

# Spontaneous deleterious mutation in *Arabidopsis thaliana*

STEWART T. SCHULTZ\*<sup>†</sup>, MICHAEL LYNCH<sup>‡</sup>, AND JOHN H. WILLIS<sup>‡</sup>

\*Department of Biology, University of Miami, Coral Gables, FL 33124; and <sup>‡</sup>Department of Biology, University of Oregon, Eugene, OR 97403

Communicated by Margaret G. Kidwell, University of Arizona, Tucson, AZ, August 5, 1999 (received for review September 21, 1998)

**ABSTRACT** The frequency and selective impact of deleterious mutations are fundamental parameters in evolutionary theory, yet they have not been directly measured in a plant species. To estimate these quantities, we allowed spontaneous mutations to accumulate for 10 generations in 1,000 inbred lines of the annual, self-fertilizing plant *Arabidopsis thaliana* and assayed fitness differences between generations 0 and 10 in a common garden. Germination rate, fruit set, and number of seeds per fruit each declined by less than 1% per generation in the mutation lines, and total fitness declined by 0.9% per generation. Among-line variances increased in the mutation lines for all traits. Application of an equal-effects model suggests a downwardly biased genomic deleterious mutation rate of 0.1 and an upwardly biased effect of individual mutations on total fitness of 20%. This genomic deleterious mutation rate is consistent with estimates of nucleotide substitution rates in flowering plants, the genome size of *Arabidopsis*, and the equilibrium inbreeding depression observed in this highly selfing plant species.

The rate of spontaneous deleterious mutation plays a central role in the evolutionary genetics of sex and recombination (1, 2), mating systems and mate choice (1, 3–7), mutation accumulation and extinction in small populations (8–12), neutral molecular variation and its maintenance in natural populations (13), and life-history variation in natural populations (14); for a review see ref. 15. Despite its theoretical importance, the total number of new deleterious mutations per zygote per generation ( $U$ ) is unknown for all but a few species.

The genomic deleterious mutation rate most commonly is estimated by using mutation accumulation experiments, in which spontaneous mutations are allowed to accumulate in replicate populations in the relative absence of natural selection. In such studies  $U$  has been estimated as 0.02–0.6 for egg-adult viability in *Drosophila* (16–21),  $2 \times 10^{-4}$  for total fitness in *Escherichia coli* (22), 0.5 for life-history traits in *Daphnia pulex* (14), and 0.003–0.060 for various life-history traits in *Caenorhabditis elegans* (23, 24).

Population-genetic theory of the balance between mutation and selection permits indirect inference of the genomic deleterious mutation rate from measures of inbreeding depression in natural populations combined with the variance among families in the relative fitness of inbred individuals (25–27). These methods suggest that  $U$  is roughly one for egg to adult viability in *Drosophila* (3, 11), from 0.2 to 1.7 for a variety of fitness traits in highly selfing plants (26, 28, 29), 0.87 for clutch size in *D. pulex* (27), and about 0.04 to lethals in humans (25, 30, 31). However, these indirect estimates require assumptions about the average degree of dominance, absence of overdominance, and mutation-selection equilibrium that may not be valid.

No direct measurements of genomic deleterious mutation rate have yet been published for a plant species, despite the

central importance of this parameter in many recent theoretical models describing the evolution of plant mating systems, especially the joint evolution of inbreeding depression and the rate of self fertilization (see refs. 3–7, 32). We performed such a mutation-accumulation experiment in a plant species, the annual, self-fertilizing *Arabidopsis thaliana*. Our major objectives were to provide conservative (minimum) estimates of (i) the total mutational decay in major fitness components, and (ii) the total diploid deleterious genomic mutation rate,  $U$ .

## MATERIALS AND METHODS

**Generation of Mutation Accumulation Lines.** The mutation accumulation experiment was started with a single *A. thaliana* individual from the ecotype *Landsberg erecta*. To minimize the possibility of heterozygosity at the start of the experiment, a progenitor line was derived from this plant by self-fertilization and single-seed descent for six generations. Four hundred selfed seed from this single generation-6 plant were divided into two blocks. Block 1 was placed in storage (dark and dry and at room temperature) to serve as a control in the fitness assay described below, and the 200 seeds of block 2 were sown. Five seeds from each of the resulting 200 adults then were used to initiate generation zero of the mutation-accumulation experiment. These 1,000 lines were maintained independently by passive self-fertilization and single-seed descent for 10 generations, during which time spontaneous mutations were expected to accumulate. *Arabidopsis* naturally self-fertilizes its seed, but is probably capable of a very low frequency of outcrossing when plants are grown in close proximity. To minimize the chances of “migration” among the 1,000 replicate lines during mutation-accumulation caused by accidental outcrossing or seed dispersal, plants were raised individually within either clear glass tubes (for generations 0–5) or clear plastic tubes (for generations 5–10).

