 0.001$).

Table 3: O-virility ANOVA

| Source | Sum of Squares | df | Mean Square | $F$ | $p$ |
|---|---|---|---|---|---|
| Block | 10.5 | 4 | 2.63 | 36.5 | <.001 |
| Age | 3.225 | 4 | .806 | 11.2 | <.001 |
| Population | .400 | 24 | .017 | .231 | 1.00 |
| Population × age | 7.22 | 96 | .075 | 1.04 | .368 |
| Error | 228 | 3,167 | .072 | | |

Note. Population refers to all 25 hybrid and parental populations.

cundity declines with age (Fig. 5). However, the ANOVA and our parametric estimates of $A^x$ suggest no significant differences between purebred B female fecundity and hybrid B female fecundity at 0, 1, 2, and 4 wk (Fig. 6). Hybrid B females showed a significantly higher fecundity than purebred B females at week 3 (Fig. 6). The only $A^x$ statistic that is significantly less than 0 is also at 3 wk (Fig. 6). While this may have be indicative of some hybrid vigor, the small sample sizes and high variance during the fourth week preclude any solid conclusions.

### *Longevity*

Overall, males lived significantly longer than females, regardless of whether they were from hybrid or purebred populations. The *t*-test revealed no statistical differences between hybrid B males (30.3 ± 0.68) and purebred B males (29.74 ± 1.14) for adult longevity. Purebred B females (25.15 ± 0.98) were not significantly shorter lived than the hybrid B females (25.31 ± 0.53; Fig. 7).

### *Male Mating Success*

The ANOVA of B-male virility shows significant effects of age and population (Table 2). In general, males show declining virility with age (Fig. 9). There are significant departures from additivity at all ages in the B populations (Fig. 8). However, the magnitude of this effect increases dramatically with age. A regression of the $A^x$ values from ages 1–4 wk produces a significant negative slope in the B populations ($p = 0.001$).

While the ANOVA detected a significant effect of age on O-male virility (Table 3), there is no obvious decline in virility with increasing age (Fig. 9). None of the tests of deviation from additivity were significant in the O populations, and the value of $A^x$ typically was very close to 0 (Fig. 8). The significant block effect indicates that the relative magnitude of virility changes in different temporal experiments. However, by running the parental populations of all hybrids in each block, we have not artificially inflated the virility of parental or hybrid lines over the whole experiment.

## Discussion

### *Experimental Evidence*

Hybrids showed no detectable superiority over their purebred counterparts, of either sex, when their longevities were assayed. The deleterious alleles that affect late-life survival, in the B populations, may have effects that are too small to be detected by the methods employed by this experiment, or the directional dominance assumption might not be correct for all survival characteristics. A recent study of much larger cohorts also found no difference in late-life mortality between uncrossed and hybrid B populations (Rose et al. 2002).

Only the week 3 fecundity value of $A^x$ is significantly less than 0. Under our hypothesis of mutation accumulation we would have expected the week 4 value to be less than 0 also. Experimental design limitations made it difficult to estimate the value of $A^x$ in week 4. The sample size dwindles as females die. The experiment began with 500 females, and by week 4 only 227 remained alive. The B3 population had only a single surviving female in week 4.

Male mating behavior is not an arbitrarily chosen character. Markow and colleagues have studied multiple mating in males extensively in both wild and laboratory populations of several

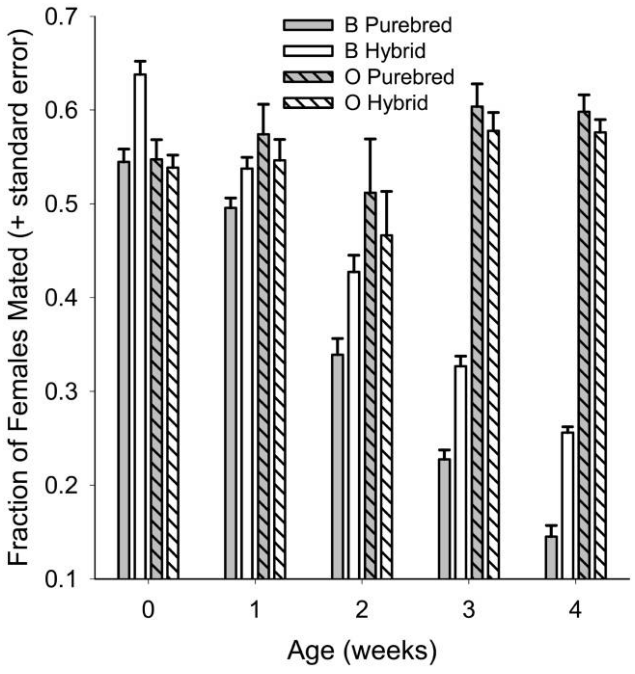


Figure 9. B and O male mating success in depictions of the age-specific decline in males for the B and O hybrid and purebred populations. Week 0 shows hybrid superiority in the B populations, which disappears during the first and second weeks of the assays (1 and 2 wk postnormal reproduction time for the B populations). At weeks 3 and 4, the hybrid superiority for male mating success reappears.

