The Distribution of Mutation Effects on Viability in *Drosophila melanogaster*

Peter D. Keightley

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

ABSTRACT

Parameters of continuous distributions of effects and rates of spontaneous mutation for relative viability in Drosophila are estimated by maximum likelihood from data of two published experiments on accumulation of mutations on protected second chromosomes. A model of equal mutant effects gives a poor fit to the data of the two experiments; higher likelihoods are obtained with leptokurtic distributions or for models in which there is more than one class of mutation effect. Minimum estimates of mutation rates (events per generation) at polygenes affecting viability on chromosome 2 are 0.14 and 0.068, but estimates are strongly confounded with other parameters in the model. Separate information on rates of molecular divergence between Drosophila species and from rates of movement of transposable elements is used to infer the overall genomic mutation rate in Drosophila, and the viability data are analyzed with mutation rate as a known parameter. If, for example, a mutation rate for chromosome 2 of 0.4 is assumed, maximum likelihood estimates of mean mutant effect on relative viability are 0.4% and 1%, but the majority of mutations have very much smaller effects than these values as distributions are highly leptokurtic. The methodology is applied to estimate viability effects of single *P* element insertional mutations. The mean effect per insertion is found to be higher, and their distribution is found to be less leptokurtic than for spontaneous mutations. The equilibrium genetic variance of viability predicted by a mutation-selection balance model with parameters estimated from the mutation accumulation experiments is similar to laboratory estimates of genetic variance of viability from natural populations of Drosophila.

IN an influential experiment, MUKAI (1964) was the first to undertake a large scale investigation of rates and effects of spontaneous mutations on viability in *Drosophila melanogaster*. MUKAI obtained a minimum estimate for the total mutation rate for polygenic effects on viability, and an upper bound for the mean mutant effect, and concluded that the overall rate of appearance of viability mutations of small effect is at least an order of magnitude higher than the rate for recessive lethals. These important conclusions were confirmed in two later experiments following similar designs (MUKAI *et al.* 1972; OHNISHI 1977). In the basic procedure (reviewed by SIMMONS and CROW 1977), a lethal-free second chromosome is extracted from a natural population and mutations are allowed to accumulate for a number of generations in replicates (sublines) in conditions where natural selection is minimized. The chromosome is protected in the heterozygous state by maintaining it against a marked balancer chromosome. Recombination is avoided due to the presence of the balancer and because the chromosome is always transmitted through males. The viability of the chromosome in the homozygous state is measured according to a method devised by WALLACE (1956) by crossing individuals from the subline *inter se* and counting the number of wild type progeny relative to the number of progeny heterozygous for the balancer (the balancer is lethal in the homozygous condition). Viability is expressed as an index, *e.g.*, as the ratio of the number of wild-type progeny emerging to the total number of progeny (MUKAI 1964). The viability is standardized by dividing by the viability index of individuals in the base population. If additivity between loci is assumed, this relative viability index can be approximated by $1 - \Sigma a_{Hi}(1 - h_i)$, where the summation is over mutation events, a_{Hi}, is the proportional change in viability of the homozygote carrying mutation i, and $a_{Hi}h_i$ is the proportional change in viability of the heterozygote (CROW and SIMMONS 1983). The behavior of relative viability as a function of time is characterized by: (1) the accumulation of lethals at a rate of about 0.006 per second chromosome per generation, (2) a decline in mean relative viability in non-lethal carrying second chromosomes at a rate in the region of 0.3% per generation, (3) an increase in variance among these chromosomes and (4) the slow accumulation of "semilethal" chromosomes with viabilities far below the majority of the "quasinormal" chromosomes.

If a model of equal mutant effects is assumed, genotypic values have the same distribution as that of numbers of mutations, namely Poisson, and the change of mean and variance provide information for inference of mutation rate and mean mutant effect (MUKAI 1964). However, any variation in effects of mutants inflates the genotypic variance relative to the change in genotypic mean, and as a consequence the mutation rate and mean mutant effect are minimum and maximum

estimates respectively. The distribution of relative viabilities of chromosomes suggest an underlying leptokurtic distribution of mutant effects. MUKAI *et al.* (1972) suggested an exponential distribution of mutant effects, in which case the estimate of the mutation rate increases by a factor of about 2 compared to equal effects. Inferences of minimum mutation rate and maximum mean gene effect under the traditional model of equal gene effects has been recently carried out for accumulated mutants on total fitness (HOULE *et al.* 1992).