At the start of each generation, four seeds were sown per line, and one adult was chosen at random from each line to provide seeds for the next generation. The small effective population sizes were necessary to insure that mutations with mild effects on fitness were nearly neutral and thus accumulated close to their rate of occurrence. Mutations of large effect were not expected to accumulate, because of selection within lines and complete loss of lines.

Within-line selection may be effective on major mutations reducing preadult survivorship. Because four seeds were chosen from one random individual per line, the effective population size for mutations affecting germination was four. Because one adult individual was chosen at random as progenitor of the next generation, the effective population size for postgermination mutations was one. Thus mutations with selection coefficients less than 0.5 to 0.125 were effectively neutral and accumulated in our mutation lines.

Selection also may have occurred among the large number of gametes produced per individual, thus preventing some deleterious mutations from accumulating. However, gametic

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at [www.pnas.org](http://www.pnas.org).

<sup>†</sup>To whom reprint requests should be addressed at: Department of Biology, University of Miami, 215 Cox Science, 1301 Memorial Drive, Coral Gables, FL 33146. E-mail: [schultz@fig.cox.miami.edu](mailto:schultz@fig.cox.miami.edu).

selection is unlikely to influence our results appreciably, for several reasons. Any mutation homozygous in the zygote would have been passed to all gametes produced by that individual. Thus the effective number of gametes for the transmission of that mutation is one, and the mutation will be nearly neutral in the gamete phase if it reduces gamete fitness by less than 0.5. If a mutation is heterozygous in the zygote it will be passed to half its pollen and ovules. Thus the effective number of gametes for the transmission of that mutation is two, and the mutation will be nearly neutral in the gamete phase if it reduces gamete fitness by less than 0.25.

In a plant with a high pollen/ovule ratio, competition among pollen grains could increase the efficiency of selection against mutations with mild effect on pollen function (e.g., rate of tube growth). This competition is unlikely in our experiments, because in *A. thaliana*, the pollen/ovule ratio is low, and the absence of wind currents and perhaps insects reduces the movement of pollen onto the stigmatic surface (see refs. 33 and 34). In our fitness experiment, an average of only 34–36 seeds per silique was present in both controls and mutation lines. Because the number of ovules per silique in *A. thaliana* is 40–60, the low seed number in both groups suggests that the number of pollen grains deposited on a stigma averaged less than the number of ovules per flower. This finding is not surprising, considering that our plants were maintained indoors, each individual enclosed laterally by a clear tube. Wind currents were nearly nonexistent, insects were absent, and pollination was entirely passive self-pollination. This low ratio of transferred pollen to ovules suggests that the potential for pollen competition was low or nonexistent.

Mutations also may have been lost if their effects were so severe that they caused complete extinction of a mutation line; during the course of mutation accumulation, 76 lines (7.6%) were lost. Although a few of these cases were the result of self-sterility, most certainly were caused by random environmental factors. Most cases of nongermination, for example, were undoubtedly nongenetic, as they occurred in the drier outer and corner wells of our planting flats (total germination rate was about 83% per generation).

The loss of severe mutations by either means would have tended to bias our estimates of mean mutational decay and genomic mutation rate downward. This bias is acceptable, because our purpose was to generate conservative estimates of these values. Moreover, numerous studies have shown that mutations with severe effects constitute a small minority of mutational damage (in *Drosophila*, for example, lethals and sublethals are less than about 5% of all new deleterious mutations; refs. 19, 21, and 35). Thus we are confident that our assay, though conservative, revealed the majority of deleterious mutations in *Arabidopsis*.

