Copyright © 2004 by the Genetics Society of America

Letter to the Editor

Comparing Analysis Methods for Mutation-Accumulation Data

Peter D. Keightley[1]

School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom

Manuscript received November 25, 2003
Accepted for publication January 19, 2004

THE genomic deleterious mutation rate (U) and the distribution of mutational effects for fitness, $f(s)$, are important parameters for several theoretical issues in evolution (Charlesworth and Charlesworth 1998), and there has been much recent work on the problem of their estimation. There are currently three statistical approaches to infer U and $f(s)$ on the basis of the distribution of fitness estimates among inbred mutation accumulation (MA) lines maintained under relaxed selection: minimum distance (MD; García-Dorado 1997), traditional maximum likelihood (ML; Keightley 1994), and Markov chain-Monte Carlo ML (Shaw *et al.* 2002). These methods extract information from the shape of the distribution of MA line means; this information is not used by the Bateman-Mukai method of moments (BM; Bateman 1959; Mukai 1964). Recently, García-Dorado and Gallego (2003) have compared the performance of the BM, MD, and ML procedures by computing means and variances of parameter estimates in replicated simulated data sets and concluded that MD tends to produce mean estimates with the lowest bias and sampling variance. In this letter, I question the evidence that led to these claims.

García-Dorado and Gallego's (2003) principal claims are that MD produces unbiased estimates of U and the mean mutational effect $E(s)$, that MD outperforms ML by producing estimates of U that have lower bias and smaller mean squared error (MSE), and that ML performs more poorly because many estimates are "large outliers." Table 1 summarizes the data on which García-Dorado and Gallego (2003) base their conclusions. In 4 of 6 cases mean MD estimates for U appear to be less biased than ML, and in 5 of 6 cases MD estimates of MSE are lower. However, there is a notable difference in the number of replicates that were excluded on the basis of failure to converge (15/62 for MD *vs.* 6/60 for ML; χ^2 1 d.f. = 3.43, $P = 0.064$). This difference presumably arises because MD and ML use different algorithms to locate maxima (or minima) in the multidimensional parameter space. ML employs numerical integration to compute likelihood of data as a function of U and $f(s)$ and combines grid searches with the simplex algorithm (Nelder and Mead 1965; Press *et al.* 1992) to attempt to locate the global maximum likelihood. Convergence is declared when the relative change in likelihood between successive iterations is less than a user-defined threshold. The algorithm is guaranteed to converge (although not necessarily to the global maximum) and to produce parameter estimates if the user sets bounds on valid parameter values. MD uses a stochastic algorithm to produce proposal distributions of line means that are functions of U and $f(s)$ and computes "distances" between the empirical and proposal distributions. A grid search is employed to attempt to find the combination of parameter values that minimizes the distance. Failure to converge is declared if the profile of distance as a function of the parameter of interest (*i.e.*, the marginal of distance minimized with respect to all but one parameter) changes

TABLE 1

Comparison of bias and frequency of rejected replicates between ML and MD mutation parameter estimation procedures

Parameter	No. of parameter values for which: ML is less biased	No. of parameter values for which: ML has lower MSE	Replicates rejected: ML	Replicates rejected: MD
U	2/6	1/6		
α	5/6	3/6		
$E(s)$	5/6	4.5/6		
All parameters	12/18	8.5/18	6/60	15/62

Data are from García-Dorado and Gallego (2003), Tables 1–4, in which MD and ML are compared using simulated data that conform to the model assumed. Methods that use the most similar information supplied by the data are compared: "CD-MD" and "ML-C", where available, otherwise "ML-W." In the case of $E(s)$ one replicate gave identical ML and MD MSE estimates, so one-half of a replicate is included in the total.

[1]*Address for correspondence:* School of Biological Sciences, University of Edinburgh, West Mains Rd., Edinburgh EH9 3JT, United Kingdom.

