

Inference of Genome-Wide Mutation Rates and Distributions of Mutation Effects for Fitness Traits: A Simulation Study

Peter D. Keightley

Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, Scotland

Manuscript received April 15, 1998
Accepted for publication July 20, 1998

ABSTRACT

The properties and limitations of maximum likelihood (ML) inference of genome-wide mutation rates (U) and parameters of distributions of mutation effects are investigated. Mutation parameters are estimated from simulated experiments in which mutations randomly accumulate in inbred lines. ML produces more accurate estimates than the procedure of Bateman and Mukai and is more robust if the data do not conform to the model assumed. Unbiased ML estimates of the mutation effects distribution parameters can be obtained if a value for U can be assumed, but if U is estimated simultaneously with the distribution parameters, likelihood may increase monotonically as a function of U . If the distribution of mutation effects is leptokurtic, the number of mutation events per line is large, or if genotypic values are poorly estimated, only a lower limit for U , an upper limit for the mean mutation effect, and a lower limit for the kurtosis of the distribution can be given. It is argued that such lower (upper) limits are appropriate minima (maxima). Estimates of the mean mutational effect are unbiased but may convey little about the properties of the distribution if it is leptokurtic.

MUTATIONS that affect fitness are usually deleterious and rarely become fixed in large populations. However, deleterious mutations may occur at a sufficiently high rate to play an important role in several key evolutionary phenomena, such as the evolution and maintenance of sex. Some evolutionary theories require estimates of the genomic rate of deleterious mutations, U , but not necessarily of the distribution of their selective effects [*e.g.*, to test the “deterministic mutation” theory for the evolution of sex (Kondrashov 1993)], while others require information both on U and the distribution of selective values [*e.g.*, to infer the probability of “mutational meltdown” (Lande 1995; Lynch *et al.* 1995)].

There are several ways to obtain information on U and distributions of selective values. One approach, proposed by Kondrashov and Crow (1993), is to compare between-species nucleotide substitution rates at a sample of regions in the genome to rates in regions evolving unconstrained by natural selection. If selection has operated to remove new mutations in the sampled regions, the rate of nucleotide substitution will be lower than in the unconstrained regions. By combining the relative substitution rates with estimates of between-species divergence times and generation intervals, it is possible, in principle, to estimate U . For example, the rate of nucleotide substitution in human pseudogenes implies that about three point mutations occur in amino acid

coding regions per zygote per generation (Drake *et al.* 1998). Protein coding sequences are under moderate constraint, so this rate is similar to the deleterious mutation rate specific to such sequences (A. Eyre-Walker and P. D. Keightley, unpublished results). The estimate refers to fitness-altering mutations occurring on an evolutionary time scale, but does not tell us about the magnitudes of the effects upon which natural selection acted (except that fitness effects were greater than the reciprocal of the effective population size), or about the effects of large-scale molecular changes such as deletions or transposable element (TE) insertions.

A second general approach to indirectly infer U and mutation parameters is from a comparison of the distributions of fitnesses of outbred (or inbred) individuals sampled from a natural population to their inbred (or outbred) progeny. A method developed by Deng and Lynch (1996) for outbreeding populations allows estimation of U and the mean mutant effect and degree of dominance. However, with this method deleterious mutation-selection balance is assumed to be the only mechanism maintaining variation for fitness, and biased estimates will result if any other mechanisms, such as balancing selection or migration, lead to the maintenance of variation (Drake *et al.* 1998). For example, negative genetic correlations between major components of fitness can lead to the maintenance of additive and nonadditive genetic variation for the major components and fitness, respectively (Rose and Charlesworth 1980; Rose 1982; Falconer and Mackay 1996, chapter 20). In plant species that normally reproduce by selfing, a comparison of the fitnesses of selfed and

Address for correspondence: Institute of Cell, Animal and Population Biology, University of Edinburgh, W. Mains Rd., Edinburgh EH9 3JT, Scotland. E-mail: p.keightley@ed.ac.uk

outcrossed progeny can provide an estimate of U (Charlesworth *et al.* 1990). As with the method applied to outbreeders, it is assumed that genetic variation is maintained solely by a balance between mutation and selection. Charlesworth *et al.*'s (1990) method provides estimates for U , but not of mutation distribution parameters such as the mean mutation effect.

A third direct way to study effects of deleterious mutations is the mutation accumulation (MA) approach in the laboratory. Mutations are allowed to randomly accumulate in benign conditions in sublines derived from an inbred base population. The sublines are maintained by close inbreeding or as replicated chromosomes protected by a balancer chromosome, so drift will tend to dominate selection. After many generations of mutation accumulation, fitnesses of the MA lines or chromosomes are compared to controls. The approach was pioneered by Mukai (1964) who used a *Cy* balancer chromosome as a control to study the viability effects of mutation accumulation in *Drosophila melanogaster* second chromosomes. In contrast to the indirect approach outlined above, mutation should be the only source of evolutionary divergence between the lines under test, but the method is restricted to laboratory populations. Inference of U and properties of the distribution of mutation effects are made by comparing the distribution of phenotypes of MA and control lines. The method of Bateman (1959), which was taken up by Mukai [1964; the Bateman-Mukai (BM) method], compares the rate of increase in among-MA line variance per generation, V_m , with the rate of change of fitness or fitness component, ΔM . Under the assumption that mutations have equal deleterious effects, an estimate of U is obtained from

$$\hat{U} = \Delta M^2 / V_m, \quad (1)$$

and an estimate of the mean mutation effect comes from

$$\hat{E}(a) = V_m / \Delta M. \quad (2)$$

If mutation effects vary, \hat{U} is underestimated and $\hat{E}(a)$ is overestimated. This is analogous to the estimation bias for the effective number of factors influencing a quantitative trait (Falconer and Mackay 1996, chapter 12). Furthermore, if the distribution of mutation effects is strongly leptokurtic, the mean mutation effect conveys little about the properties of the distribution. Genomes contain sites that vary greatly in functional significance, so the distribution of selective values of new mutations is of central interest.

To infer U and distributions of mutation effects from MA experiments, an alternative approach is to assume that the true distribution of mutation effects follows some family of distributions. Theoretically, the distribution of effects of all mutations that occurred in an experiment could be estimated, but in practice there are insufficient degrees of freedom, so a family of distributions must be assumed. The family is chosen so that

changes in its parameters produce distributions with a wide range of properties. Numerical techniques involving maximum likelihood (ML; Keightley 1994) or minimum distance (Garcia-Dorado 1997) have been applied to find values for U and the parameters of the assumed distribution that best fit phenotypic data from the MA experiment. Fit is measured either in terms of $\Sigma \log$ likelihood of data or distance between the fitted and empirical distributions of line means. Following the nearly neutral model (Kimura 1983; Ohta 1992), gamma distributions of mutation effects have mostly been assumed. The gamma distribution has two parameters, α and β , specifying scale and shape, respectively. With small values of β , the distribution is leptokurtic: most mutations have effects close to zero; larger effects occur with diminishing frequency and contribute most of the between-line variance. Special cases are $\beta = 1$ (the exponential distribution) and $\beta = 1/2$ (the chi-square distribution with 1 d.f.). Gamma distributions with higher values of β approach symmetry about the mean (β/α). The case of $\beta \rightarrow \infty$ corresponds to equal mutation effects, the model that is generally used to obtain BM estimates.