**Control Lines and Design of Fitness Assay.** After 10 generations of mutation accumulation, we began a common garden experiment that enabled us to measure major components of fitness for both the mutation accumulation lines and ancestral control lines. Control lines were formed from the seeds stored as block 1 at the start of the mutation accumulation experiment. Our goal in forming the control lines was to minimize environmental effects, such as seed age or maternal environment, that could cause phenotypic differences between the control and mutation accumulation lines. Two generations before the start of the common garden experiment, the seeds from block 1 were sown to produce 200 adult plants. These 200 plants were raised in flats adjacent to the generation-9 mutation accumulation lines in the same growth chamber. At the start of the next generation, five seeds from each of the 200 plants were sown to found 1,000 control lines. At the same time, the seeds of the generation-10 mutation accumulation lines were sown in a checkerboard fashion with the control lines in 72-well flats. A single seedling from each line was chosen randomly to provide seed for the fitness assay in the

common garden experiment. In this manner, the seeds planted for both the control lines and the generation-10 lines in the actual fitness assay were of the same age and had been produced from parents raised in a common environment and grandparents raised in a nearly identical environment. Nongenetic differences between the control and generation-10 lines therefore should have been minimal.

**Measurement of Fitness Components.** Seeds from the 924 surviving mutation accumulation lines and from a random sample of 924 of the 1,000 control lines then were used in a common garden experiment. After planting these seeds, we measured three major components of fitness. These included the proportion of seeds that germinated, fruit set per individual germinated seed, and the number of seeds per fruit produced. The product of the three components of fitness is an estimate of the total number of seeds produced per seed sown. This composite trait is a reasonable measure of lifetime total fitness for this self-fertilizing annual plant, because total seed set encompasses both male and female function as well as viability.

These traits were assayed in a competition-free common garden experiment in a large growth chamber. Plants were grown under cultural and environmental conditions that were essentially identical to those used in the five generations before the fitness assay and in the growth of the seeds used in the fitness assay. The growth chamber environment was maintained at a constant 23°C with 60% humidity and alternating 12 hr of dark and light from 40-W cool white fluorescent bulbs. Plants were grown in 72-well flats in Sunshine Aggregate No. 4 potting mix (a mixture of peat moss, perlite, nutrients, gypsum, wetting agent, and dolomitic lime). Flats were bottom-watered without fertilizer every 2 or 3 days to maintain approximately uniform soil moisture. In the common garden experiment, the position of flats in the growth chamber was randomized weekly. (Because the large growth chamber was not purchased until the experiment had begun, the environment used during the first five generations of mutation accumulation differed slightly from that used in later generations. In the earlier generations, plants were grown in 96-well flats in a similarly lit growth room that was maintained at approximately 23°C.)

At the start of the common garden experiment, we planted seeds from one line per well. Lines were ordered in the 72-well flats such that in each flat 36 different randomly chosen mutation accumulation lines were sown in wells alternating with 36 different randomly chosen control lines. The layout for a given flat was replicated three times so that every line was represented in triplicate. Four seeds were sown per well, giving 12 seeds per line. Seeds were allowed to germinate, and all seedlings were counted. The germination rate was measured as the fraction of seeds planted that produced seedlings. All but one of the resulting seedlings per well were discarded at random, leaving three surviving plants per line, each in a unique flat. Because 76 lines were lost over the 10 generations, the total experiment comprised roughly 22,176 seeds (two treatments  $\times$  924 lines per treatment  $\times$  12 seeds per line) and 5,544 adults (two treatments  $\times$  924 lines  $\times$  three adults per line). Plants were allowed to senesce naturally, and all fruits were counted on each plant. It was not possible to count all the seeds produced per plant, so we counted seed set per fruit for two random fruits per plant, three plants per line, in 144 random lines of both mutation accumulation lines and control lines.

**Data Analysis and Estimation of Mutational Parameters.** As mutations accumulate they have two expected measurable consequences. First, because mutations are generally deleterious, they reduce mean fitness. Second, because they occur independently in different lines, they increase the variance in fitness among lines. If mutation is a Poisson process, then for alleles of moderate effect, the genomic deleterious mutation

rate ( $U$ ) and mean homozygous effect of a mutation ( $\bar{s}$ , the reduction in fitness relative to the nonmutant homozygote, taken here to be the fitness of the controls) can be estimated from the per-generation decline in mean fitness ( $R$ , the “mutational decay”) and increase in fitness variance among lines ( $V_b$ ) (16, 17, 36, 37).

