

QUO VADIS?

¹Ward Wheeler

¹Department of Invertebrates, American Museum of Natural History,
Central Park West at 79th Street, New York 10024, U.S.A.

Received for publication 2 December 1991; accepted 2 December 1991

Albert and Mishler (1992) raise several points in the criticism of my two papers (Wheeler, 1990a,b) describing and using combinatorial weights. Overall, the points raised are divisible into two types, those which arise from a misconception as to the meaning of the weights and those which are based on a probabilistic model of their construction. I will discuss their criticism in this light.

The central point of Albert and Mishler's argument—namely that the values of the **T** matrix are inappropriate as probabilities—proceeds from their mathematical description of the character transformation process and from their notion as to what these **T** values are. When the authors discuss the **T** values they are confusing conditional (**T** matrix) and absolute (their model) probabilities. This confusion is central to their argument. Without treating the **T** values as they treat their probability values, there can be no comparison.

The **T** matrix describes "transformation frequencies from a specific starting point to each of the three alternatives" (p. 271). The **T** values then are most similar to conditional probabilities of change from a base to other bases given that a change occurs and the starting point is known. There is no term for the actual rate of change along lineages, nor is there any term for substitution in general. By some sort of multiplicative process these values might be shoehorned into Albert and Mishler's framework, but the "probabilities" of change would then no longer be so great as to cause the authors concern. Albert and Mishler's probabilities and the **T** matrix are simply not comparable.

The conditional nature of the transformation weights I proposed also goes to the issue of when to apply the weights. Since these probabilities were not absolute, I am uncomfortable with their application *ab initio*. This is certainly possible with new versions of PAUP, as pointed out by Albert and Mishler, but not necessary (although PAUP 3.0 may have been available since March 1989, the paper was submitted in November 1988).

These misreadings aside, the most important points of disagreement between myself and Albert and Mishler concern the model they employ or, more specifically, that they used a model at all.

It seems to me illogical to hold one mathematical description up to another, point out that they disagree, and then champion the superiority of one or another without any empirical reference whatsoever. It would seem that data is required in this process at some level.

This is the general concern I have with the procedure Albert and Mishler employed in their discussion. The foremost problem with maximum-likelihood and other such procedures is their requirement of an external model. A model which is neither derived from nor tested by the data at hand. This is what I was trying to avoid in the first place.

The use of Poisson models of whatever simplicity or complexity, or in fact any probabilistic model which involves rates of change or theory generated transformation modes, fails for precisely the same reason. If the model is correct, fine. But if it is not, what to do then? And how do we know which situation applies—especially in light of the fact that no attempt is made to verify the model? This applies to all cavalier assumptions, however “reasonable”, of transition–transversion ratio, codon bias or mathematical framework.

Albert and Mishler base their ideas on notions which are erroneous (conflation of conditional and absolute probability) and methods I reject (unsubstantiated modelling). It seems misdirected to spurn weights based on observation in favor of those derived from probability models devoid of empirical content. We should be following the path of empiricism and not be distracted by synthetic mathematical descriptions.

Acknowledgments

I would like to thank Paul Vrana, Rob DeSalle, Norman Platnick and two anonymous reviewers for comments.

REFERENCES

- ALBERT, V. A. AND B. D. MISHLER. 1992. On the rationale and utility of weighting nucleotide sequence data. *Cladistics* 8: 73–83.
- WHEELER, W. C. 1990a. Combinatorial weights in phylogenetic analysis: a statistical parsimony procedure. *Cladistics* 6: 269–275.
- WHEELER, W. C. 1990b. Nucleic acid sequence phylogeny and random outgroups. *Cladistics* 6: 363–367.