The distribution of effects of mutations is a key parameter in many models involving the maintenance or dynamics of quantitative genetic variation (KEIGHTLEY and HILL 1988). Here, a method is developed to infer mutation rates and parameters of the distribution of mutant effects. It is assumed that the distribution of mutant effects is continuous. The analysis is by maximum likelihood based on the observed distribution of line means following mutation accumulation, and allows support limits to be obtained for such quantities as the mean mutant effect, kurtosis of the distribution of mutant effects, and the mutation rate. Data from two of the mutation accumulation experiments are reanalyzed. The results of the mutation accumulation experiments are also analyzed under the assumption that the mutation rate is a known parameter, based on alternative information on the total genomic mutation rate in Drosophila from the rate of DNA base pair substitution and rates of movement of transposable elements. Finally, the distribution of effects of *P* element insertional mutations on viability is compared to that of spontaneous mutations, using data on effects of multiple *P* element insertions of MACKAY *et al.* (1992).

DESCRIPTION OF THE DATA ON SPONTANEOUS MUTATION ACCUMULATION

Published data from the spontaneous mutation accumulation experiments of MUKAI *et al.* (1972) and OHNISHI (1974) were used in the analysis. Both involved the maintenance of sets of protected chromosomes for 40 generations. The present analysis uses data from the last available generation in both cases. A greater amount of information would be extracted by simultaneous analysis of the data from all available generations, but this was not possible because the line numbers across the different generations were not available (*i.e.*, the links between the generations were missing). The results from separate analysis of the earlier generations did not appear to be qualitatively different from generation 40 (data not shown). Details of the data sets are as follows.

MUKAI *et al.* (1972): The data at generation 40 consists of relative viabilities of 43 sublines from one of the three groups of lines (CH) group, and were taken from a histogram in the paper. Summary statistics from two other groups of lines were given, but these could not be

used, because relative viabilities of each subline are needed for the present analysis.

OHNISHI (1974): The data are relative viabilities, again from a histogram, of 106 spontaneous mutation lines at generation 40. Summary statistics from these lines and others treated with the chemical mutagen ethyl methanesulfonate are given in OHNISHI (1977).

Lines with lethals: Lines with relative viability of less than 10% of the control were designated lethals in both of the above studies, and precise estimates of viability were not given. Lethals appear to represent a true discontinuity in the distribution of mutant effects. Following previous analyses, such lines were excluded.

MODEL

Mutant effects: The number of new mutations appearing per chromosome 2 per generation was assumed to be Poisson distributed with parameter λ. Mutations were assumed to accumulate in each chromosome independently in a time-invariant manner, to act additively between loci, and selection on the mutants in the generation in which they appeared was assumed to be weak. Mutants were assumed unconditionally to reduce relative viability. The reduction of relative viability between the mutant homozygote and the heterozygote at a locus was a. The main requirements for modeling the distribution of mutation effects were that the distribution should have few parameters, and that its shape could be varied over a wide range. The gamma distribution was chosen as it has only two parameters, changes in the values of which produce distributions with a broad range of characteristics. The density function for the gamma distribution is

$$g(a) = \alpha^\beta a^{\beta-1} e^{-\alpha a}/\Gamma(\beta), \qquad 0 < a < \infty, \qquad (1)$$

where $\Gamma(\ )$ is the gamma function and α and β are scale and shape parameters, respectively, of the distribution. The moments, k, of the distribution are $E(a^k) = \beta(\beta + 1) \ldots (\beta + k - 1)/\alpha^k$. The mean of the distribution is therefore $E(a) = \beta/\alpha$, and $\gamma_2 = E(a^4)/E^2(a^2) = (\beta + 2)(\beta + 3)/[\beta(\beta + 1)]$ is a parameter that describes the kurtosis of the distribution. So for example, $\beta \to \infty$ is the limiting case for all effects being equal, in which case $\gamma_2 = 1$. Conversely, as $\beta \to 0$, $\gamma_2 \to \infty$, and the distribution becomes increasingly leptokurtic.

METHOD

Likelihood analysis: Let $Z_{i,t}$ be the estimated phenotypic value (*i.e.*, relative viability) of chromosome i at generation t. Assume that this value is the sum of effects of some number of mutations plus a sampling effect, assumed to be normally distributed with mean μ and variance σ_E^2. Let $p(j \mid \lambda t)$ be the probability function for the Poisson distribution with parameter λt for j events, and let $f(x \mid \mu, \sigma_E^2)$ be the density of the normal

distribution at point x. The relative likelihood of $Z_{i,t}$ is

$$L(Z_{i,t}) = p(0 \mid \lambda t) f(Z_{i,t} \mid \mu, \sigma_E^2) + \sum_{j=1}^{\infty} p(j \mid \lambda t)$$

$$\times \int \int \cdots \int f(Z_{i,t} + a_1 + a_2 \cdots + a_j \mid \mu, \sigma_E^2) g(a_1) g(a_2) \qquad (2)$$

$$\cdots g(a_j) \, da_1 \, da_2 \cdots da_j.$$

The first term is the likelihood for zero mutations (genotypic value zero). This is the height of the normal distribution with mean μ and variance σ_E^2 at point $Z_{i,t}$, weighted by the probability of zero mutations. Terms inside the summation are weighted by the probability of occurrence of j mutations, and are integrals in j dimensions over the gamma distribution.