The expectation of  $R$  is  $(1/2) UE(s)$  and that of  $V_b$  is  $(1/2) UE(s^2)$ . If there is no variation among mutations in their effect, then  $E(s)^2 = E(s^2)$  and  $U$  can be estimated as  $2R^2/V_b$ . An estimate of  $\bar{s}$  is  $V_b/(\bar{z}_0 R)$ , where  $\bar{z}_0$  is the mean fitness of the controls. Because mutations must vary in their effects on fitness, the mutation rate estimated as outlined above must be a downwardly biased estimate,  $U_{\min}$ , and the selection coefficient must be an upwardly biased estimate,  $\bar{s}_{\max}$ . The degree of bias in these estimates depends on the distribution of  $s$  among mutated alleles (see refs. 35 and 38). If the distribution is exponential, an unbiased estimate of  $U$  is exactly twice the estimated  $U_{\min}$ , and an unbiased estimate of  $\bar{s}$  is exactly half the estimated  $\bar{s}_{\max}$  (35). If the distribution is more leptokurtic than exponential, as is suggested by data in *Drosophila* (38), the unbiased estimate of  $U$  would be even larger and that of  $\bar{s}$  even smaller. These equations differ from those used in *Drosophila* mutation-accumulation experiments because of their use of balancer chromosomes, which prevent segregational loss of mutations.

The increase per generation in the genetic variance for an additive trait caused by mutation is the mutational variance,  $V_m$ . This quantity can be estimated as half the rate of increase in the among-line variance ( $V_b/2$ ). The ratio of mutational variance to the environmental variance,  $V_e$ , is known as the mutational heritability,  $h_m^2$ .

To obtain estimates of the mutational heritability, we estimated  $V_e$  as the within-line variance for the generation-0 plants by using ANOVA. Confidence intervals for all parameters estimated were obtained on bootstrapped data. Each generation was resampled 1,000 times randomly, with replacement, and the full analysis was redone on each resampled data set. Resampling was performed at the level of line: a line was chosen randomly, with replacement, until the number of these resampled lines equaled that in the original data set, for both control lines and mutation lines. For each line chosen, all replicates of that line were automatically chosen. Ninety-five percent confidence intervals and the proportion of bootstrap estimates less than zero were obtained from these 1,000 bootstrapped data sets.

For those bootstraps yielding negative estimates of mutational variance (indicated by the probability values cited in Table 3), we did not estimate mutational parameters. The true mutational variance cannot be negative, and substituting such impossible variances into the Bateman-Mukai equations would yield nonsense estimates. Nevertheless, negative estimates of mutational variance are potentially informative. Negative variances could be interpreted as evidence that the true (positive but unknown) mutational variance is extremely small. Thus, eliminating these cases from our estimates of  $U$  and  $s$  prevent these small mutational variances from influencing our results. Thus, our results tend to overestimate the mutational variance (and selection coefficient  $s$ ), and underestimate the genomic mutation rate  $U$ . Because our purpose in this experiment is to provide a conservative minimum estimate of  $U$  (thus a maximum estimate of  $s$ ), this bias is acceptable.

During the course of the common garden experiment it became apparent that the three genetically uniform, replicated flats for each unique set of control and generation-10 lines varied in obvious traits like time of flowering and stature despite weekly flat randomization. Unfortunately, our large number of lines (the minimum needed to demonstrate significant mean effects, as shown by a prior analysis of statistical power) precluded the use of a single, fully crossed ANOVA to determine the statistical significance of this potentially large source of environmental variation. Instead, we conducted a maximum-likelihood analysis of variance (39, 40) of the fitness traits in which the replicate flats were nested within the unique groups of control and experimental lines. With both group and flat taken as random effects, the ANOVAs confirmed that there was significant variation among replicate flats for all traits.  $\chi^2$  values with one degree of freedom were 25, 344, and 80 for the flat effect on germination rate, fruit set, and seed set, respectively (all these are significant at  $P = 0.001$ ). To reduce some of this noise, we estimated mutation parameters from variances calculated from the deviation of each trait value from the flat mean (Table 1 and see Table 3). In these ANOVAs, the error term estimated the among-flat variance within a line.

## RESULTS AND DISCUSSION

Individual fitness components declined by 0.05–0.4% per generation, and total fitness declined by 0.9% per generation (Tables 2 and 3). The estimates of mutational decay,  $R$ , were significantly different from zero for seed set and total fitness,

Table 1. ANOVA of fitness components for control lines and mutation accumulation lines