ML or minimum distance methods have been applied to data from several MA experiments (Keightley 1994, 1996; Garcia-Dorado 1997; Keightley and Caballero 1997; Keightley and Ohnishi 1998). The algorithms to fit the mutation parameters have previously relied on highly computer-intensive Monte Carlo methods, and only limited investigations of the performance of the procedures have been possible (Keightley 1994; Garcia-Dorado 1997). Recently, an improved version of the ML procedure has been developed with algorithms ~ 2 orders of magnitude less demanding of computer time based on numerical integration (Keightley and Ohnishi 1998). An extensive study of the properties of ML estimation of mutation parameters is now feasible.

The purpose of this article is to explore the properties and limitations of ML inference of mutation parameters by simulation. Simulated rather than real MA data are analyzed, because the true parameter values are known, the model assumptions are not violated, and replication to detect significant deviations from expectation is possible.

MATERIALS AND METHODS

Model and simulation of data: The data available for analysis are assumed to be phenotypic means from a set of control lines (generation 0 of the MA experiment) and a set of MA lines (from generation t). In principle, data from several time points could be analyzed simultaneously, and an algorithm has been proposed (Keightley 1994). The power of the procedure can also be improved by making use of within-MA line replicate information (J. D. Fry, P. D. Keightley, S. L. Heinsohn and S. V. Nuzhdin, unpublished results). Control line phenotypic values were random normal deviates with

TABLE 1
Comparison of BM and ML procedures in which fixed values of β (β^*) are assumed, with $\sigma_g^2/\sigma_e^2 = 20$

Model	Simulated values		Estimation procedure	Estimated values						
	U	$E(a)$		β^*	\bar{U}	SD	$\widehat{E(a)}$	SD	$\log L$	SD
Equal effects ($\beta \rightarrow \infty$)	0.5	6.32	BM	$\rightarrow \infty$	0.50	(0.079)	6.35	(0.68)	—	—
			ML	$\rightarrow \infty$	0.50	(0.047)	6.32	(0.095)	0	—
				2	0.61	(0.093)	5.36	(0.35)	-48	9.3
				1	0.80	(0.13)	4.12	(0.34)	-55	10.0
				0.5	1.15	(0.20)	2.94	(0.27)	-61	10.6
$\beta = 0.5$	0.5	3.65	BM	$\rightarrow \infty$	0.17	(0.034)	11.1	(2.13)	—	—
			ML	$\rightarrow \infty$	0.32	(0.071)	5.49	(0.86)	-32	12
				2	0.28	(0.044)	6.45	(1.25)	-1.2	1.8
				1	0.36	(0.057)	5.14	(1.07)	-0.0	0.8
				0.5	0.51	(0.086)	3.65	(0.78)	0	—
Equal effects ($\beta \rightarrow \infty$)	2.5	2.83	BM	$\rightarrow \infty$	2.52	(0.29)	2.83	(0.29)	—	—
			ML	$\rightarrow \infty$	2.53	(0.20)	2.80	(0.17)	0	—
				2	3.66	(0.39)	1.97	(0.18)	-3.0	2.1
				1	4.77	(0.51)	1.51	(0.14)	-4.3	2.6
				0.5	6.99	(0.76)	1.04	(0.10)	-6.0	3.2
$\beta = 0.5$	2.5	1.63	BM	$\rightarrow \infty$	0.84	(0.12)	4.92	(0.71)	—	—
			ML	$\rightarrow \infty$	1.29	(0.17)	3.14	(0.33)	-18	6.2
				2	1.37	(0.18)	2.99	(0.32)	-1.8	1.9
				1	1.76	(0.24)	2.35	(0.26)	-0.3	0.9
				0.5	2.54	(0.35)	1.64	(0.19)	0	—
	0.25	4.10	(0.57)	1.02	(0.12)	-0.3	0.6			

mean μ and variance σ_e^2 . Phenotypic values of MA lines were sums of independent random normal deviates, mean μ , variance σ_e^2 as above, plus mutational effects generated by summing n random deviates from a gamma distribution, parameters α and β . Epistasis between mutations was not modeled. n was sampled from a Poisson distribution with parameter U (or $n = 1$ as a special case, where the absolute number of mutation events was assumed to be known). Here, U is the mean number of mutations accumulated per MA line, and would be divided by t to estimate the mutation rate per generation for a real experiment. Following Garcia-Dorado (1997), to compare simulations with different β and U , the variance of genotypic values $\sigma_g^2 = U\beta(\beta + 1)/\alpha^2$ was set at a fixed multiplier of the among-line error variance (σ_e^2), often 5 or 20. Precision levels (expressed as σ_g^2/σ_e^2) achieved in some previous large-scale MA experiments are, for example, as follows: Mukai *et al.* (1972) ~ 7 (one replicate); Ohnishi (1974) ~ 8 ; Keightley and Ohnishi (1998) mean 12 for nine traits, range 1.0–56.

ML analysis: The numerical integration procedure to estimate mutation parameters is fully described elsewhere (Keightley and Ohnishi 1998). The same model was assumed as was used to generate the simulated data. The parameters estimated in the model were μ , σ_e^2 , U , α , and β . Distributions reflecting about zero, with a parameter P , the proportion of positive mutant effects, have been investigated previously (Keightley 1994; Keightley and Ohnishi 1998), but are not investigated here. The α and β parameters fully specify the properties of the distribution of mutation effects. Traditionally the mean mutational effect has been estimated from MA experiments. If the distribution is leptokurtic, it is most logical to estimate α and β , but the mean mutant effect is likely to be of continued interest, so β and $E(a) = \beta/\alpha$ are given

here. Parameter estimates are based on “profile likelihoods,” computed by maximizing the likelihood for a series of fixed values of one parameter. Likelihood surfaces often become very flat, so maximization could fail if all parameters are fitted simultaneously. ML estimates were obtained from profile likelihoods, typically involving 10–20 points, by fitting a quadratic curve to the highest likelihood point and the nearest points on either side. Tests in which additional points were added about the ML did not significantly change the results. C computer code to carry out the likelihood calculations is available on request.