It was not practical to evaluate (2) directly, so Monte Carlo methods were used. A random integer, j, was drawn from the Poisson distribution with parameter λt. A genotypic value, X, was simulated by adding together j random deviates from the gamma distribution with parameters α and β. gamma deviates with arbitrary shape parameter β were generated by algorithms of AHRENS and DIETER (1974, 1982). The relative likelihood of observation $Z_{i,t}$ given a genotypic value X was

$$L(Z_{i,t} \mid X) = f(Z_{i,t} + X \mid \mu, \sigma_E^2). \qquad (3)$$

The relative likelihood of observation $Z_{i,t}$ was taken as the average of many replicates (typically 5×10^6) of (3) for independent X. The overall log likelihood of the data was the sum of log likelihoods of the independent observations in the data set. Evaluation of likelihood was highly demanding of computer time, at the time of writing.

Maximization of likelihood: The three basic parameters with respect to which the likelihood function was maximized were λ, $E(a)$ and γ_2. The downhill simplex method (NELDER and MEAD 1965; PRESS et al. 1992) was used for this purpose. The main interest was in the evaluation of the likelihood over a range of fixed values of one parameter, but maximized with respect to the remainder ("profile likelihoods"). Support limits with respect to the fixed parameter were obtained by linear interpolation from differences in Log likelihood from the maximum over the profile. The fitted parameter value corresponding to a change in natural log likelihood of 2 was taken as the support limit, which is asymptotically equivalent to a 95% confidence limit.

Estimation of population mean and error variance: OHNISHI (1974) estimated relative viabilities in the population of chromosomes lines at generation $t = 1$, for which it can reasonably be assumed that the estimated viabilities were independent from viabilities measured in subsequent generations. In this case data from $t = 1$ were included with, but treated as independent of, the

generation's data under analysis. Their inclusion provided information for simultaneous estimation of μ and σ_E^2. Generation 10 was the earliest for which relative viabilities were available in MUKAI et al. (1972), and the above assumption of independence from subsequent generations was unreasonable. The base population mean and sampling variance were estimated from the intercepts of regression lines fitted to the estimated mean and variance of relative viability respectively at generations 10, 20, 30 and 40. Note that in this case μ and σ_E^2 were assumed to be known with certainty, and were not estimated simultaneously with the other parameters, as above.

Use of information from more than one generation: Although data were not available, it is straightforward to extend the Monte Carlo method for evaluation of likelihoods (Equation 3) to the simultaneous analysis of data from more than one time point. If, say, there are two time points t_1 and t_2, the relative likelihood of observations $Z_{i,t1}$ and $Z_{i,t2}$ from the same subline i is the average of many evaluations of:

$$L(Z_{i,t_1}, Z_{i,t_2}) = f(Z_{i,t_1} + X_1 \mid \mu, \sigma_E^2) f(Z_{i,t_2} + X_1 + X_2 \mid \mu, \sigma_E^2),$$

where X_1 and X_2 are sums of j_1 and j_2 random gamma deviates, and j_1 and j_2 are random integers sampled from Poisson distributions with parameters λt_1 and $\lambda(t_2 - t_1)$ respectively.

Distribution of effects of P element insertional mutations: Parameters of the distribution of homozygous effects of single P element insertions on relative viability were inferred by similar methods to the above from data on estimated viabilities of lines containing multiple P element insertions (MACKAY et al. 1992). In this case the number of mutations is the number of insertion events (measured by in situ hybridization of a P specific probe to polytene chromosomes). The likelihood of observation i with phenotypic value Z_i, given that there are j P element insertions is

$$L(Z_i) = \int \int \cdots \int f(Z_i + a_1 + a_2 \cdots + a_j \mid \mu, \sigma_E^2)$$
$$\times g(a_1) g(a_2) \cdots g(a_j) \, da_1 \, da_2 \cdots da_j, \qquad (4)$$

where in this case a_j is the homozygous effect of insertion on relative viability. In practice, likelihoods were evaluated by Monte Carlo methods (Equation 3). The population mean and variance were estimated along with the mutation distribution parameters, and lines with zero insertions were included in the analysis.