Trait	Generation	Source	Raw data			Transformed data		
			df	MS	F-value	df	MS	F-value
Germination	Controls	Line	933	0.0495	1.18***	933	0.0457	1.13***
		Error	1,758	0.0419		1,758	0.0403	
	Mutation accumulation	Line	925	0.0528	1.41***	925	0.0491	1.35***
		Error	1,744	0.0376		1,744	0.0364	
Fruit set	Controls	Line	932	1635.2	3.61***	932	505.5	1.27***
		Error	1,808	452.9		1,808	397.9	
	Mutation accumulation	Line	923	1670.4	3.24***	923	553.7	1.28***
		Error	1,791	516.0		1,791	433.3	
Seeds per fruit	Controls	Line	143	160.6	1.83***	143	80.6	1.21
		Error	277	87.6		277	66.5	
	Mutation accumulation	Line	142	223.9	2.90***	142	112.5	1.71***
		Error	273	77.2		272	65.6	
Total fitness, $\times 10^{-3}$	Controls	Line	124	1584.8	2.55***	124	708.5	1.13
		Error	250	621.0		250	625.5	
	Mutation accumulation	Line	126	1352.0	2.58***	126	715.7	1.41*
		Error	254	523.6		254	507.3	

ANOVAs are presented for both the raw values of the fitness components and the transformed values, where each observation was calculated as the deviation from the mean value for the flat (see text for details). \*,  $P < 0.05$ ; \*\*\*,  $P < 0.001$ . MS, mean square.

Table 2. Trait means at generations 0 and 10

Trait		Gen. 0	Gen. 10
Prop. germ.	Mean	0.845	0.839
	SE	0.003	0.004
	<i>N</i>	10,768	10,680
Fruit set	Mean	47.5	47.3
	SE	0.40	0.40
	<i>N</i>	2,741	2,715
Seed set	Mean	36.3	34.7
	SE	0.41	0.44
	<i>N</i>	421	416
Total fitness	Mean	1,460	1,380
	SE	20	22
	<i>N</i>	414	408

Means, SEs, and total sample size for proportion of seeds germinated, total fruit set, and number of seeds per fruit.

but not for germination success or fruit set. Minimum detectable differences in means (control-mutation lines) for germination, fruit set, and seed set were 0.03, 6, and 1.6, respectively, for a statistical power of 0.9 and  $\alpha = 0.05$  (ref. 41, see Tables 1 and 2).

We found significant variation among lines in both the controls and the generation-10 mutation accumulation lines for all three fitness components, regardless of whether the data were transformed as deviations from the flat means (Table 1). It is unlikely that mutational variance could account for the among-line variance in the control, because all such lines descended from a single plant within two generations. It is more likely that the among-line variance in the controls was caused by variation in environmental effects experienced by the maternal plants that produced the seeds used in the common garden experiment. If micro-environmental differences caused variation among the maternal plants in traits like seed size, then such effects could result in variation in maternal environmental effects for later components of fitness in the next generation. Because the 36 seeds and three seedlings from each line used in the common garden experiment all descended from a single maternal plant, variance in maternal environmental effects could result in among-line variance.

Because the same methods were used to generate both the control and mutation accumulation lines, these potential effects can be accounted for in the estimation of the mutational variance during the 10 generations of mutation accumulation. For all of our analyses on all of the traits, we assumed that the among-line variance in the controls provided an unbiased estimate of the sum of maternal environment effect variance and the mutational variance accrued during line formation. We also assumed that the among-line variance in the generation-10 lines provided an unbiased estimate of the sum of three sources of among-line variance: genetic variance caused by mutations that accumulated during the 10 generations, mutational variance because of the two generations of line formation, and the maternal environmental effect variance. Under these assumptions, the difference in among-line variance between the generation-10 lines and the controls provides

an estimate of the genetic variance caused by 10 generations of mutation accumulation ( $10 V_b$  or  $20 V_m$ ).

We assume here that the maternal environmental effect variance is the same in generations 1 and 10. Although this assumption seems reasonable, it is possible that deleterious mutations caused the generation-10 maternal plants to be more sensitive to the slight environmental variation. If this were the case, then by making the assumption of equal maternal variances we might overestimate the mutational variance, and therefore underestimate the downwardly biased estimate  $U_{min}$  and overestimate upwardly biased estimate  $\bar{s}_{max}$ . Second, there is the remote possibility of grandmaternal environmental effects on fitness. If they exist, then we might expect them to be slightly greater in the generation-10 plants, owing to the smaller number of grandparents in the controls (200) than the mutation-accumulation lines (1,000). Again, if true, such effects would lead us to overestimate  $V_m$ , underestimate  $U_{min}$ , and overestimate  $\bar{s}_{max}$ . Because our main goal is to place lower limits on the genomic deleterious mutation rate and upper limits on the average effect, the assumption of equal maternal environmental variance is a conservative one.