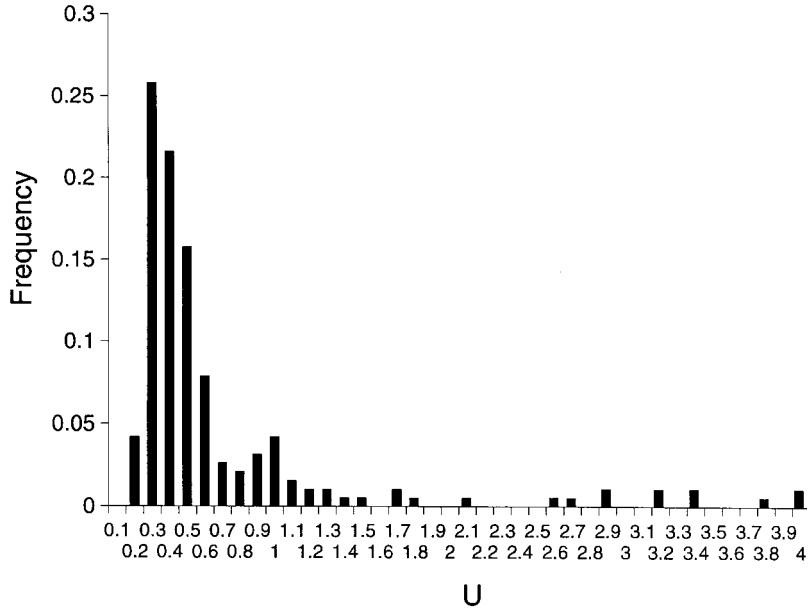

FIGURE 1.—Frequency distribution of estimates of U from 200 replicated simulations of 200 MA lines with parameters $U = 0.5$, $\alpha = 0.5$, and the ratio of genetic:environmental variance = 20. U, α, and $E(s)$ were fitted as unknowns in the model. There were two further replicates that resulted in stable estimates of $U > 4$ ($U = 4.48, 4.84$) and eight further replicates that appeared to result in estimates of $U \rightarrow \infty$.

nonsignificantly over a range of three times the parameter value. This implies that MD can fail to provide estimates if the profile is flat in the region of the minimum. GARCÍA-DORADO and GALLEGO (2003) exclude all MD replicates that fail to converge and any ML replicates for which the ML U estimate exceeds 50.

There is therefore an important difference in the criteria that were used to exclude replicates. Under ML, the set of nonexcluded replicates can contain some very large U estimates below the cutoff of 50 (see Figure 1). I argue that it is highly likely that the excluded MD replicates also tended to be at the upper end of the distribution of U values and that the exclusion of a higher proportion of these extreme replicates led to lower bias and lower sampling variance (Table 1). Replicates giving high U values tend to be excluded under the MD criterion because profiles of distance or likelihood frequently reach plateaus or asymptotically approach limits as a function of increasing U. The existence of such flat profiles has been demonstrated in empirical investigations of MD (GARCÍA-DORADO and MARIN 1998) and ML (KEIGHTLEY 1994; LOEWE *et al.* 2003) and in simulations of MD (GARCÍA-DORADO 1997) and ML (KEIGHTLEY 1998). The behavior does not depend on the way in which the data are analyzed and can be explained by considering the way in which the moments of the distribution of genotypic values of line means (X) change as a function of U: for high values of U the moments of the distribution of X can be held approximately constant by making compensatory changes upward and downward in the values of U and α, respectively (KEIGHTLEY 1998); as U increases, the shape of the distribution of X (*i.e.*, the proposal distribution under MD) can remain almost unchanged.

MA line data often contain insufficient information to allow unbiased estimation of mutational parameters simultaneously. The parameters are confounded in such a way that the best estimate of the mutation rate is often near a plateau in the profile of distance or likelihood. An estimation procedure that rejects nearly one-quarter of such values (Table 1) should not be claimed to show "no bias" (GARCÍA-DORADO and GALLEGO 2003). Furthermore, in cases where U, α, and $E(s)$ are estimated simultaneously, a comparison of means or variances of parameter estimates cannot substantiate a claim that one estimation procedure outperforms another if a significant proportion of replicates are excluded and different exclusion criteria are employed.

LITERATURE CITED

BATEMAN, A. J., 1959 The viability of near-normal irradiated chromosomes. Int. J. Radiat. Biol. **1:** 170–180.

CHARLESWORTH, B., and D. CHARLESWORTH, 1998 Some evolutionary consequences of deleterious mutations. Genetica **103:** 3–19.

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

GARCÍA-DORADO, A., and A. GALLEGO, 2003 Comparing analysis methods for mutation-accumulation data: a simulation study. Genetics **164:** 807–819.

GARCÍA-DORADO, A., and J. M. MARIN, 1998 Minimum distance estimation of mutational parameters for quantitative traits. Biometrics **54:** 1097–1114.

KEIGHTLEY, P. D., 1994 The distribution of mutation effects on viability in *Drosophila 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.

LOEWE, L., V. TEXTOR and S. SCHERER, 2003 High deleterious genomic mutation rate in stationary phase of *Escherichia coli*. Science **302:** 1558–1560.

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., and R. Mead, 1965 A simplex method for function minimization. Comput. J. **7:** 308–313.

Press, W. H., S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, 1992 *Numerical Recipes in C*, Ed. 2. Cambridge University Press, Cambridge, UK.

Shaw, F. H., C. J. Geyer and R. G. Shaw, 2002 A comprehensive model of mutations affecting fitness and inferences for *Arabidopsis thaliana.* Evolution **56:** 453–463.

Communicating editor: S. P. Otto