RESULTS

Comparison with original estimates: MUKAI et al. (1972) obtained minimum estimates of mutation rates for cases of equal and exponentially distributed mutant effects, based on the observed rate of change of mean and variance of relative viability. Mutation rate estimates
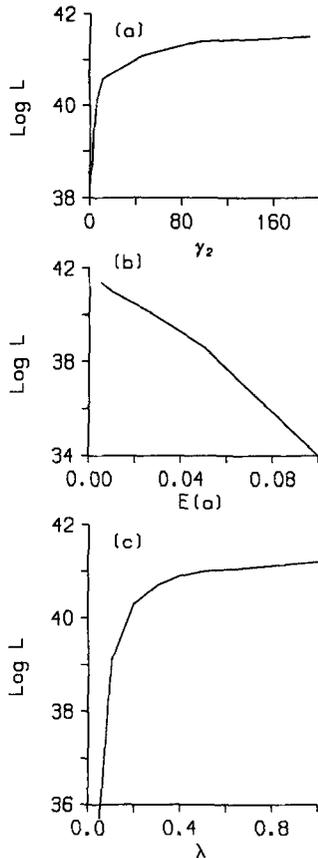
FIGURE 1.—Profile of natural Log likelihood of data as a function of parameters of the model for data of MUKAI et al. (1972). (a) γ_2; (b) $E(a)$; (c) λ.

for the group of lines in question (CH) were 0.097 (equal effects) and 0.194 (exponential). Maximum likelihood (ML) estimates with the present analysis are in good agreement: 0.087 (equal effects, $\gamma_2 = 1$) and 0.18 (exponential, $\gamma_2 = 6$). OHNISHI's (1977) estimate of the mutation rate for equal effects was 0.020, while the ML estimate is 0.025, also in good agreement.

Global maximum likelihood: A distribution of mutant effects with infinitely high kurtosis and with mean approaching zero, and an infinitely high mutation rate appears to give the highest likelihood for data of both MUKAI et al (1972) and OHNISHI (1974). However, there is other information on mutation rates in Drosophila. Consequently, it is possible to infer both upper and lower bounds for γ_2 and $E(a)$. This will be discussed in a later section.

Profile likelihoods: Natural Log likelihood of the two data sets as functions of γ_2, $E(a)$, and λ are shown in Figures 1 and 2. Note that likelihood approaches limits in all cases (see above). The two sets of graphs show how the three parameters are confounded with one another in the model. The likelihoods become flat with increasing mutation rate, increasing kurtosis of the distribution and decreasing average effect of mutants. A model of equal mutant effects ($\gamma_2 = 1$) fits both sets of data poorly,
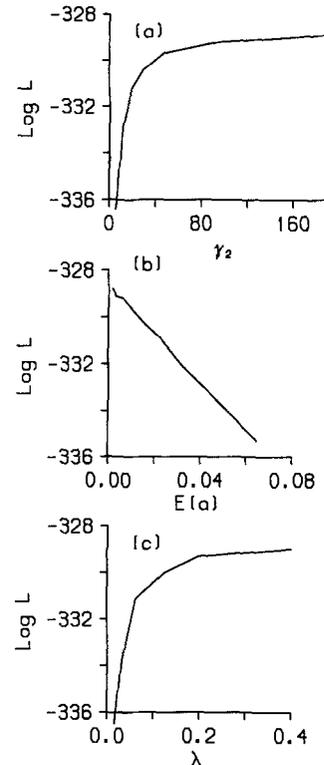

FIGURE 2.—Profile of log likelihood of data as a function of model parameters for data of OHNISHI (1974). (a) γ_2; (b) $E(a)$; (c) λ.

however. The likelihood ratios for this model are MUKAI et al. (1972): log L_{MAX} - log L = 3.3, so the ML is $e^{3.3}$ = 27 times more likely; OHNISHI (1974): log L_{MAX} - log L = 23.5, so the ML is $e^{23.5}$ = 1.6 × 10^{10} times more likely. Support limits for the three parameters based on differences in log likelihood from the asymptotic values are shown in Table 1. The lower limits of γ_2 imply at least a mildly leptokurtic distribution for MUKAI et al. (1972), and an extremely leptokurtic distribution for OHNISHI (1974). Density functions with parameters corresponding to the lower limit estimates of γ_2 and upper limit estimates of $E(a)$ are shown in Figure 3. The precision of the estimates from OHNISHI (1974) is higher than from MUKAI et al. (1972) because the former has more than twice as many observations, and smaller error variance (1.1 ×10^{-3} compared to 1.8 ×10^{-3}). Furthermore, OHNISHI's data set contains many more "semilethal" chromosomes. Note that this is reflected in the steepness of the likelihoods.

Fixed mutation rate: The mutation rate parameter is the total number of events per generation with any effect on the trait. Direct estimates of λ are not available (with the exception of rates of transposable element insertion, see below). However, data on rates of nucleotide substitution between Drosophila taxa allow indirect inference of the mutation rate per base pair per generation (see also KONDRASHOV and TURELLI 1992). The D. melanogaster genome has ca. 1.7 × 10^8 base pairs (ASHBURNER

TABLE 1

Lower and upper support limits for $\gamma_2 = E(a^4)/E^2(a^2)$, mean mutant effect and the mutation rate

| Parameter	Experiment	Support limits	
		Lower	Upper
γ_2	MUKAI et al. (1972)	3.8	$\to \infty$
	OHNISHI (1974)	23	$\to \infty$
$E(a)$	MUKAI et al. (1972)	$\to 0$	0.034
	OHNISHI (1974)	$\to 0$	0.020
λ	MUKAI et al. (1972)	0.14	$\to \infty$
	OHNISHI (1974)	0.068	$\to \infty$

FIGURE 3.—Plots of gamma distributions with parameters corresponding to the lower support limit for γ_2 and the upper support limit for $E(a)$ (Table 1).