Finally, there is the hypothetical possibility that storage of our control seeds (dry and at room temperature) for 3 years may have altered their subsequent germination, growth, and reproduction, and this altered mean performance may have been retained across three generations during which offspring of controls and mutation lines were interdigitated in the same flats. Because we have no evidence of any such storage effects in *Arabidopsis*, let alone that they could have persisted across three generations, we assumed that any such effects are inconsequential.

The highly significant among-flat variation discussed above led us to use transformed data, in which each trait value was the taken as the deviation from the flat mean, to estimate mutational parameters. We also have estimated mutational parameters by using the raw data and found that the transformation made little difference to our estimates of mean mutational damage or total genomic mutation rate. However, we present the ANOVAs for line effects for both the raw data and the transformed data (Table 1).

Estimates of the mutational variance,  $V_m$ , were positive for all traits, with mutational heritabilities ( $V_m/V_e$ ) ranging from 0.0005 to 0.0030 (Table 3). The increase in variance among mutation lines was significant for seed set ( $F_{142, 143} = 1.40, P = 0.023$ ; see Table 1). The minimum detectable among-line variance ratios (mutation lines: control lines) for germination, fruit set, seed set, and total fitness were 1.11, 1.11, 1.31, and 1.34, respectively, for a statistical power of 0.9 at  $\alpha = 0.05$  (41). Despite the significant variance increase for seed set, error propagation in the estimates of mutational variances and heritabilities made these composite quantities insignificantly different from zero. We might have obtained significant mutational variances had the lines accumulated mutations for many more additional generations; however, continuing the experiment for 5–10 more years (yielding approximately 40 generations of mutation accumulation) was beyond the scope of our study. We note that it is not unusual for well-designed

Table 3. Mutational decay, variance, heritability, rate, and mean effect

Trait	$R/\bar{z}_0$	$V_b/(\bar{z}_0)^2$	$h_m^2$	$U_{min}$	$s_{max}$
Germ. est.	0.00062	0.00041	0.0032	0.0024*	0.62
Limits	(-0.0008, 0.002)	( $-2 \times 10^{-6}$ , 0.0009)	(-0.001, 0.007)	( $3 \times 10^{-5}$ , 0.05)	(-1, +1)
Fruit est.	0.00054	0.00024	0.00051	0.0030*	0.42
Limits	(-0.001, 0.004)	(-0.002, 0.002)	(-0.003, 0.005)	( $2 \times 10^{-5}$ , 0.2)	(-1, +1)
Seed est.	0.0044*	0.00083	0.0081	0.054*	0.24*
Limits	(0.001, 0.009)	(-0.0007, 0.004)	(-0.003, 0.05)	(0.005, 0.4)	(0.02, 0.6)
Total est.	0.0089*	0.0020	0.0031	0.10*	0.23*
Limits	(0.001, 0.03)	(-0.005, 0.008)	(-0.006, 0.01)	(0.004, 0.8)	(0.02, 0.9)

Estimates of mutational effects (\* for  $P < 0.05$ ) and their 2.5 and 97.5 percentiles from 1,000 bootstraps of the data set.

mutation accumulation experiments to fail to obtain mutational variances that are significantly different from zero (e.g., ref. 23; those authors, however, limited their full analyses to traits exhibiting significant mutational variances). Moreover, prolonging such an experiment presents some disadvantages, such as increased loss of lines caused by environmental variation, and epistatic interactions among accumulated mutations that bias the estimators.

From the per-generation decline in mean fitness and increase in fitness variance, the minimum genomic deleterious mutation rate to alleles reducing total fitness,  $U_{\min}$ , in *Arabidopsis* is estimated to be 0.1 (Table 3). The 95% confidence intervals for this estimate are 0.005 and 0.600. Estimates of the genomic deleterious mutation rate for fitness components ranged from 0.002 to 0.050. The upwardly biased estimate of the mean homozygous effect of a mutation,  $\bar{s}_{\max}$ , is 0.2 for total fitness, with 95% confidence intervals of 0.006 and 0.9 (Table 3). Estimates of  $\bar{s}_{\max}$  for the fitness components ranged from 0.2 to 0.6, with large confidence intervals.

Our estimated mean effects of deleterious mutations are larger than those estimated from other mutation-accumulation studies. For example,  $\bar{s}_{\max}$  was estimated to be about 0.06 in *Drosophila* (16, 18) and 0.01 in *E. coli*, (the haploid effect, see ref. 22). However, the large confidence limits on our estimates make them consistent with the possibility that the majority of deleterious mutations are of mild effect. Theoretical studies (see ref. 42) have shown that mutations of intermediate to small effect (depending on the population size) are maximally effective in causing extinction of small populations, because of their fixation during random genetic drift.

