SHORT PAPER

Analysis of mutation accumulation experiments: response to Deng, Li and Li

PETER D. KEIGHTLEY*

Institute of Cell, Animal and Population Biology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, Scotland, UK

(*Received 26 November 1998 and in revised form 16 February 1999*)

Summary

A recent paper in this journal by Deng, Li and Li has investigated methods to estimate rates and effects of polygenic mutations using data from mutation accumulation experiments. Here, I evaluate a number of critical points in this paper concerning a maximum likelihood (ML) procedure to analyse mutation accumulation data. I show that Deng, Li and Li's criticisms are based on misunderstandings, or numerical problems they encountered that could have been readily overcome. In Monte Carlo simulations, I show that ML can give a considerable increase in precision over the method of moments that is traditionally used to analyse mutation accumulation data. Furthermore, ML allows the comparison of the fit of different models for the distribution of mutation effects.

1. Introduction

There has been a recent upsurge of interest in the experimental estimation of genome-wide mutation rates and properties of polygenic mutations by mutation accumulation (MA). The standard design for a MA experiment is to allow mutations to accumulate at random is inbred lines. Mutations may also be induced, and bred to fixation. Rates of change of measurable properties of the phenotypic distribution of a trait are used to infer the genome-wide mutation rate (U) and properties of the distribution of their effects. There has also been interest in methods to analyse data from MA experiments. A recent paper (Deng *et al.*, 1998 – DLL98 henceforth) was an extensive investigation of the Bateman–Mukai (BM) analysis method (Bateman, 1959; Mukai, 1964). The BM method uses the rate of change of phenotypic mean and the rate of increase of among-line variance of the MA lines to estimate U and the average effect of a new mutation (s), under the assumption that mutations have equal effects. DLL98 performed a more limited investigation of a maximum likelihood (ML) method to infer mutation parameters (Keightley, 1994; Keightley & Ohnishi, 1998), in which models

with variable mutation effects can be assumed. DLL98 describe a number of problems and difficulties with the ML procedure, principally that it is not more informative than the BM approach, and that problems occur in detecting global likelihood maxima. The purpose of this short paper is to explain the nature of these problems, and to compare the BM and ML approaches.

2. Materials and methods

(i) *Simulation of MA experiments*

The simulation assumptions were the same as DLL98. Each simulated experiment is assumed to consist of a set of mutation-free control lines and a set of MA lines that have accumulated mutations for several generations. The MA lines each have an average of U fixed mutations, the actual number being sampled from a Poisson distribution. Phenotypic values for the analysis were line means, with environmental deviations normally distributed with variance σ_e^2, the environmental variance of line means. Mutation effects were assumed to be gamma distributed with location and shape parameters α and β respectively. The mean effect is $\bar{s} = \beta/\alpha$, and the variance is β/α^2. The gamma distribution was chosen as it can take a wide range of shapes by varying one parameter, although natural distributions of mutation effects

* Telephone: +44 (0)131 650 5443. Fax: +44 (0)131 650 6564. e-mail: p.keightley@ed.ac.uk.

could be more complex. Cases with equal mutation effects, corresponding to $\beta \to \infty$ (the model generally assumed with the BM analysis method), were also simulated.

(ii) *Analysis by the BM method*

The change in phenotypic mean between the MA and control lines, ΔM, and the increase in among-line variance σ_{b}^2 are used to infer U and s, under the assumption of equal mutation effects:

$$\hat{U} = \Delta M^2 / \sigma_{\mathrm{b}}^2, \tag{1}$$

$$\hat{s} = \sigma_{\mathrm{b}}^2 / \Delta M. \tag{2}$$

BM estimates of U (s) are often presented as minima (maxima), since variation in selective values will lead to underestimation (overestimation) of the true parameter values. However, they could only be minima (maxima) in an experiment with zero sampling error. The estimates are also often presented as model free, in contrast to ML for which a distribution of mutation effects is assumed. However, under BM there is no obvious way to compare the fit of different distributions of mutation effects, so the simplest model of equal effects needs to be assumed.

(iii) *Analysis by ML*

The method for analysis has been described elsewhere (Keightley & Ohnishi, 1998). Data were line means for MA and control lines, as for the BM method. The parameters estimated are M, σ_{e}^2, α, β and U. To simplify the interpretation of the results, U and $\bar{s}$ were estimated for each data set using a series of different models, corresponding to fixed values of β. Likelihood was maximized by the simplex method (Nelder & Mead, 1965). To ensure convergence to the global maximum likelihood, a starting value for U was obtained by maximizing likelihood for five fixed values of U ranging from $0 \cdot 25 \times$ to $4 \times$ the expected value, then using the U value which gave the highest likelihood as the starting value for the global maximization (within a fixed β model). Within the initial maximizations using a fixed U value, the starting value for α was set to its expected value, and starting values of M and σ_{e}^2 were calculated from the control line data. However, investigation of a subset of the runs showed that this initial line search approach seemed to be unnecessary, as the global maximum was reached as long as a fixed value for β was assumed, the starting values for the variable parameters were plausible, and the data reasonably informative. Convergence to the ML was checked by restarting the procedure after convergence had apparently occurred, until there was no further significant increase in likelihood.