1989). An estimate of the substitution rate at silent sites showing a low level of constraint is 16×10^{-9} per year (SHARP and LI 1989), but this may underestimate the mutation rate if purifying selection operates at these sites. Ignoring this source of bias, an estimate of the genomic mutation rate per year for base substitutions is 2.72. Similar rates can be inferred from molecular divergence data in Hawaiian Drosophila (ROWAN and HUNT 1991). If it is assumed that Drosophila goes through 5 generations per year in nature, the genomic mutation rate per generation for single base pair changes based on the above figures would be about 0.5. To this needs to be added the rate for events other than base substitutions. Rates of insertion and excision of transposable elements (TEs) are appreciable. EGGLESTON et al. (1988) measured the rate of excision and insertion on the X chromosome for 19 TE families in 18 sublines. A total transposition rate of 0.036 per generation can be inferred (the average of rates for dysgenic and non-dysgenic lines, excluding transposition of P elements in dysgenic lines). The D. melanogaster genome has a total of about 50 families of TEs (FINNEGAN 1992), and the X chromosome constitutes 20% of the genome, so the total genomic mutation rate for TEs is estimated to be 0.5. HARADA et al. (1990) measured insertion and excision rates of 4 TE families in 70 protected second chromosome lines for 400 generations. The overall rate per generation in these families was 0.007, so a rate of 0.22 can be inferred for 50 families in the whole genome. More recently, NUZHDIN and MACKAY (1994) measured the rate of transposition of 5 families of elements in initially inbred lines, and by similar calculations obtained a genomic rate of 0.6. Although estimates of transposition rates vary, figures are of the same order as the the spontaneous mutation rate for base substitutions. A reasonable estimate of the total genomic mutation rate is therefore about 1, or 0.4 for chromosome 2.

ML estimates of γ_2 and $E(a)$ along with support limits are shown in Table 2 for the above value of λ (chromosome 2 only) and values on either side of it. The estimates of mean mutant effect are smaller than in the original papers of MUKAI et al. (1972) and OHNISHI
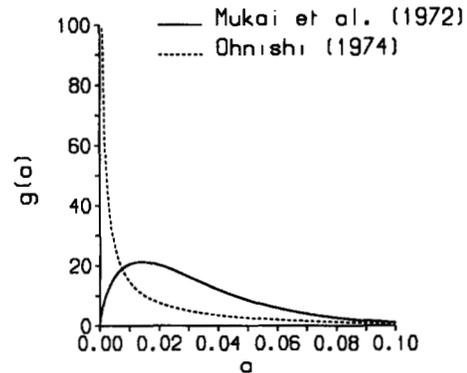
(1974) (which assumed equal gene effects and did not include prior information on λ). ML estimates of γ_2 imply extremely leptokurtic distributions of mutant effects in both cases. Gamma distributions with parameters corresponding to the ML under the assumption that $\lambda = 0.4$ are shown in Figure 4.

Dependence of parameter estimates on the details of the distribution: To investigate the degree to which the moments of the distribution of mutant effects depend on the precise details of the distribution assumed, an alternative model was explored in which there were two discrete classes of mutant effect, a_1 and a_2, and each class had a different mutation rate, λ_1 and λ_2. There were therefore four parameters whose values could vary in the model, plus the base population mean and variance where applicable, and likelihood was maximized with respect to these. Attempts to maximize likelihood with respect to more than two classes of effect were unsuccessful. ML estimates of a_1, a_2, λ_1 and λ_2 for a fixed overall mutation rate for chromosome 2 of 0.4 ($\lambda = \lambda_1 + \lambda_2 = 0.4$), along with resulting parameter estimates and log likelihoods are compared to ML estimates for the gamma distribution in Table 3. For both data sets, the two effects model which gives the best fit to the observations involves a large proportion of mutations with small effects and a small proportion with large effects. The estimate of mean mutant effect is fairly independent of the model. The two effects model gives lower estimates of γ_2 compared to the gamma distribution, but the likelihoods for the continuous distribution are considerably higher.

Simulated data: The behavior of the estimation procedure was investigated by analyzing sets of simulated data. Profile likelihoods with respect to λt for 5 independent data sets of 100 sublines are shown in Figure 5. In contrast to the experimental data (Figures 1 and 2) the ML estimates of λt are finite, and appear to center around the simulated value. Note that in the simulation, gene effects were large relative to the environmental variance.