Our estimate of the genomic deleterious mutation rate can be compared with the expected substitution rate by noting that the average substitution rate per nucleotide site in plants is approximately  $6 \times 10^{-9}$  per year (43), that the diploid genome size in *Arabidopsis* is roughly  $1.4 \times 10^8$  nt (44), and that transposable-element activity in *A. thaliana* is negligible. With 2–3 generations per year (probably a maximum over most of the natural range of this winter annual), the molecular data predict a genomic mutation rate of 0.28–0.42 per diploid genome per generation, or 3–4 times our minimum estimate of the deleterious rate. Similar calculations suggest that the total mutation rate in *Drosophila melanogaster* not attributable to transposons is 2–8 times the deleterious rate (see refs. 16, 18, 35, and 45–47), and in *E. coli* about 10 times the deleterious rate (22). We should note that the estimates for *Drosophila* differ greatly among several recent studies (e.g., refs. 20 and 48–51), and discussion of the methods, assumptions, and implications of these conflicting investigations is ongoing (19, 21).

Our estimate of the genomic deleterious mutation rate can be used to predict levels of inbreeding depression in *Arabidopsis* at equilibrium of mutation and selection, and in the absence of heterozygote advantage, using the method of Charlesworth *et al.* (26). This method requires an estimate of the arithmetic mean dominance of mutations and is relatively insensitive to assumptions regarding the distribution of dominance or selective effects of mutations. The Charlesworth *et al.* (26) method, however, encompasses all mutations, whereas our study necessarily excluded mutations of large effect. If such mutations are rare (they constitute only a few percent of all mutations in *Drosophila*, ref. 35), then the two methods should yield comparable results. Indeed in this case, they do: if the mutations in our study are partially recessive with arithmetic mean dominance  $h = 0.3$  (as estimated in one species of highly selfing *Amsinckia*, ref. 29; see also ref. 35), and our estimate of genomic deleterious mutation rate is roughly half the true value (as predicted if mutation effects follow an exponential distribution; see ref. 35 and *Materials and Methods*), then its confidence limits predict that the equilibrium ratio of fitnesses in selfed to randomly outcrossed offspring should lie from 0.81

to near 1 (26). In empirical studies, this ratio ranged from 0.88 to 0.93 for total plant dry weight in *Arabidopsis* (see ref. 26). The observed values of inbreeding depression in *Arabidopsis* are within the range predicted by our direct estimate of genomic deleterious mutation rate. Although these calculations, of course, should be treated with caution, they provide direct confirmation of the utility of inbreeding depression as an assay of mutation rate in large, stable, and highly inbred populations.

This paper is dedicated to the memory of Frank L. Allard. We thank D. Charlesworth, D. Houle, A. Kondrashov, R. Lande, R. Shaw, and three anonymous reviewers for helpful comments on the manuscript and experimental design, and J. Houle, S. Durand, and a team of tireless research assistants for plant cultivation and data collection. The work was supported by National Science Foundation Grants BSR-9024977 and DEB-9419884, and National Institutes of Health Grant RO1-GM36827.