*Drosophila* species (Markow et al. 1978; Gromko and Markow 1993; Pitnick and Markow 1994). Others have found that the trait is influenced by male age and competitive access to females (Service 1993). After week 1, $A^x$ clearly declines with age in the B flies. In direct contrast to this pattern are the O males, which show no evidence of hybrid vigor over the 5-wk period. In fact the O males show no decline in the fraction of females mated over the first 5 wk of adult life. This is expected since reproduction in the O populations requires that males mate at least until the fourth week of adult life (A. Joshi and L. D. Mueller, unpublished data).

The surprising result from these data is the hybrid vigor expressed at time 0, the normal age of reproduction in the B populations. We think this is unlikely to be due to inbreeding depression. If these populations have suffered increased homozygosity due to inbreeding, other traits should have been affected, and that was not observed. The likelihood of a bottleneck or extended period of small population size giving rise to inbreeding depression would also be related to the effective population size. We have made an extensive examination of the effective population size in both the B and the O populations including 3 yr of complete census counts of all 10 populations. The effective size in these populations is generally around 950–1,000 (L. D. Mueller, unpublished data). If anything, the effective size in the O populations has been less than in the B's. Because there is no evidence of any type of inbreeding depression in the O virility, we think it unlikely that demographic factors have caused the reduction in early B-male virility, compared with their hybrids.

The action of natural selection, however, may account for these patterns. Joshi et al. (1999) demonstrated that under conditions similar to the normal B-culture regime, few males mate more than once. Thus, in this environment, the virility phenotype measured here may be only weakly related to fitness. This would mean that deleterious alleles at all ages might rise in frequency. However, the pattern of age specificity in hybrid vigor in the B populations is consistent with the prediction of the mutation accumulation theory of higher levels of deleterious genetic variation at older ages (Charlesworth and Hughes 1996). In fact, it is probably a relatively weak correlation to fitness that has made virility the most likely phenotype to show hybrid vigor in these experiments, compared with longevity and female fecundity.

### *Mutation Accumulation in the B's or Mutation Elimination in the O's?*

The B and O populations used here have been used extensively in the past to study the genetic and physiological properties of aging (Rose et al. 2004). Some of the most important observations made in the past are that in response to culturing only long-lived adults, the O populations show increased longevity, reduced early fecundity, and an increase in a number of correlated physiological traits relative to the B populations. Some of these observations have been used as evidence in support of the antagonistic pleiotropy theory of aging. Promislow and Tatar (1998) have suggested that these conclusions may be erroneous. Their argument is that before the start of selection in the B and O populations, the base stock had undergone about 120 generations of selection in a B-type laboratory environment. Promislow and Tatar suggest that this was sufficient time for mutations to accumulate and thus cause up to a 50% decline in late-life fitness. This effect, they reason, is sufficiently powerful that it is well known that wild-caught flies are more "robust" than lab stocks. This last conclusion was supported with a citation of Dobzhansky et al. (1964), who in fact demonstrated that over time, laboratory populations exhibit an increase in the demographic fitness measure $r_m$.

Nevertheless, perhaps the B hybrids simply show a decline in virility similar to their feeble ancestors, and the O populations are just like wild-type flies having had the opportunity to eliminate the deleterious alleles that accumulated in the base populations. Without having samples of frozen flies from the original wild population, it is somewhat difficult to determine this unambiguously. However, other data exist that do shed light on the problem suggested by Promislow and Tatar.

Kosuda (1985) measured male virility in chromosome extraction lines and in heterozygous lines that were wild-type flies from his initial collection in Katsunuma, Yamanashi, Japan. The essential features of Kosuda's study were (i) he examined females for presence of sperm in his virility assays, (ii) single males were allowed to mate for 24 h, (iii) males were measured at 3 d of age and 28 d of age, and (iv) males had access to 12 virgin females.

Our study differed from Kosuda's in several ways. We measured virility at 4–5 and 25–26 d of age (week 3). However, we looked for the presence of viable progeny that would reduce our estimates of virility compared with Kosuda's. Kosuda presented males with 50% more females than we did, which might lower his estimate of virility. Compensating for this was the fact that Kosuda's males had three times longer to complete mating. So while our experiments are not exactly comparable, there is no striking difference that would cause one to think the virility measures would differ systematically between our two studies. In Kosuda's study, his wild-caught young males inseminated 77% of the females, while our young B-hybrid males fertilized 64% of the females. Kosuda's old males inseminated only 38% of all females, while our old B hybrids fertilized 33%.