ML estimates of mean mutant effect and γ_2 along with support limits for a range of fixed values of the mutation rate for chromosome 2

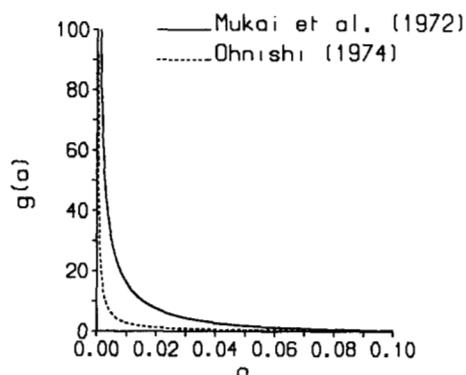| λ	Experiment	$E(a)$ (support limits)	γ_2 (support limits)
0.2	MUKAI et al. (1972)	0.022 (0.019–0.026)	8.5 (3.5–16)
	OHNISHI (1974)	0.0071 (0.0042–0.0098)	86 (46–150)
0.4	MUKAI et al. (1972)	0.011 (0.0092–0.013)	22 (11–40)
	OHNISHI (1974)	0.0039 (0.0024–0.0048)	180 (90–300)
0.8	MUKAI et al. (1972)	0.0057 (0.0047–0.0064)	49 (28–86)
	OHNISHI (1974)	0.0018 (0.0014–0.0025)	370 (200–620)

FIGURE 4.—Plots of gamma distributions with values of γ_2 and $E(a)$ corresponding to the ML estimates for $\lambda = 0.4$ (Table 2).

Distribution of effects of P element insertional mutations: Profile likelihoods with respect to $E(a)$ and γ_2 were computed as described above using estimates of relative viability for 94 Drosophila lines containing an average of 3.1 P element insertions (MACKAY et al. 1992). In this case, mutation effects were measured in the homozygous state relative to an insert-free control. In contrast to the distributions for spontaneous mutations, the ML estimate for γ_2 is 1, i.e., a model of equal viability effects of P elements fits the data best, but the upper support limit is $\gamma_2 = 26$, so a leptokurtic distribution cannot be ruled out. The ML estimate of $E(a)$ is 0.042, with lower and upper support limits of 0.026 and 0.057, respectively. This ML estimate of mean viability effect per homozygous insertion is actually lower than the estimates of the mean effect of spontaneous mutations if equal effects are assumed, but substantially higher if, e.g., the spontaneous mutation rate in the whole genome is assumed to be 1.

DISCUSSION

Distribution of mutant effects on viability: Data from two mutation accumulation experiments have been reanalyzed under the assumption of a continuous distribution of effects of mutations on viability. Lower support limits for the estimates of γ_2 (the ratio of the fourth moment to the square of the second moment of the distribution) are 3.8 and 23 for experiments of MUKAI et al. (1972) and OHNISHI (1974), respectively. The parameter estimates for the latter

are somewhat more precise because the error variance is smaller, and had data on more than twice as many sublines. In both cases a distribution with infinitely high kurtosis, mean mutant effect approaching zero, and a mutation rate approaching infinity gives the highest likelihood. As was pointed out in the original studies (e.g., MUKAI 1964), the mutation rate and distribution parameters are confounded with one other and it is only possible to obtain lower limits for e.g., the mutation rate.

An alternative approach is to assume a value for the mutation rate based on information external to the mutation accumulation experiment. Such information is now available from studies of rates of molecular evolution and movement of transposable elements. With the mutation rate for chromosome 2 fixed at 0.4 (a reasonable estimate from the available data), estimates of the mean mutant effect on relative viability are about 1% for MUKAI et al. (1972), and about 0.4% for OHNISHI (1974), and estimates of γ_2 are 22 and 180, respectively, which imply extremely leptokurtic distributions with the majority of the mutations having tiny effects.

The true distribution of mutant effects on viability is certain not to be gama, but the gamma distribution is useful for modeling purposes as it has useful properties. It is questionable, however, whether the distribution of viability effects is continuous, as it is likely to be a complex mixture of distributions. Mutation events fall into several classes, of which base substitutions and transpositions are probably the most important. Several classes of target sites for mutation events can also be recognized. In decreasing order of constraint, these might be listed: base pairs which lead to an amino acid replacement; upstream and downstream "control" sequences of genes; synonymous substitution sites of coding sequences; introns; intergenic "space"; pseudogenes. Transposition events anywhere in the genome appear to be deleterious, however (LANGLEY et al. 1988). It is possible that there is a class of mutations with almost zero effect on viability which represent a true discontinuity in the distribution. It would be possible to change the model to include such a class of mutation events by assuming a lower total genomic mutation rate. Note, however, that the total mutation rate at loci affecting viability on chromosome 2 excluding this class is not likely to be much less than 0.1 (see Table 1).