It should be noted that assuming a normal distribution of environmental deviations could make ML more sensitive to departures from normality than the BM method. One way to alleviate this potential problem is to use line means in the calculations (as is done here), as these will follow a normal distribution more closely than replicate values (perhaps leading to a loss of some information). If the distribution of residuals is not normal, transformation of values to an appropriate scale may also be possible. For categorical data (e.g. dead or alive), an appropriate liability model, analysed via likelihood, may be the appropriate solution (T. A. Bataillon and P. D. Keightley, unpublished).

3. Results and discussion

(i) *BM and ML: comparison of equivalent models*

In their comparison of the BM and ML methods, DLL98 do not evaluate the procedures under equivalent models. Since there is no obvious way to compare the fit of different distributions of mutation effects under the BM approach, the simplest model of equal effects is assumed. This is therefore the appropriate model for comparing the performance of the BM and ML procedures. Table 1 shows mean estimates for U and s for a range of simulated values from the analysis of 1000 simulated data sets by the BM and ML procedures, assuming equal mutation effects, and σ_{e}^2 arbitrarily set to one. The absolute values of s simulated are not relevant here, as the population mean is not specified, and estimates of s for life history traits would normally be scaled by the population mean. The informativeness of an experiment depends on the number of lines and the ratio $\sigma_{\mathrm{b}}^2 / \sigma_{\mathrm{e}}^2$ (García-Dorado, 1997), so a range of plausible values of this ratio were simulated. The performance of the procedures can be evaluated in two ways: by comparing the mean estimates of U and s or by comparing the variances among the estimates. Table 1 shows that both methods give mean estimates close to the simulated values over the range of parameter values simulated. However, if $\sigma_{\mathrm{b}}^2 / \sigma_{\mathrm{e}}^2$ is small, both methods show a small but appreciable upward bias. Under ML, this is due to a few data sets for which likelihood is extremely flat as a function of U, and occurs if the data do not contain sufficient information to reliably distinguish the model parameters. If an experiment is noisy ($\sigma_{\mathrm{b}}^2 / \sigma_{\mathrm{e}}^2 < \sim 1$), likelihood can go on increasing with increasing U, and therefore give infinite sampling variance. A similar problem also occurs with the BM method if the denominator in (1) or (2) can approach zero, and explains the extremely high sampling variance in the case of $U = 0 \cdot 2$, $s = 2 \cdot 24$ (Table 1).

Since mean estimated values are close to those simulated, the precision of the methods can be compared from their estimation variances. Table 1

Table 1. *Comparison of means and variances (VAR) of estimates from the BM and ML procedures under a model of equal mutation effects*

Simulated values			Estimated values							
			BM		ML		BM		ML	
σ_b^2/σ_e^2	U	s	$\bar{U}$	VAR	$\bar{U}$	VAR	$\bar{s}$	VAR	$\bar{s}$	VAR
1	0·2	2·24	0·24	0·045	0·24	0·014	3·08	21·1	2·19	0·38
	1·0	1·0	1·17	0·69	1·17	0·49	1·01	0·13	1·00	0·11
	5·0	0·447	5·87	10·9	5·87	10·8	0·44	0·022	0·44	0·020
5	0·2	7·07	0·21	0·0051	0·20	0·0021	7·23	1·84	7·07	0·042
	1·0	2·24	1·04	0·065	1·06	0·057	2·23	0·17	2·20	0·13
	5	1·0	5·12	0·96	5·17	0·98	1·01	0·031	1·00	0·030
20	0·2	14·14	0·21	0·0034	0·20	0·0020	14·1	4·13	14·1	0·051
	1·0	4·47	1·01	0·037	1·00	0·010	4·51	0·48	4·46	0·013
	5	2·0	5·08	0·67	5·17	0·62	2·01	0·091	1·97	0·078

Results shown are based on 1000 replicates with 100 MA and 100 control lines.

shows that ML can provide a worthwhile increase in precision for estimation of U, and a considerable increase for estimation of s, compared with BM. The benefit is greatest for cases with relatively few mutations per line or large σ_b^2/σ_e^2 relative to s. If there are many mutations per line or the experiment is noisy, data will, presumably, be close to normally distributed, and the information that can be gleaned by ML comes almost exclusively from changes of mean and between-line variance.