RESULTS

Comparison to the Bateman-Mukai method with shape of the distribution of mutation effects assumed:

In principle, the shape of the distribution of mutation effects will always need to be estimated. However, it is useful to compare the fit of models with different fixed values of β to explore the behavior of the ML procedure. By assuming equal mutation effects in the ML analysis ($\beta \rightarrow \infty$), the performance of the ML and BM procedures can also be compared. Tables 1 and 2 show means and standard deviations (SDs) for estimates of U and $E(a)$ from 30 replicate simulations (1000 replicates for $\beta \rightarrow \infty$). Two values of U were simulated, with equal mutation effects, or a gamma distribution with $\beta = 0.5$ (corresponding to a strongly leptokurtic distribution). The genetic variance was $20\sigma_e^2$ (Table 1) or $5\sigma_e^2$ (Table

TABLE 2
Comparison of BM and ML procedures with $\sigma_g^2/\sigma_e^2 = 5$

Model	Simulated values		Estimation procedure	Estimated values						
	U	$E(a)$		β^*	\bar{U}	SD	$\widehat{E(a)}$	SD	$\overline{\log L}$	SD
Equal effects ($\beta \rightarrow \infty$)	0.5	3.16	BM	$\rightarrow \infty$	0.50	(0.10)	3.21	(0.44)	—	—
			ML	$\rightarrow \infty$	0.51	(0.064)	3.15	(0.19)	0	—
				2	0.79	(0.14)	2.13	(0.25)	-6.1	3.2
				1	1.08	(0.19)	1.60	(0.20)	-8.2	3.8
				0.5	1.62	(0.29)	1.08	(0.14)	-10.2	4.3
$\beta = 0.5$	0.5	1.83	BM	$\rightarrow \infty$	0.18	(0.056)	5.46	(1.41)	—	—
			ML	$\rightarrow \infty$	0.20	(0.058)	4.08	(0.81)	-11.3	5.8
				2	0.24	(0.049)	3.64	(0.79)	-0.9	1.4
				1	0.33	(0.067)	2.75	(0.63)	-0.2	0.6
				0.5	0.51	(0.10)	1.86	(0.46)	0	—
Equal effects ($\beta \rightarrow \infty$)	2.5	1.41	BM	$\rightarrow \infty$	2.51	(0.36)	1.42	(0.18)	—	—
			ML	$\rightarrow \infty$	2.52	(0.36)	1.42	(0.18)	0	—
				2	3.75	(0.41)	0.96	(0.092)	-0.7	1.1
				1	4.98	(0.54)	0.72	(0.070)	-1.4	1.7
				0.5	7.46	(0.82)	0.49	(0.047)	-2.1	2.1
$\beta = 0.5$	2.5	0.82	BM	$\rightarrow \infty$	12.4	(1.37)	0.29	(0.029)	-2.8	2.5
			ML	$\rightarrow \infty$	0.85	(0.16)	2.44	(0.39)	—	—
				2	0.92	(0.16)	2.22	(0.30)	-4.5	3.8
				1	1.25	(0.24)	1.64	(0.27)	-0.8	1.4
				0.5	1.66	(0.32)	1.25	(0.21)	-0.2	0.6
	0.5	2.49	(0.47)	0.84	(0.15)	0	—			
	0.25	4.13	(0.79)	0.51	(0.091)	-0.1	0.4			

2), and 200 control and MA lines were simulated. The data for each replicate were analyzed by ML as described above, and by the BM method (Equations 1 and 2), with V_m estimated as the difference between the among-MA line and control line variances. Tables 1 and 2 illustrate a number of interesting results: (1) When the data conform to the model assumed (*i.e.*, simulated and assumed β are the same), the ML and BM procedures give mean estimates very close to the simulated values. ML provides good mean parameter estimates if the model conforms with the data irrespective of the shape of the distribution, *i.e.*, the mean $\log L$ is highest for

the β value corresponding to the simulated distribution. (2) If a model corresponding exactly to the data is assumed, ML provides more accurate estimates than BM (*i.e.*, coefficients of variation for the estimates are lower). This effect is particularly apparent for the case of few mutations with equal effects measured with low error (Table 1), presumably because the MA line data tend to fall into discrete classes. (3) If the model does not correspond to the data (*e.g.*, $\beta = 0.5$ simulated, but $\beta \rightarrow \infty$ assumed), ML provides mean parameter estimates closer to the values simulated than BM. ML is therefore more robust to deviations from the true distribution than BM. (4) For the U values simulated, ML can distinguish better between distributions if there are few mutations per MA line. (5) ML can distinguish better between different distributions if the true distribution is platykurtic (*i.e.*, with $\beta \rightarrow \infty$ simulated, the average change in $\log L$ between fitted distributions is very large). If the true distribution is leptokurtic, little information can be obtained on β , beyond inferring that the model of equal effects gives a poor fit.

Number of mutation events known: In certain experimental situations the number of mutation events per genome is known. In *Drosophila*, mobilization of *P* elements has been used to generate lines with single independent insertions (Lyman *et al.* 1996), and in *Escherichia coli*, strains with fixed numbers of *Tn10* insertions

TABLE 3

Parameter estimates from analysis of simulated data sets with the expected number of mutation events (U^*) fitted as a fixed parameter

U^*	Simulated values		Estimated values			
	β	$\overline{E(a)}$	$\bar{\beta}$	SD	$\widehat{E(a)}$	SD
0.5	0.5	3.65	0.59	(0.23)	3.99	(0.81)
	1.0	4.46	1.16	(0.43)	4.34	(0.55)
	2.0	5.16	2.52	(1.05)	5.26	(0.60)
5.0	0.5	1.16	0.54	(0.11)	1.17	(0.079)
	1.0	1.41	1.04	(0.22)	1.42	(0.066)
	2.0	1.63	2.09	(0.78)	1.64	(0.054)

TABLE 4
Analysis of simulated data where simulated values are favorable.

Simulated values			Estimated values								
			\bar{U}			$\bar{\beta}$			$\overline{E(a)}$		
U	β	$E(a)$	Min	ML	Max	Min	ML	Max	Min	ML	Max
2.5	1.0	4.46	1.74 (0.75)	2.87 (1.32)	$\rightarrow \infty$ [4]	0.37 (0.31)	1.20 (0.70)	3.04 (1.71)	2.22 (1.42)	4.38 (1.25)	6.19 (0.80)

The variance of genotypic values was $100\sigma_g^2$.

Standard deviations of estimates are shown in parentheses; numbers of runs (out of 30) for which a minimum or maximum estimate occurs at a boundary of 0 or ∞ are shown in brackets.

have been generated (Elena and Lenski 1997; Elena *et al.* 1998). Effects of TE insertion on the distribution of quantitative trait phenotypes can then be measured relative to nonmutagenized control lines. Analysis of simulated data with exactly one mutation per MA line showed that mean parameter estimates are very close to the values simulated, with no directional bias, and that the mean mutation effect is estimated with considerably higher accuracy than the distribution shape parameter (data not shown).

There are certain experimental situations where the expected rather than absolute number of mutation events can be estimated. For example, rates of accumulation of spontaneous TE insertions and base pair substitutions can be used to indirectly estimate the per genome mutation rate in *Drosophila*, albeit imprecisely (Keightley 1994; Drake *et al.* 1998). Table 3 compares estimates of β and $E(a)$ for simulations with “low” ($U = 0.5$) and “high” ($U = 5$) numbers of mutations. The number of mutation events per line was Poisson distributed. In the analysis the expected number was fitted as a known parameter. Mean estimates of β are close to the simulated values, but there is a slight but consistent upward bias. The bias, which is explored more fully in the next sections, implies that the parameters in the model are confounded with one another. There does not appear to be any directional bias for estimates of $E(a)$, which is again estimated much more precisely than β .