TABLE 3

ML estimates of mutation rates λ_1 and λ_2 for a total mutation rate of 0.4 and estimates of corresponding mutation effects a_1 and a_2 for models involving either one or two classes of mutant effect

Model	Data set	λ_1	λ_2	a_1	a_2	$E(a)$	γ_2	Log L
One effect	MUKAI et al.	0.4		0.012		0.012	1	18.2
Two effects		0.072	0.33	0.056	0.0010	0.011	5.5	38.4
gamma						0.011	22	41.0
One effect	OHNISHI	0.4		0.0048		0.0048	1	−420.3
Two effects		0.0087	0.39	0.11	0.0022	0.0045	44	−336.9
gamma						0.0039	180	−329.0

Resulting estimates of mean mutant effect and γ_2, and log likelihoods for the two data sets are also shown. ML estimates for the gamma distribution are shown for comparison.
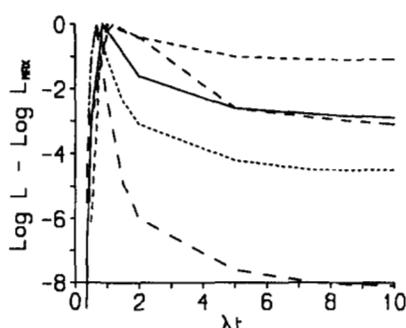

FIGURE 5.—Profile of Log likelihood as a function of λt for 5 sets of simulated data of 100 observations each. The parameters used to generate the observations were $\lambda t = 1$ and $E(a)/\sigma_E = 3.2$, with effects sampled from an exponential distribution.

It has been suggested that mutation accumulation experiments performed prior to knowledge of hybrid dysgenesis (KIDWELL et al. 1977) could have been subject to elevated mutation rates from transposition of P elements because the balancer strains are long-established laboratory stocks, while the mutation accumulation chromosomes are from the wild. However, such an argument does not seem tenable because it is implausible that such strong mutator activity would have gone unnoticed (CROW and SIMMONS 1983). Furthermore, a recent mutation accumulation experiment in which P element hybrid dysgenesis was deliberately avoided produced similar estimates for the rate of recessive lethal mutations (HOULE et al. 1992). It is a possible, however, that mobilization of other families of elements could generate new mutational variation in an experiment involving balancers.

Distribution of effects of P element insertions: The analysis of data on multiple P element insertions using data of MACKAY et al. (1992) provides evidence that the distribution of homozygous effects of single P element insertions is different to that of spontaneous mutations, because a model of equal mutant effects provides the best fit to the data. The ML estimate of the mean effect of insertion is somewhat lower than that for spontaneous mutations if equal effects are assumed, but very much larger if the spontaneous mutation rate for chromosome 2 is assumed to be 0.4.

The distributions of effects of single P element insertions on abdominal and sternopleural bristle number were also investigated. In this case, a gamma distribution reflected about $a = 0$ was assumed, and an additional parameter, the proportion P of the distribution greater than zero was estimated. The ML estimate of P was close to 0.5 for both types of bristles, implying a symmetrical distribution of effects, but the profile likelihood was very flat. Estimates of variances of effects of inserts were similar to those obtained by HILL (1992), who used a method of moments. ML estimates of γ_2 were 37 and 140 for abdominal and sternopleural bristle number, respectively, which imply rather more leptokurtic distributions than inferred by HILL (1992).

Variance maintained at mutation-selection balance: If it is assumed that viability equates with fitness, the variance of viability at equilibrium in a infinite population for a model in which mutant alleles are unconditionally deleterious with $E(a_H)$ the mean homozygous effect of mutants is $V_g = \lambda E(a_H)$. Under the assumption that the mutation rate for chromosome 2 is 0.4, and that therefore the genomic mutation rate is 1, and using estimates of $E(a)$ from Table 2 as values of $E(a_H)$, the predicted equilibrium variances are 1.1×10^{-2} for MUKAI et al. (1972) and 3.9×10^{-3} for OHNISHI (1974). Note that estimates of $E(a)$ in Table 2 refer to the difference in relative viability between the mutant homozygote and the heterozygote. It seems likely that viability mutations are partially recessive (MUKAI et al. 1972), so the predicted equilibrium variance would be somewhat higher. The predicted equilibrium variance is almost independent of the mutation rate because the assumption of e.g., a higher value for λ implies a lower value of $E(a)$ (Table 2). Estimates of the total genetic variance of viability estimated in the laboratory for natural populations of $D.$ melanogaster (MUKAI 1985) are of the same order as the predicted values (see CABALLERO and KEIGHTLEY (1994) who address this question in greater detail). Viability differences between natural Drosophila chromosomes account for about one-third of the differences in total fitness (SVED 1971, 1975; MACKAY 1985), so this factor would lead to a reduction in genetic variance predicted by the mutation-selection balance

model. The appearance of even a small number of mutants with deleterious effects on viability, but with beneficial effects on some other major fitness component could offset this, however.