(ii) *Can ML distinguish between distributions of mutation effects?*

In their evaluation of the ML procedure DLL98 state that 'Keightley's method *cannot* estimate U, and all the distribution parameters α and β *simultaneously* and *individually* from M-A data…One parameter must be assumed in order to estimate the other parameters…Therefore, contrary to the general belief, Keightley's method (1994) *does not* yield estimates on more parameters from M-A data about deleterious genomic mutations than Bateman–Mukai's method of moments.' To evaluate the validity of this claim, simulated MA experiments with mutation effects either from a platykurtic ($\beta = 4$) or a leptokurtic ($\beta = 0.5$) distribution are analysed for a series of models with fixed values of β including the case of equal effects ($\beta \to \infty$). On average, the model giving the highest log likelihood corresponds with the model simulated (Table 2). If the true distribution is moderately platykurtic ($\beta = 4$), equal mutation effects and distributions much more leptokurtic than the exponential distribution ($\beta = 1$) can be excluded, on average. If the true distribution is leptokurtic, distributions more platykurtic than $\beta = 2$ are excluded, on average, but distributions more leptokurtic than that simulated usually cannot be excluded with any

confidence. The behaviour of the moments of the distribution of genotypic values can provide an explanation for the difficulty in placing an upper limit for the kurtosis of the distribution (Keightley, 1998).

(iii) *The problem of flat likelihoods*

If variable mutation effects are assumed, parameters that we wish to estimate from MA experiments become strongly confounded with one another, as seen in results of analysis of real data by ML (Keightley, 1994; Keightley & Ohnishi, 1998; Fry *et al.*, 1999). In all published studies so far with ML, estimates of mutation parameters have come from 'profile likelihoods' in which likelihood is maximized for a series of fixed values of one parameter of interest. This is a well-known technique in, for example, animal breeding (e.g. Graser *et al.*, 1987; Visscher *et al.* 1991), where computer-intensive likelihood-based estimation procedures have become standard. One reason for performing profile likelihoods is to overcome the problem of locating the global maximum where multi-dimensional likelihood surfaces are flat. If the parameters are confounded very severely, it may be necessary to perform multi-dimensional grid searches. In their analysis of simulated MA data, DLL98 highlight the difficulty in locating global likelihood maxima as a fundamental problem of the method: 'Data not shown revealed that Keightley's M-L program may fail to find global maxima, even with the starting value of the other parameter(s) set close to the true (but generally unknown) values' (DLL98). However, the responsibility to find global maxima lies with the program user. The simplex algorithm can find local maxima, depending on the starting values it is given, and is not guaranteed to reach the global maximum if the likelihood surface is flat. Clearly, if the quality of the data is poor,

Table 2. *Performance of the ML procedure illustrated by analysis of data simulated with two different gamma distributions, and analysed under the assumption of gamma distributions with a range of shape parameter β**

Simulated values				Estimated values					
Model	U	$\bar{s}$	β*	$\hat{U}$	SD	$\hat{\bar{s}}$	SD	$\overline{LogL}$	SD
$\beta = 4$	1	2·83	$\rightarrow \infty$	0·94	0·15	3·06	0·42	−2·4	2·3
			10	0·93	0·14	3·10	0·36	−0·4	1·0
			4	1·02	0·15	2·84	0·33	0·0	—
			2	1·21	0·18	2·42	0·29	−0·4	0·9
			1	1·59	0·24	1·86	0·23	−1·5	1·7
			0·5	2·35	0·36	1·27	0·16	−3·1	2·5
$\beta = 0·5$	1	1·83	$\rightarrow \infty$	0·46	0·11	3·85	0·63	−16·3	6·5
			10	0·41	0·079	4·25	0·52	−8·0	4·1
			4	0·44	0·080	4·01	0·53	−3·8	2·7
			2	0·52	0·094	3·49	0·49	−1·4	1·6
			1	0·69	0·12	2·70	0·40	−0·2	0·7
			0·5	1·02	0·18	1·87	0·29	0·0	—
			0·25	1·67	0·30	1·15	0·18	−0·2	0·5
			0·125	2·97	0·54	0·66	0·11	−0·4	0·8

Results shown are based on 40 replicates. There were 200 control and 200 MA lines, and the ratio σ_b^2/σ_e^2 was 10. Mean LogL ($\overline{LogL}$) and its standard deviation (SD) refer to means and mean squared differences from the LogL of the 'correct' model.

maximization of likelihood by any method may be difficult. These points were emphasized in the paper describing the ML method (Keightley & Ohnishi, 1998), and in the release notes provided with the computer code: 'It is strongly recommended to generate profile likelihoods, by keeping one of the parameters fixed. Convergence over the full likelihood surface is often a problem. Also, some runs should be checked with different starting values, again to check convergence, and for multiple peaks.' It is an option for the user to program-in an alternative maximization procedure if the one provided is found to be unsatisfactory.