Overall, these numbers are very similar. However, the most important aspect of these numbers is the relative reduction from young to old ages. In Kosuda's study, average virility of the wild males declined by 49%, while in our study, the B-hybrid males showed a decline of 52%. These results suggest that the declining performance of B-hybrid males is not unlike that of wild-type males. By contrast, the O males show no decline in virility over

a similar time period. In fact, both hybrid and purebred O males show an increase in virility over this period of time. Thus, we think that the virility of B hybrids is probably close to the level of virility in natural populations of *Drosophila melanogaster*, the virility of the purebred B populations reflecting the consequence of deleterious alleles accumulating over that last 400 generations, while O virility shows the beneficial effects of experimental evolution in the laboratory purging naturally occurring deleterious alleles and possibly increasing the frequency of beneficial alleles present in the ancestral populations at low frequency. These results do not support Promislow and Tatar's idea that O flies are similar to wild *Drosophila*.

### *Heterogeneity of Trait Results*

The results of this study vary quite widely between each fitness trait studied. This is not unexpected. While it is true that the level of drift is the same for genetic variation at each trait, other factors are likely to vary. For instance, late traits may not be strictly neutral, even under the conditions of this experiment, if there are any correlated negative effects on early fitness. This type of pleiotropy would have two consequences. (i) The equilibrium frequency of mutants with deleterious effects at early ages would be expected to be smaller due to their early effects on fitness, all other things being equal. (ii) Drift would be less effective at increasing these types of variants because they would still retain some negative fitness effects even in the demographic regime of the B populations. Both of these effects would reduce the chance of increasing frequencies for late-acting deleterious alleles in this experiment. There may also be a large difference in these types of correlations between different traits.

In the B-culture regime, it appears that most males only mate a single time before egg laying; very few males mate more than once (Joshi et al. 1999). Thus, even if there are correlations between early- and late-life male virility, their effects on fitness may be small. This suggests that male virility should respond more rapidly to drift. The strong decay in value of the $A^x$ statistic for virility at late age and the lack of any hybrid vigor in the O populations suggests that the B-culture regime has undergone mutation accumulation for age-specific virility. These results support the notion that at least in *Drosophila*, mutation accumulation may be an important cause of aging.

## Acknowledgments

We would like to thank those who assisted in the experiments, in particular G. Elias, D. Nguyen, H. Tran, and S. L. Nguyen. Thanks to Y. T. Morimoto and G. T. Ho for their expert assistance with the endless virgin collections. Special thanks to Amitabh Joshi, for initial motivation. We would also like to thank A. D. Long, M. D. Drapeau, T. J. Bradley, and L. Nunney for comments on the manuscript.

## Literature Cited

Charlesworth B. 1980. Evolution in Age-Structured Populations. Cambridge University Press, Cambridge.

Charlesworth B. and K.A. Hughes. 1996. Age-specific inbreeding depression and components of genetic variance in relation to the evolution of senescence. Proc Natl Acad Sci USA 93:6140–6145.

Crow J.F. and M. Kimura. 1970. An Introduction to Population Genetics Theory. Harper & Row, New York.

Dobzhansky T., R.C. Lewontin, and O. Pavlovsky. 1964. The capacity for increase in chromosomally polymorphic and monomorphic populations of *Drosophila pseudoobscura*. Heredity 19:597–614.

Edney E.B. and R.W. Gill. 1968. Evolution of senescence and specific longevity. Nature 220:281–282.

Engström G., L.E. Liljedahl, M. Rasmuson, and T. Björklund. 1989. Expression of genetic and environmental variation during aging. I. Estimation of variance components for number of adult offspring in *Drosophila melanogaster*. Theor Appl Genet 77:119–122.

Ewens W.J. 1979. Mathematical Population Genetics. Springer, Berlin.

Graves J.L. and M.R. Rose. 1989. Population genetics of senescence in *Drosophila*. Genetics 8:45–55.

Graves J.L., E.C. Toolson, C. Jeong, L.N. Vu, and M.R. Rose. 1992. Desiccation, flight, glycogen, and postponed senescence in *Drosophila melanogaster*. Physiol Zool 65:268–286.

Gromko M.H. and T.A. Markow. 1993. Courtship and remating in field populations of *Drosophila*. Anim Behav 45:253–263.

Hamilton W.D. 1966. The moulding of senescence by natural selection. J Theor Biol 12:12–45.

Hughes K.A. 1995. The evolutionary genetics of male life-history characters in *Drosophila melanogaster*. Evolution 49: 521–537.