LITERATURE CITED

AHRENS, J. H., and U. DIETER, 1974 Computer methods for sampling from gamma, beta, Poisson and binomial distributions. Computing **12:** 223–246.

AHRENS, J. H., and U. DIETER, 1982 Generating gamma variates by a modified rejection method. Commun. ACM **15:** 873–882.

ASHBURNER, M., 1989 *Drosophila: A Laboratory Manual.* Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

CABALLERO, A., and P. D. KEIGHTLEY, 1994 A pleiotropic nonadditive model of variation in quantitative traits. Genetics **138:** 883–900.

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.

EGGLESTON, W. B., D. M. JOHNSON-SCHLITZ and W. R. ENGELS, 1988 P-M hybrid dysgenesis does not mobilize other transposable element families in *D. melanogaster.* Nature **331:** 368–370

FINNEGAN, D. J., 1992 Transposable elements, pp. 1096–1107 in *The Genome of Drosophila Melanogaster,* edited by D. L. LINDSLEY and G. G. ZIMM. Academic Press, San Diego, Calif.

HARADA, K., K. YUKUHIRO and T. MUKAI, 1990 Transposition rates of movable genetic elements in *Drosophila melanogaster.* Proc. Natl. Acad. Sci. USA **87:** 3248–3252.

HILL, W. G., 1992 Inferences of moments of effects of individual inserts. Genetics **130:** 331–332.

HOULE, D., D. K. HOFFMASTER, S. ASSIMACOPOULOS and B. CHARLESWORTH, 1992 The genomic mutation rate for fitness in *Drosophila.* Nature **359:** 58–60.

KEIGHTLEY, P. D., and W. G. HILL, W. G., 1988 Quantitative variability maintained by mutation-stabilizing selection balance in finite populations. Genet. Res. **52:** 33–43.

KIDWELL, M. G., J. F. KIDWELL and J. A. SVED, 1977 Hybrid dysgenesis in *Drosophila melanogaster:* a syndrome of aberrant traits including mutation, sterility, and male recombination. Genetics **86:** 813–833.

KONDRASHOV, A. S., and M. TURELLI, 1992 Deleterious mutations, apparent stabilizing selection and the maintenance of quantitative variation. Genetics **132:** 603–618.

LANGLEY, C. H., E. A. MONTGOMERY, R. R. HUDSON, N. L. KAPLAN and B. CHARLESWORTH, 1988 On the role of unequal exchange in the containment of transposable element copy number. Genet. Res. **52:** 223–235.

MACKAY, T. F. C., 1985 A quantitative genetic analysis of fitness and its components in *Drosophila melanogaster.* Genet. Res. **47:** 59–70.

MACKAY, T. F. C., R. LYMAN and M. S. JACKSON, 1992 Effects of P elements 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., 1985 Experimental verification of the neutral theory, pp. 125–145 in *Population Genetics and Evolution,* edited by T. OHTA and K. AOKI. Japan Scientific Societies Press, Tokyo.

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.

NELDER, J. A., and R. MEAD, 1965 Comput. J. **7:** 308–313.

NUZHDIN, S. V., and T. F. C. MACKAY, 1994 Direct determination of retrotransposon transposition rates in *Drosophila melanogaster.* Genet. Res. **63:** 139–144.

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.

PRESS, W. H., S. A. TEUKOLSKY, W. T. VETTERLING and B. P. FLANNERY, 1992 *Numerical Recipes in C,* Ed. 2. Cambridge University Press, Cambridge.

ROWAN, R. G., and J. A. HUNT, 1991 Rates of DNA change and phylogeny from the DNA sequences of the alcohol dehydrogenase gene for five closely related species of Hawaiian *Drosophila.* Mol. Biol. Evol. **8:** 49–70.

SHARP, P. M., and W.-H. LI, 1989 On the rate of DNA sequence evolution in *Drosophila.* J. Mol. Evol. **28:** 398–402.

SIMMONS, M. J., and J. F. CROW, 1977 Mutations affecting fitness in *Drosophila* populations. Annu. Rev. Genet. **11:** 49–78.

SVED, J. A., 1971 An estimate of heterosis in *Drosophila melanogaster.* Genet. Res. **18:** 97–105.

SVED, J. A., 1975 Fitness of third chromosome homozygotes in *Drosophila melanogaster.* Genet. Res. **25:** 197–200.

WALLACE, B., 1956 Studies of irradiated populations of *Drosophila melanogaster.* J. Genet. **56:** 280–293.