Unknown numbers of mutation events, “favorable” data: Highly precise measurement of genotypic values is unrealistic in an experimental setting, but should be favorable for disentangling the parameters. To model such a situation, 30 data sets were independently generated with the among-line variance of genotypic values, $\sigma_g^2 = 100\sigma_e^2$. Two hundred control and MA lines were simulated with $U = 2.5$ and $\beta = 1$. Table 4 shows mean ML parameter estimates along with mean lower and upper support limits based on \log_e likelihood drops of 2 from the MLs (asymptotically approximately equivalent to 95% confidence limits). Mean ML estimates

agree reasonably closely with their simulated parameter values. Frequency distributions of the estimates (Figure 1) show, however, that distributions are skewed upward in the case of \bar{U} , and downward in the case of $\overline{E(a)}$. The presence of skew turns out to be a consistent feature where U is estimated as an unknown parameter. A small number of \bar{U} are very much larger than the simulated value, while $\overline{E(a)}$ for the corresponding simulations

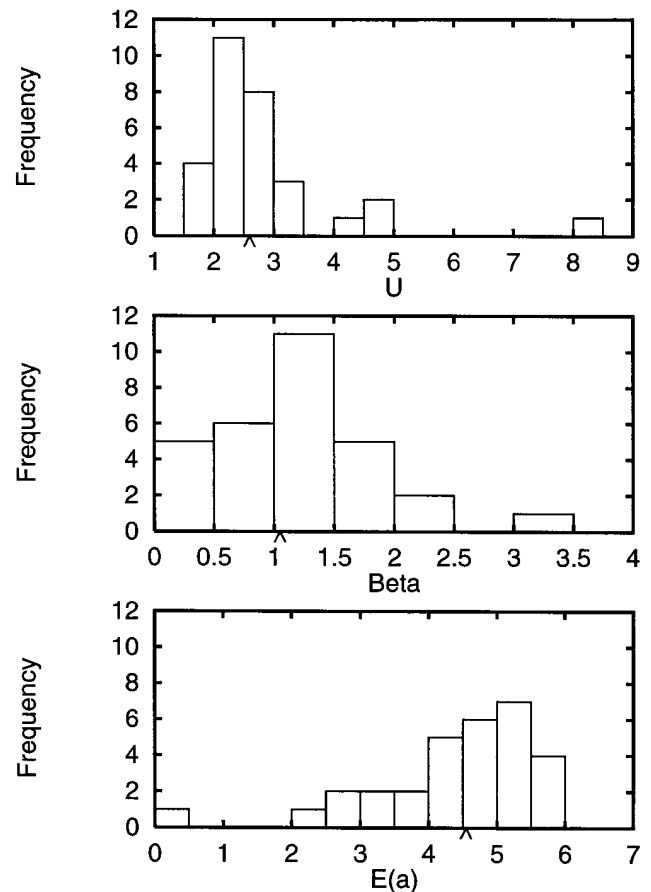


Figure 1.—Frequency distributions of ML estimates of mutation parameters from the favorable data sets. Simulated parameter values are indicated by “^.”

are very much smaller than their true values. For such data sets, the $\hat{\beta}$ tend to be close to zero. An explanation for this behavior can be found in the moments of the distribution of genotypic values.

The wide range for the mean lower and upper support limits for \hat{U} and $\hat{\beta}$ (Table 4) implies wide confidence intervals, even for the “favorable” parameter values. The upper support limit for $\hat{U} \rightarrow \infty$ because in 4 of the 30 simulations, log likelihood leveled out at a value < 2 different from the ML. The criterion for the confidence interval is a drop in log likelihood of 2, and is based on asymptotic properties, which, presumably, are not being met. In these 4 pathological cases, no upper limit for U can be given. In the same 4 simulation runs, the lower support limits for both $\hat{\beta}$ and $\hat{E}(a)$ approached zero, so only upper limits can be given. The data would be consistent, therefore, with an extremely high mutation rate, with effects sampled from a strongly leptokurtic distribution having a mean effect close to zero.

Behavior of the moments of the distribution of genotypic values: Previous investigations of the ML procedure have shown that estimates of β and U tend to be strongly confounded with one another (Keightley 1994; Keightley and Ohnishi 1998). Likelihood often becomes flat if the fitted value of U is increased while, for constant α , β is simultaneously decreased. Since $E(a) = \beta/\alpha$ the estimated mean mutant effect also decreases as the fitted value of U increases. The correlation between the parameters generates characteristically shaped profile likelihoods. Profile likelihoods for one data set investigated in the previous section are shown in Figure 2. In Figure 2a (the likelihood profile for U), the ML is close to the U value of 2.5 simulated, but log likelihood quickly becomes flat as a function of increasing U . This behavior can be explained as follows: The moments of the distribution of genotypic values, X , are given by

$$E(X^n) = \sum_{i=0}^{\infty} p(i|U) \frac{i\beta(i\beta + 1) \dots (i\beta + n - 1)}{\alpha^n}. \quad (3)$$

The first four moments are therefore

$$E(X) = \frac{U\beta}{\alpha}, \quad (4)$$

$$E(X^2) = E^2(X) \left(1 + \frac{1}{U} + \frac{1}{U\beta} \right), \quad (5)$$

$$E(X^3) = E^3(X) \times \left(1 + \frac{3}{U} + \frac{3}{U\beta} + \frac{1}{U^2} + \frac{3}{U^2\beta} + \frac{2}{U^2\beta^2} \right), \quad (6)$$

$$E(X^4) = E^4(X) \left(1 + \frac{6}{U} + \frac{6}{U\beta} + \frac{7}{U^2} + \frac{18}{U^2\beta} + \frac{11}{U^2\beta^2} + \frac{1}{U^3} \right)$$

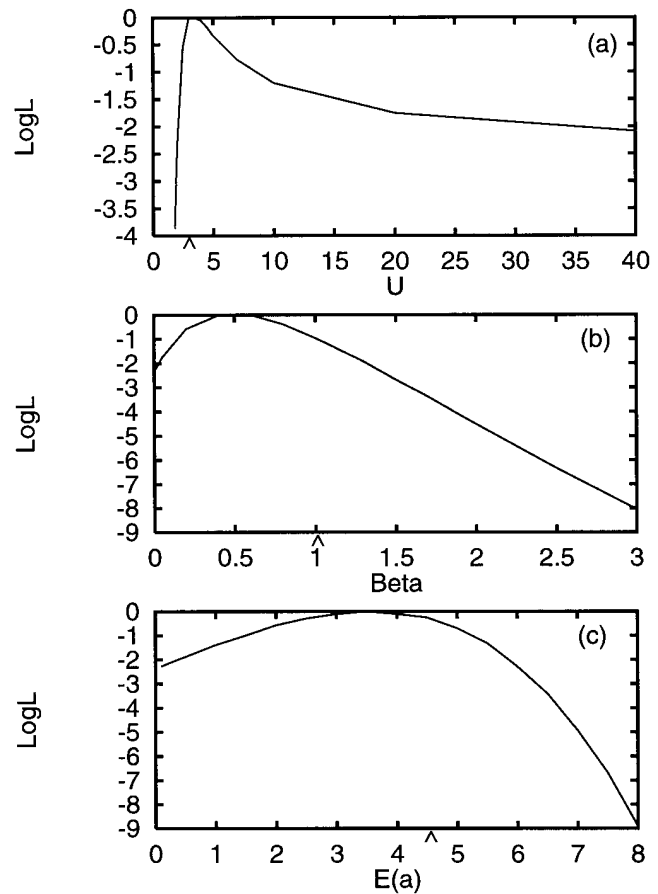


Figure 2.—Profile likelihoods as a function of U , β , and $E(a)$ for one of the favorable data sets. The stimulated parameter values are indicated by “ \wedge .”

$$+ \frac{6}{U^3\beta} + \frac{11}{U^3\beta^2} + \frac{6}{U^3\beta^3}). \quad (7)$$

Under ML the moments of the fitted distribution will be close to those of the empirical distribution. The mean of the fitted distribution, $E(X)$, is the most important constraint determining the fit. With a fixed $E(X)$ and a large value of U , the higher order moments can be held constant by increasing U while adjusting β downward in a compensatory manner. The important terms are $1/U\beta$, $2/U^2\beta^2$, $6/U^3\beta^3$, etc., in which U and B are of the same order in the denominator, while the terms where U is of higher order than β become small (*i.e.*, $1/U$, $1/U^2$, $3/U^2\beta$, etc.). The behavior of the moments of the genotypic distribution implies that increasing U can eventually always lead to an essentially unchanging fitted distribution, and to flat profile likelihoods.