Hughes K.A. and B. Charlesworth. 1994. A genetic analysis of senescence in *Drosophila*. Nature 367:64–66.

Hutchinson E.W. and M.R. Rose. 1991. Quantitative genetics of postponed aging in *Drosophila melanogaster*. I. Analysis of outbred populations. Genetics 127:719–727.

Joshi A., M.H. Do, and L.D. Mueller. 1999. Poisson distribution of male mating success in laboratory populations of *Drosophila melanogaster*. Genet Res 73:239–249.

Kosuda K. 1985. The aging effect on male mating activity in *Drosophila melanogaster*. Behav Genet 15:297–303.

Lefevre G., Jr., and U.B. Jonsson. 1962. Sperm transfer, storage, displacement, and utilization in *Drosophila melanogaster*. Genetics 47:1719–1736.

Leroi A.M., A.K. Chippindale, and M.R. Rose. 1994. Long-term evolution of a genetic life-history trade-off in *Drosophila melanogaster*. I. The role of genotype-by-environment interaction. Evolution 48:1244–1258.

Luckinbill L.S., R. Arking, M.J. Clare, W.C. Cirocco, and S.A.

Buck. 1984. Selection for delayed senescence in *Drosophila melanogaster*. Evolution 38:996–1003.

Markow T.A., M. Quaid, and S. Kerr. 1978. Male mating experience and competitive courtship success in *Drosophila melanogaster*. Nature 276:821–822.

Medawar P.B. 1952. An Unsolved Problem of Biology. Lewis, London.

Mueller L.D. 1987. Evolution of accelerated senescence in laboratory populations of *Drosophila*. Proc Natl Acad Sci USA 84:1974–1977.

Partridge L. and K. Fowler. 1992. Direct and correlated responses to selection on age at reproduction in *Drosophila melanogaster*. Evolution 46:76–91.

Pitnick S. and T.A. Markow. 1994. Male gametic strategies: sperm size, testes size, and the allocation of ejaculate among successive mates by the sperm-limited fly *Drosophila pachea* and its relatives. Am Nat 143:785–820.

Promislow D.E.L. and M. Tatar. 1998. Mutation and senescence: where genetics and demography meet. Genetica 102/103: 299–314.

Promislow D.E.L., M. Tatar, A. Khazaeli, and J.W. Curtsinger. 1996. Age-specific patterns of genetic variance in *Drosophila melanogaster*. I. Mortality. Genetics 143:839–848.

Rose M.R. 1984. Laboratory evolution of postponed senescence in *Drosophila melanogaster*. Evolution 38:1004–1010.

———. 1985. Life history evolution with antagonistic pleiotropy and overlapping generations. Theor Popul Biol 28:342–358.

———. 1991. Evolutionary Biology of Aging. Oxford University Press, New York.

Rose M.R. and B. Charlesworth. 1980. A test of evolutionary theories of senescence. Nature 287:141–142.

———. 1981. Genetics of life-history evolution in *Drosophila melanogaster*. I. Sib analysis of adult females. Genetics 97: 173–186.

Rose M.R., M.D. Drapeau, P.G. Yazdi, K.H. Shah, D.B. Moise, R.R. Thakar, C.L. Rauser, and L.D. Mueller. 2002. The evolution of late-life mortality in *Drosophila melanogaster*. Evolution 56:1982–1991.

Rose M.R., H.B. Passananti, and M. Matos, eds. 2004. Methuselah Flies: A Case Study in the Evolution of Aging. World Scientific, Singapore.

Service P.M. 1993. Laboratory evolution of longevity and reproductive fitness components in male fruit flies: mating ability. Evolution 47:387–399.

Service P.M., E.W. Hutchinson, and M.R. Rose. 1988. Multiple genetic mechanisms for the evolution of senescence in *Drosophila melanogaster*. Evolution 39:943–945.

Shaw F., D.E.L. Promislow, M. Tatar, K.A. Hughes, and C.J. Geyer. 1999. Toward reconciling inferences concerning genetic variation in senescence in *Drosophila melanogaster*. Genetics 152:553–566.

Tatar M., D.E.L. Promislow, A. Khazaeli, and J. Curtsinger. 1996. Age-specific patterns of genetic variance in *Drosophila melanogaster*. II. Fecundity and its correlation with age-specific mortality. Genetics 143:849–858.

Teotónio H. and M.R. Rose. 2000. Variation in the reversibility of evolution. Nature 408:463–466.

Williams G.C. 1957. Pleiotropy, natural selection, and the evolution of senescence. Evolution 11:398–411.

Wright S. 1977. Evolution and the Genetics of Populations. Vol. 3. Experimental Results and Evolutionary Deductions. University of Chicago Press, Chicago.