- Kondrashov, A. S. (1988) *Nature (London)* **336**, 435–440.
- Charlesworth, B. (1990) *Genet. Res.* **55**, 199–222.
- Charlesworth, D. & Charlesworth, B. (1987) *Annu. Rev. Ecol. System.* **18**, 237–268.
- Lande, R. & Schemske, D. W. (1985) *Evolution* **39**, 24–40.
- Charlesworth, D., Morgan, M. T. & Charlesworth, B. (1990) *Evolution* **44**, 1469–1489.
- Lande, R., Schemske, D. W. & Schultz, S. T. (1994) *Evolution* **48**, 965–978.
- Schultz, S. T. & Willis, J. H. (1995) *Genetics* **141**, 1209–1223.
- Lynch, M., Burger, R., Butcher, D. & Gabriel, W. (1993) *J. Hered.* **84**, 339–344.
- Lande, R. (1994) *Evolution* **48**, 1460–1469.
- Lynch, M., Conery, J. & Burger, R. (1995) *Evolution* **49**, 1067–1080.
- Lynch, M., Conery, J. & Buerger, R. (1995) *Am. Nat.* **146**, 489–518.
- Schultz, S. T. & Lynch, M. (1997) *Evolution* **51**, 1363–1371.
- Charlesworth, B., Morgan, M. T. & Charlesworth, D. (1993) *Genetics* **134**, 1289–1303.
- Lynch, M., Latta, L., Hicks, J. & Giorgianni, M. (1998) *Evolution* **52**, 727–733.
- Charlesworth, B. & Charlesworth, D. (1998) *Genetica* **103**, 3–19.
- Mukai, T. (1964) *Genetics* **50**, 1–19.
- Mukai, T., Chigusa, S. I., Mettler, L. E. & Crow, J. F. (1972) *Genetics* **72**, 335–355.
- Ohnishi, O. (1977) *Genetics* **141**, 1209–1223.
- Drake, J. W., Charlesworth, B., Charlesworth, D. & Crow, J. F. (1998) *Genetics* **148**, 1667–1686.
- Fry, J. D., Keightley, P. D., Heinsohn, S. L. & Nuzhdin, S. V. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 574–579.
- Lynch, M., Blanchard, J., Houle, D., Kibota, T., Schultz, S., Vassilieva, L. & Willis, J. (1999) *Evolution* **53**, 645–663.
- Kibota, T. T. & Lynch, M. (1996) *Nature (London)* **381**, 694–696.
- Keightley, P. D. & Caballero, A. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 3823–3827.
- Vassilieva, L. L. & Lynch, M. (1999) *Genetics* **151**, 119–129.
- Morton, N. E., Crow, J. F. & Muller, H. J. (1956) *Proc. Natl. Acad. Sci. USA* **42**, 855–863.
- Charlesworth, B., Charlesworth, D. & Morgan, M. T. (1990) *Nature (London)* **347**, 380–382.
- Deng, H. W. & Lynch, M. (1997) *Genetics* **147**, 147–155.
- Charlesworth, D., Lyons, E. E. & Litchfield, L. B. (1994) *Proc. R. Soc. London Ser. B* **258**, 209–214.
- Johnston, M. O. & Schoen, D. J. (1995) *Science* **267**, 226–229.
- Crow, J. F. (1993) *Environ. Mol. Mutagen.* **21**, 122–129.
- Crow, J. F. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 8380–8386.
- Schultz, S. T. & Ganders, F. R. (1996) *Evolution* **50**, 842–855.
- Wu, H.-M. & Cheung, A. Y. (1998) In *Arabidopsis*, eds. Anderson, M. & Roberts, J. (CRC, Boca Raton, FL), pp. 181–222.
- Bowman, J. (1994) *Arabidopsis: An Atlas of Morphology and Development* (Springer, New York).
- Simmons, M. J. & Crow, J. F. (1977) *Annu. Rev. Genet.* **11**, 49–78.
- Lynch, M. (1988) *Genet. Res.* **51**, 137–148.
- Bateman, A. (1959) *Int. J. Radiat. Biol.* **1**, 170–180.
- Keightley, P. D. (1994) *Genetics* **138**, 1315–1322.
- Harville, D. A. (1977) *J. Am. Stat. Assoc.* **72**, 320–340.
- Dixon, W. J. (1992) *BMDP* (Statistical Software, Los Angeles).

41. Zar, J. H. (1996) *Biostatistical Analysis* (Prentice Hall, Upper Saddle River, NJ).
42. Gabriel, W., Lynch, M. & Buerger, R. (1993) *Evolution* **47**, 1744–1757.
43. Wolfe, K. H., Li, W.-H. & Sharp, P. M. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 9054–9058.
44. Shields, R. (1989) *Nature (London)* **337**, 308.
45. Catcheside, D. G. (1977) *The Genetics of Recombination* (University Park Press, Baltimore).
46. Sharp, P. M. & Li, W. H. (1989) *J. Mol. Evol.* **28**, 398–402.
47. Nuzhdin, S. V., Pasyukova, E. G. & Mackay, T. F. C. (1995) *Mol. Biol. Evol.* **12**, 180–181.
48. Garcia-Dorado, A. (1997) *Evolution* **51**, 1130–1139.
49. Garcia-Dorado, A., Monedero, J. L. & Lopez-Fanjul, C. (1998) *Genetica* **103**, 255–265.
50. Keightley, P. D. (1996) *Genetics* **144**, 1993–1999.
51. Shabalina, S. A., Yampolsky, L. Y. & Kondrashov, A. S. (1997) *Proc. Natl. Acad. Sci. USA* **94**, 13034–13039.