With the favorable simulated data sets, there were always maxima in the likelihood profiles. However, less favorable situations, such as a strongly leptokurtic distribution of mutation effects, a large number of mutations per MA line, or poorly estimated genotypic values, often give no likelihood maxima, so only lower or upper support limits for the parameters can be given.

TABLE 5
Effect of changes in U and β on parameter value estimates

U	β	$E(a)$	\bar{U}			$\bar{\beta}$			$\widehat{E(a)}$		
			Min	ML	Max	Min	ML	Max	Min	ML	Max
0.5	0.5	3.65	0.22 (0.03)	0.95 (1.24)	$\rightarrow \infty$	0.056 (0.18)	0.52 (0.52)	2.00 (0.93)	0.39 (1.22)	3.27 (1.49)	7.52 (0.60)
	1.0	4.46	0.33 (0.09)	0.59 (0.27)	$\rightarrow \infty$	0.19 (0.19)	1.18 (0.64)	3.59 (1.76)	1.25 (1.16)	4.20 (1.25)	6.53 (1.11)
	2.0	5.17	0.40 (0.09)	0.65 (0.26)	$\rightarrow \infty$	0.45 (0.36)	1.66 (0.84)	4.44 (2.06)	2.25 (1.42)	4.50 (1.10)	6.26 (0.92)
2.5	0.5	1.63	1.33 (0.23)	3.11 (1.56)	$\rightarrow \infty$	0.23 (0.52)	0.71 (0.81)	$\rightarrow \infty$	0.21 (0.57)	1.65 (0.70)	3.09 (0.62)
	1.0	2.00	1.57 (0.26)	4.05 (2.75)	$\rightarrow \infty$	0.10 (0.25)	0.98 (0.98)	$\rightarrow \infty$	0.28 (0.61)	1.77 (0.62)	3.11 (0.39)
	2.0	2.30	1.85 (0.36)	3.38 (1.94)	$\rightarrow \infty$	0.45 (0.58)	2.37 ^a (2.32)	$\rightarrow \infty$	0.83 (0.84)	2.00 (0.70)	3.01 (0.48)
5.0	0.5	1.15	2.21 (0.29)	5.93 (4.05)	$\rightarrow \infty$	0.059 (0.18)	1.05 (1.45)	$\rightarrow \infty$	0.11 (0.33)	1.28 (0.52)	2.58 (0.37)
	1.0	1.41	2.87 (0.49)	7.83 (6.90)	$\rightarrow \infty$	0.18 (0.28)	1.70 ^b (2.34)	$\rightarrow \infty$	0.32 (0.44)	1.36 (0.63)	2.48 (0.45)
	2.0	1.63	3.24 (0.43)	7.70 (6.17)	$\rightarrow \infty$	0.16 (0.22)	2.98 ^a (3.52)	$\rightarrow \infty$	0.27 (0.38)	1.42 (0.50)	2.40 (0.27)

The variance of genotypic values was $20\sigma_g^2$. Means and standard deviations are from 10 simulation runs except for $U = 2.5$, $\beta = 1$, for which there were 30 runs.

^a One run gave an ML estimate $\rightarrow \infty$ and was not included.

^b Two runs gave ML estimates $\rightarrow \infty$ and were not included.

U unknown, more realistic data: To investigate ML with more realistic data, simulations in which $\sigma_g^2 = 20\sigma_e^2$ were analyzed. Mean ML estimates of U (Table 5) are of the correct order, and individually nonsignificantly different from the values simulated, but there appears to be a general upward bias in the estimated values. The bias is more serious than noted for simulations with U fixed. Individual simulation runs usually allow upper support limits for U to be obtained, but each set included at least one with an upper support limit $\rightarrow \infty$. Lower support limits for U tend to be closer to the simulated values for platykurtic distributions. Table 5 shows that information on the shape of the distribution of mutation effects is difficult to obtain. In some cases, likelihood increased monotonically as a function of β , hence the mean ML estimate $\rightarrow \infty$. In contrast to the other parameters, estimates of $E(a)$ seem to be unbiased, but individual runs give lower support limits for $E(a)$ of zero. Figures 3 and 4 compare results for mutation rates of 0.5 and 5, respectively, under a platykurtic distribution of mutation effects ($\beta = 2$). The lower U gives steeper likelihood profiles, so more information on the mutation parameters can be obtained. As U increases, the distribution of genotypic values will become increasingly normal, presumably increasing the difficulty in disentangling the parameters.

Trade-off between number of lines and precision for individual lines? To investigate the effect of varying the number of lines (N), while maintaining the same total

“effort” in the experiment, simulation runs were compared for constant $N \times \sigma_g^2$ (Table 6). Two levels of experimental precision were investigated: 50 or 200 lines with σ_g^2 set to 80 or 20, respectively (the upper two sets of parameters in Table 6), or the same numbers of lines with σ_g^2 set to 20 or 5, respectively (the lower two sets of parameters in Table 6). An exponential distribution of mutation effects with $U = 2.5$ was simulated. The results suggest that a higher number of lines with reduced effort per line leads to proportionally narrower bounds for $E(a)$, and fewer runs in which minimum, ML, or maximum parameter estimates were 0 or ∞ . Table 6 also illustrates that it is generally possible only to obtain lower support limits for U and β , and an upper limit for $E(a)$ (*cf.* Table 5).