The problem of flat profile likelihoods also occurs, but is not explained in DLL98 fig. 7*e*, panels A and B, where the relationship between fixed 'ML' estimates of U and β is shown. Curves are not monotonic in these figures in cases where β is very far from the value which best fits the data, and a broad combination of parameter values can give essentially equally poor fits. This would have been exemplified more clearly as a two-dimensional grid plot. The true relationship between the parameters is monotonic.

(iv) Concluding remarks

MA experiments are time-consuming, tedious and costly, and often give noisy or inconclusive results. It is therefore desirable to employ analysis methods that extract the maximum amount of information from the hard-won data. The results presented here and previously (Keightley, 1998) suggest that using ML can give worthwhile increases in precision and decreased bias compared with the BM approach. Likelihood allows the analysis of more complex models, unbalanced data, and makes fuller use of the information available in the data. In general, likelihood-based approaches have become the method of choice in many areas of statistical genetics (Lynch & Walsh, 1998, chaps. 13, 15, 16, 27). For example, restricted ML methods have largely replaced analysis of variance for estimation of genetic parameters in animal breeding. The ML approach has recently been extended to handle experiments with several assay generations, and analysis of data from two multi-assay MA experiments in *C. elegans* provided parameter estimates with sampling variances of the order of 100 times smaller than obtained by the BM method of moments (Keightley & Bataillon, 1999). Furthermore, including intermediate generations in the analysis led to considerable increases in precision. Deng & Fu (1998) and DLL98 concluded that including intermediate generations adds little information, but confined their analysis to the BM method. In contrast to the *C. elegans* experiments, similar levels of precision were obtained in BM and ML analysis of *Drosophila* MA data for viability (Fry *et al.*, 1999).

Although MA experiments can provide some information on properties of new mutations affecting quantitative traits, it is doubtful that results of MA experiments alone can allow rejection of evolutionary

hypotheses that they purport to test. The presence of a large class of mutations with very small phenotypic effects can never be excluded on the basis of phenotypic measures made in the laboratory.

References

Bateman, A. J. (1959). The viability of near-normal irradiated chromosomes. *International Journal of Radiation Biology* **1**, 170–180.

Deng, H.-W. & Fu, Y.-X. (1998). On the three methods for estimating deleterious genomic mutation parameters. *Genetical Research* **71**, 223–236.

Deng, H.-W., Li, J. & Li, J.-L. (1999). On the experimental design and data analysis of mutation accumulation experiments. *Genetical Research* **73**, 147–164.

Fry, J. D., Keightley, P. D., Heinsohn, S. L. & Nuzhdin, S. V. (1999). New estimates of rates and effects of mildly deleterious mutation in *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the USA* **96**, 574–579.

García-Dorado, A. (1997). The rate and effects distribution of viability mutation in *Drosophila*: minimum distance estimation. *Evolution* **51**, 1130–1139.

Graser, H. U., Smith, S. P. & Tier, B. (1987). A derivative free approach for estimating variance components in animal models by restricted maximum likelihood. *Journal of Animal Science* **64**, 1362–1370.

Keightley, P. D. (1994). The distribution of mutation effects on viability in *Drosphila melanogaster*. *Genetics* **138**, 1315–1322.

Keightley, P. D. (1998). Inference of genome wide mutation rates and distributions of mutation effects for fitness traits: a simulation study. *Genetics* **150**, 1283–1293.

Keightley, P. D. & Bataillon, T. M. (1999). Multi-generation maximum likelihood analysis applied to mutation accumulation experiments in *Caenorhabditis elegans*. *Genetics*, submitted.

Keightley, P. D. & Ohnishi, O. (1998). EMS induced polygenic mutation rates for nine quantitative characters in *Drosophila melanogaster*. *Genetics* **148**, 753–766.

Lynch, M. & Walsh, B. (1998). *Genetics and Analysis of Quantitative Traits*. Sunderland, MA: Sinauer.

Mukai, T. (1964). The genetic structure of natural populations of *Drosophila melanogaster*. I. Spontaneous mutation rate of polygenes controlling viability. *Genetics* **50**, 1–19.

Nelder, J. A. & Mead, R. (1965). A simplex method for function minimization. *Computer Journal* **7**, 308–313.

Visscher, P. M., Thompson, R. & Hill, W. G. (1991). Estimation of genetic and environmental variances for fat yield in individual herds and an investigation into heterogeneity of variance between herds. *Livestock Production Science* **28**, 273–290.