DISCUSSION

The development of methods to infer distributions of mutation effects has been motivated by the crucial importance of these distributions for evolutionary models. This article has investigated ML as an approach to obtain estimates of U and mutation distribution parameters. One key parameter estimated by ML, but not estimated by the BM procedure, is a distribution shape parameter. The ML procedure behaves well if the number of mutation events per MA line is known, but if U has to be estimated simultaneously with the distribution parameters, the mutation parameters become strongly

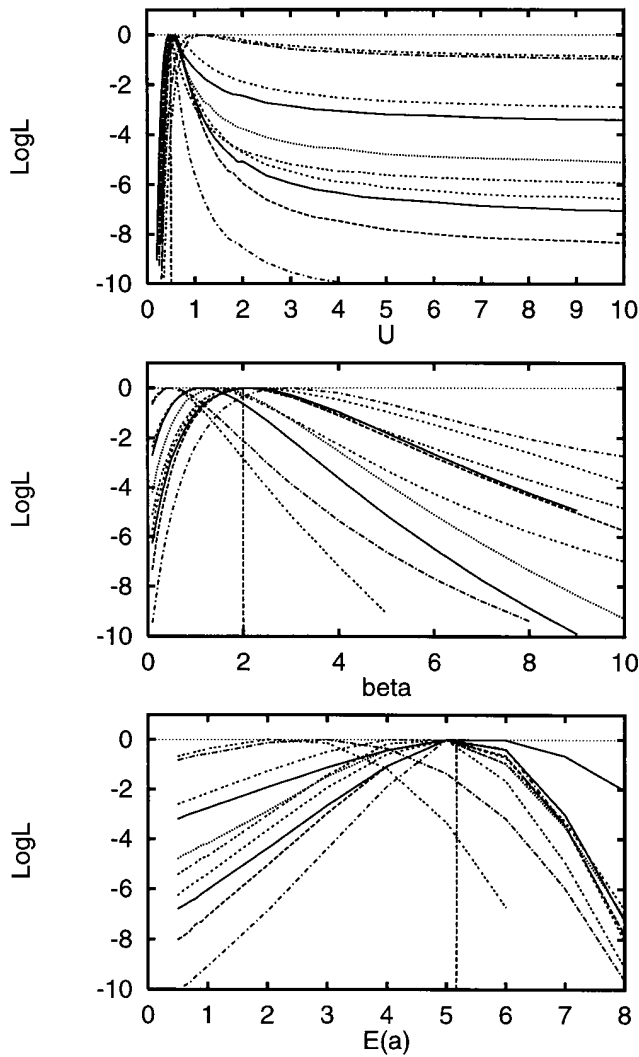


Figure 3.—Likelihood profiles as a function of U , β , and $E(a)$ for simulations with U set at 0.5 and β at 2. The value of α was adjusted such that the variance of genotypic values was $20\sigma_e^2$. The simulated parameter values are indicated by vertical broken lines.

confounded with one another, and profile likelihoods tend to be flat and asymmetrical about their maxima. Often, ML estimates for $U \rightarrow \infty$ and $\beta \rightarrow 0$. The flatness of the profile likelihoods can be explained by the behavior of the moments of the distribution of genotypic values. The moments contain terms in the product $\beta \times U$ that can be held constant by making compensatory changes in β and U . By assuming asymptotic properties of likelihood (*i.e.*, large sample size), lower support limits for U can be obtained (and sometimes upper limits as well), but the tendency for profile likelihoods to be asymmetrical suggests that the asymptotic properties are not being met. Further investigation of this aspect is needed.

By assuming a specific distribution of mutation effects with a fixed shape, much of the difficulty in disentangling the remaining parameters in the model disap-

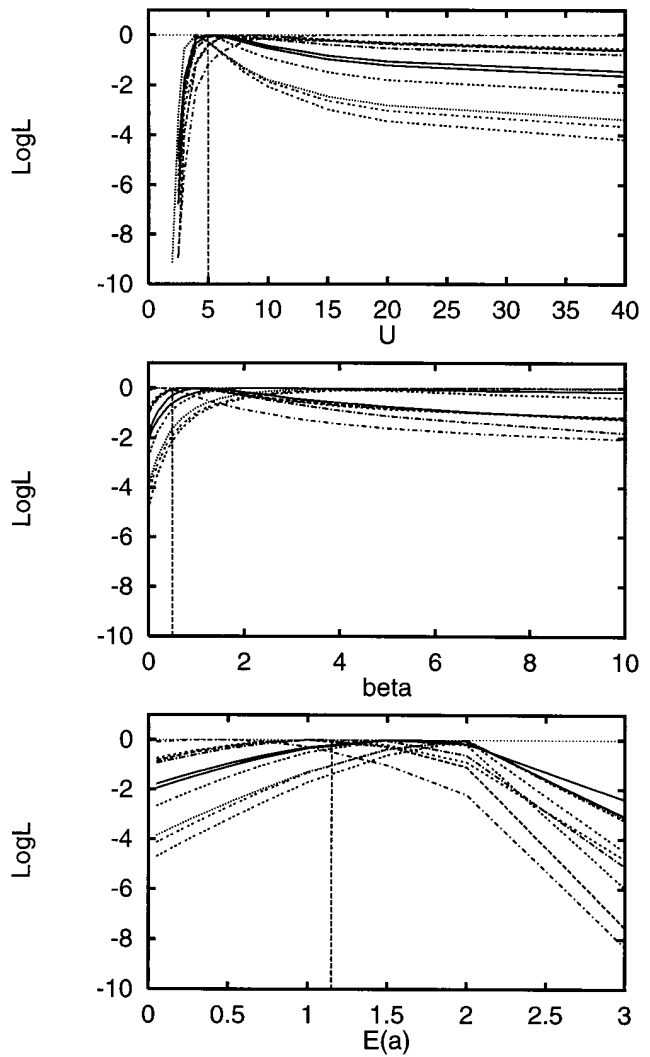


Figure 4.—As shown in Figure 3 with the simulated value of U set at 5 and β at 2.

pears, and point estimates for U and $E(a)$ can be obtained (Tables 1 and 2). An alternative to ML, the BM method, generally assumes that mutation effects are equal, and produces point estimates from the rate of change of mean and variance between MA and control lines. If the true distribution of mutation effects is gamma, the BM method underestimates U by a factor $1 + 1/\beta$, and overestimates $E(a)$ by the same factor:

$$U_{\text{BM}} = \Delta M^2 / V_m = U \left(1 + \frac{1}{\beta} \right), \quad (8)$$

$$E(a)_{\text{BM}} = V_m / \Delta M = E(a) \left(1 + \frac{1}{\beta} \right). \quad (9)$$

The BM method does not provide an obvious method to compare the fit of different distributions to the data, so the simplest model (*i.e.*, equal effects) needs to be assumed. By employing the BM method with bootstrapping to obtain a confidence interval (Houle *et al.* 1992), lower and upper limits for point estimates of U and $E(a)$

TABLE 6
Effect of change in number of lines measured (N) on parameter value estimates

Simulated values						Estimated values								
						\bar{U}			$\bar{\beta}$			$\widehat{E}(a)$		
N	U	β	α	$E(a)$	σ_g^2	Min	ML	Max	Min	ML	Max	Min	ML	Max
50	2.5	1.0	0.25	4.0	80	1.47 (0.36)	2.58 (1.14)	$\rightarrow \infty$ [3]	0.19 (0.38)	1.97 (1.71)	$\rightarrow \infty$ [10]	0.82 (1.35)	4.04 (1.85)	7.01 (1.59)
200	2.5	1.0	0.5	2.0	20	1.57 (0.26)	4.05 (2.75)	$\rightarrow \infty$ [19]	0.10 (0.25)	0.98 (0.98)	$\rightarrow \infty$ [1]	0.28 (0.61)	1.77 (0.62)	3.11 (0.39)
50	2.5	1.0	0.5	2.0	20	1.19 (0.33)	2.91 (2.62)	$\rightarrow \infty$ [16]	0.11 (0.31)	1.51 (1.79)	$\rightarrow \infty$ [2]	0.20 (0.57)	1.89 (1.25)	4.09 (0.90)
200	2.5	1.0	1.0	1.0	5	1.16 (0.27)	2.37 (1.22)	$\rightarrow \infty$ [24]	0.072 (0.21)	1.92 (1.77)	$\rightarrow \infty$ [14]	0.10 (0.28)	1.14 (0.48)	2.12 (0.35)
								[2]	[23]	[23]	[1]	[18]	[23]	

Standard deviations of estimates are shown in parentheses; number of runs (out of a total of 30) for which minimum or maximum support limits appear to occur at a boundary of 0 or ∞ , or ML estimates that occur at ∞ are shown in brackets. Runs for which the ML estimate $\rightarrow \infty$ were not included in the means.

can be obtained. These are asymptotically equivalent to the lower or upper support limits obtained by ML based on the properties of likelihood, if equal mutation effects are assumed. The lower (or upper) support limits for U and $E(a)$ obtained by ML are *not* equivalent in a statistical sense to the point estimates obtained by the BM method. With ML there may be no point estimate because likelihood often increases monotonically as a function of one parameter. A difficulty with the BM method to estimate U is the denominator of Equation 1, V_m , which could approach zero (hence $\widehat{U}_{BM} \rightarrow \infty$) if the experiment had low power. The problem is illustrated in Figure 5, where BM estimates and corresponding coefficients of variation (CVs) for U are shown for a range of simulated values. If V_m is poorly estimated (corresponding to the lowest simulated values of U in the figure), estimates can become meaningless because the $CV \rightarrow \infty$. A similar effect also occurs with BM estimation of $E(a)$, if ΔX could plausibly be zero (Equation 2). Other properties of estimation of mutation parameters by the BM method have been investigated by Deng and Fu (1998).

In principle, parameters that can be reliably estimated by the MA approach with inbred sublines are the rates of change of mean and variance from fixation of mutations (Crow and Simmons 1983). Two experiments have estimated rates of change of mean fitness in a recently caught outbred population of *Drosophila* maintained in the laboratory (Gilligan *et al.* 1997; Shabalina *et al.* 1997), and apparently produced conflicting results, but it has been argued that effects of genetic adaptation, inbreeding depression, and an accumulation of deleterious mutations cannot be distinguished in outbred populations (Keightley *et al.* 1998). The meaning of the

point (or minimum) estimates of U by the BM or ML methods must also be born in mind. They are conditional on the form of the distribution of mutation effects assumed, equal for BM, and gamma for ML in the present case. Mutations with tiny, but evolutionarily important effects, could make up a substantial fraction of the genomic deleterious mutation rate, but a laboratory experiment cannot hope to detect these by measuring

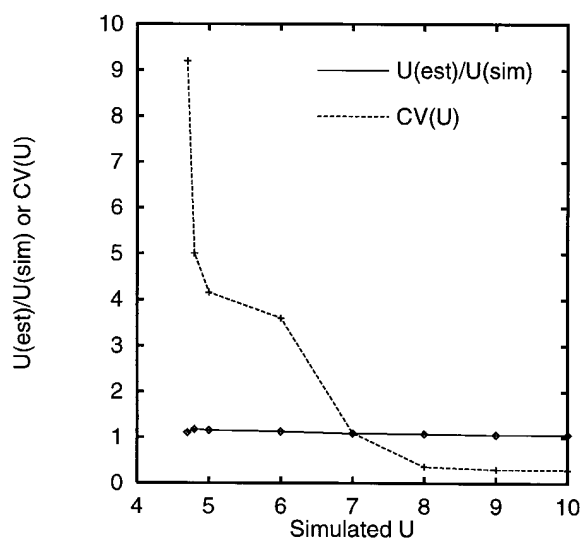


Figure 5.—Ratio of estimated to simulated mutation rate [$U(est)/U(sim)$] and coefficient of variation of estimate [$CV(U)$] plotted against simulated mutation rate. Estimates were obtained by the BM method from analysis of 100,000 simulated data sets of 100 MA and control lines, with five replicates in each line. The between-line variance was computed by an ANOVA. Mutant effects were 0.2, and the between-replicate environmental variance was 1.0.

their "pressure" on the population mean phenotype or their contribution to the mutational variance. The genomic mutation rate estimate, if it is a point estimate, must for this reason also be a minimum estimate.

The strong correlation between the parameter estimates implies that global maximization of likelihood is often problematical when all parameters are fitted simultaneously. This problem can be overcome by fixing one parameter and generating profile likelihoods. The properties of the multidimensional likelihood surface can be explored in this way, and it is recommended that this is done for analysis of real data. A further potential problem with very small simulated U values and high σ_g^2 relative to σ_e^2 (*i.e.*, very large mutation effects) is that likelihood profiles may have multiple peaks. The present version of the procedure makes use of information from a "base population" (generation 0) and from generation t . In principle, intermediate generations of data could also be utilized and may add greatly to the power to distinguish the different mutation events that occurred in the experiment. (Note that the present simulations suggest that fewer generations of mutation accumulation can give more precise estimates, but there will be an optimum, which will depend on the true mutation rate and distribution parameters.) T. Bataillon (personal communication) has devised an ML method assuming equal mutation effects to obtain estimates of U and $E(a)$ with more than one generation, and found increased power from doing so.

The confounded nature of the mutation distribution parameters can be largely overcome if the number of mutation events is known. Artificial induction of independent random TE insertions is one way to generate MA lines with known numbers of mutation events. Fitness or quantitative trait assays have allowed estimates to be obtained for mean fitness effects of insertion in *Drosophila* (Eanes *et al.* 1988; Mackay *et al.* 1992) and in *E. coli* (Elena *et al.* 1998). The latter study has shown that the distribution of fitness effects of single elements is discontinuous, with a preponderance of roughly equivalent effects of $\sim 3\%$. A discontinuous distribution of mutation effects was also inferred by reanalysis of mutation accumulation experiments involving balancers in *Drosophila*, and it was argued that TE mutations could be implicated (Keightley 1996).

Recently, minimum distance (MD) methods to infer U and mutation distribution parameters have been proposed in which data from a base population (generation 0) are not included in the analysis (Garcia-Dorado 1997). Instead, all the information comes from the MA lines themselves, with an estimate of the between-MA line variance coming from an ANOVA. Reanalysis of single generations of data of Mukai *et al.* (1972) and Ohnishi (1977), with the assumption of a gamma distribution of mutation effects, produced point estimates of U . These estimates, which did not take into account the large apparent changes of mean viability of the MA

lines, were about one order of magnitude smaller than the original BM estimates, which used information from the change of mean viability estimated by regression from several generations of data. However, an ML analysis of Ohnishi's (1977) data with the same model as described by Garcia-Dorado (1997) does not provide a point estimate; instead, the ML estimate of $U \rightarrow \infty$ (data not shown). This difference between ML and MD is probably due to the fact that ML takes into account variation in the true population mean, while MD does not. An ML analysis of data from Mukai *et al.* (1972) produces a similar point estimate of U as inferred by Garcia-Dorado (1997) if the same model is assumed [about 10 times smaller than by Mukai *et al.*'s (1972) original analysis], but again, the ML analysis gives an estimate of $U \rightarrow \infty$ if a mixed distribution of mutation effects (*e.g.*, gamma + normal) is assumed (data not shown). Garcia-Dorado's (1997) analysis of data from the MA experiment of Fernandez and Lopez-Fanjul (1996) included information on the ancestral population mean viability, so the estimates are, presumably, less sensitive to the model assumed.

The analysis presented here suggests that very large MA experiments can give some insight into genomic mutation rates and distributions of mutation effects. Small-scale experiments will have difficulty in detecting significant mutation-induced changes in mean or variance, and estimates of the underlying mutation parameters will therefore have little meaning. Analysis of the results of MA experiments by ML allows the fit of different distributions of effects to be compared, and some kinds of distributions may be rejected. ML lower bounds (or support limits) are appropriate minimum estimates of U . Estimates of distribution parameters are often unbounded. The critical problem is that the estimates depend on the model assumed for the distribution of mutation effects and may not account for mutations with small, but biologically important, effects. Analysis of DNA sequence data in which estimates of U are obtained by comparing the level of constraint in different parts of the genome (Kondrashov and Crow 1993) may ultimately be more informative, although a quantum leap in the understanding of the functional significance of noncoding DNA will be needed.

I thank Thomas Bataillon, Brian Charlesworth, Esther Davies, Jim Fry, Mike Lynch, Andy Peters, Stuart West, Alexey Kondrashov, and an anonymous reviewer for helpful comments, and the Royal Society for support.

LITERATURE CITED

- Bateman, A. J., 1959 The viability of near-normal irradiated chromosomes. *Int. J. Radiat. Biol.* **1**: 170-180.
 Charlesworth, B., D. Charlesworth and M. T. Morgan, 1990 Genetic loads and estimates of mutation rates in highly inbred plant populations. *Nature* **347**: 380-382.
 Crow, J. F., and M. J. Simmons, 1983 The mutation load in *Drosophila*, pp. 1-35 in *The Genetics and Biology of Drosophila*, Vol. 3C,

- edited by M. Ashburner, H. L. Carson and J. N. Thompson. Academic Press, London.
- Deng, H.-W., and M. Lynch, 1996 Estimation of deleterious mutation parameters in natural populations. *Genetics* **144**: 349–360.
- Deng, H.-W., and Y.-X. Fu, 1998 On the three methods for estimating deleterious mutation parameters. *Genet. Res.* **71**: 223–236.
- Drake, J. W., B. Charlesworth, D. Charlesworth and J. F. Crow, 1998 Rates of spontaneous mutation. *Genetics* **148**: 1667–1686.
- Eanes, W. F., C. Wesley, J. Hey, D. Houle and J. W. Ajioka, 1988 The fitness consequences of *P* element insertion in *Drosophila melanogaster*. *Genet. Res.* **52**: 17–26.
- Elena, S. F., and R. E. Lenski, 1997 Test of synergistic interactions among deleterious mutations in bacteria. *Nature* **390**: 395–398.
- Elena, S. F., L. Ekunwe, N. Hajela, S. A. Oden and R. E. Lenski, 1998 Distribution of fitness effects caused by random insertion mutations in *Escherichia coli*. *Genetica* **102/103**: 349–358.
- Falconer, D. S., and T. F. C. Mackay, 1996 *Introduction to Quantitative Genetics*, Ed. 4. Longman Scientific and Technical, Harlow, Essex, UK.
- Fernandez, J., and C. Lopez-Fanjul, 1996 Spontaneous mutational variances and covariances for fitness-related traits in *Drosophila melanogaster*. *Genetics* **143**: 829–837.
- Garcia-Dorado, A., 1997 The rate and effects distribution of viability mutation in *Drosophila*: minimum distance estimation. *Evolution* **51**: 1130–1139.
- Gilligan, D. M., L. M. Woodworth, M. E. Montgomery, D. A. Briscoe and R. Frankham, 1997 Is mutation accumulation a threat to the survival of endangered populations? *Conserv. Biol.* **11**: 1235–1241.
- Houle, D., D. K. Hoffmaster, S. Assimakopoulos and B. Charlesworth, 1992 The genomic mutation rate for fitness in *Drosophila*. *Nature* **359**: 58–60.
- Keightley, P. D., 1994 The distribution of mutation effects on viability in *Drosophila melanogaster*. *Genetics* **138**: 1315–1322.
- Keightley, P. D., 1996 Nature of deleterious mutation load in *Drosophila*. *Genetics* **144**: 1993–1999.
- Keightley, P. D., and A. Caballero, 1997 Genomic mutation rates for lifetime reproductive output and lifespan in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* **94**: 3823–3827.
- Keightley, P. D., and O. Ohnishi, 1998 EMS induced polygenic mutation rates for nine quantitative characters in *Drosophila melanogaster*. *Genetics* **148**: 753–766.
- Keightley, P. D., A. Caballero and A. Garcia-Dorado, 1998 Surviving under mutation pressure. *Curr. Biol.* **8**: R235–237.
- Kimura, M., 1983 *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge, UK.
- Kondrashov, A. S., 1993 Classification of hypotheses on the advantage of amphimixis. *J. Hered.* **84**: 372–387.
- Kondrashov, A. S., and J. F. Crow, 1993 A molecular approach to estimating the human deleterious mutation rate. *Hum. Mutat.* **2**: 229–234.
- Lande, R., 1995 Mutation and conservation. *Conserv. Biol.* **9**: 782–791.
- Lyman, R. F., F. Lawrence, S. V. Nuzhdin and T. F. C. Mackay, 1996 Effects of single *P*-element insertions on bristle number and viability in *Drosophila melanogaster*. *Genetics* **143**: 277–292.
- Lynch, M., J. Conery and R. Burger, 1995 Mutation accumulation and the extinction of small populations. *Am. Nat.* **146**: 489–518.
- Mackay, T. F. C., R. Lyman and M. S. Jackson, 1992 Effects of *P* element insertions on quantitative traits in *Drosophila melanogaster*. *Genetics* **130**: 315–332.
- Mukai, T., 1964 The genetic structure of natural populations of *Drosophila melanogaster*. I. Spontaneous mutation rate of polygenes controlling viability. *Genetics* **50**: 1–19.
- Mukai, T., S. I. Chigusa, L. E. Mettler and J. F. Crow, 1972 Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster*. *Genetics* **72**: 333–355.
- Ohnishi, O., 1974 Spontaneous and ethyl methanesulfonate-induced mutations controlling viability in *Drosophila melanogaster*. Ph.D. Thesis, University of Wisconsin.
- Ohnishi, O., 1977 Spontaneous and ethyl methanesulfonate-induced mutations controlling viability in *Drosophila melanogaster*. II. Homozygous effect of polygenic mutations. *Genetics* **87**: 529–545.
- Ohta, T., 1992 The nearly neutral theory of molecular evolution. *Annu. Rev. Ecol. Syst.* **23**: 263–286.
- Rose, M., 1982 Antagonistic pleiotropy, dominance, and genetic variation. *Heredity* **48**: 63–78.
- Rose, M., and B. Charlesworth, 1980 A test of evolutionary theories of senescence. *Nature* **287**: 141–142.
- Shabalina, S. A., L. Y. Yampolsky and A. S. Kondrashov, 1997 Rapid decline of fitness in panmictic populations of *Drosophila melanogaster* maintained under relaxed natural selection. *Proc. Natl. Acad. Sci. USA* **94**: 13034–13039.

Communicating editor: A. G. Clark

