

Book Reviews

Phylogenetic Trees Made Easy: A How-To Manual, 3rd edition by Barry G. Hall. 2008. 233 pp. ISBN 0-87893-310-5. \$39.95 (pbk). Sinauer Associates, PO Box 407, Sunderland, MA 01375-0407 U.S.A.

This, the third rendition of a popular text on phylogenetics, could rank among the most talked about books within the systematics community today. Its message, that the production of a *publishable* phylogenetic tree can be straightforward and worry-free, is one that raises either hope or ire among its audience. The present review attempts to focus on the utility of the book as both a text for student education and as a manual for the systematics laboratory.

Phylogenetic Trees Made Easy aims at attractiveness to the reader on many levels. It is short (just over 200 pages of text) and nicely illustrated with images of software windows from a few phylogenetics programs. Its organization has improved from previous editions and there are several highlighted boxes ("Learn More" sections) that present the rudiments of major phylogenetic frameworks as well as relevant special topics (tree shapes, rooting, bootstrapping). MEGA, a free and somewhat under-utilized program in systematics, is central to the text. Example data are provided and step-by-step instructions are given for a variety of tasks that range from the search and assembly of GenBank sequences to the manipulation of trees for publication. A handy chapter and several appendices advise on platform-based issues with the software, such as those pesky endline characters in text files that can differ between Unix and Windows, often tripping up our phylogenetic programs. The writing is clear and engaging, and the tone is undeniably optimistic. The book is largely presented in a tutorial style, with significant (though reader-avoidable) interruptions for basic theory behind the methods.

Much has been changed in this edition. The focus on PAUP in edition 2 is now completely gone and nearly all examples are worked in the MEGA framework (with PHYML for maximum likelihood applications and MrBayes for Bayesian inference). Several chapter sections have been promoted to full chapters, which improves the organization and ability to quickly reference a topic of interest. There is now an "Interlude 1" which introduces several methods of phylogenetic inference concurrently, and this treatment of the various frameworks has noticeably improved. Caveats continue to multiply in the text and boxes, which is encouraging to see.

Unfortunately, there are also several questionable statements and stances in the book. I began marking them with sticky notes as I reviewed the text, but these became bothersome from their sheer number. Some of the tagged comments have to do with the "rules of thumb" that are presented. For example, an alignment is declared "acceptable" for neighbor-joining analysis (p. 56) if the *p*-distance is less than 0.8, which presents an idea from Nei and Kumar (2000) almost as if it was a *p* value of attained significance. The remedy proposed for using neighbor-joining with more diverged data than 0.8 is simply to remove sequences from the alignment that are more divergent. For any sequence alignment that contains a large number of gaps when using CLUSTAL, gap exclusion from the analysis is not recommended because "it removes a large fraction of the sites from consideration" (p. 72), irrespective of possible homology issues in regions where alignment might be ambiguous. Several times in the text the author reports that he selects models for the data by choosing the one with the greatest log likelihood, seemingly unaware that log likelihood for his nested models will improve simply by the addition of more parameters. Not surprisingly, he reports choosing his most parameter-rich model, GTR, in all cases.

Although the author is energetic and optimistic, he seems to lack an appreciation for many key issues in the phylogenetic inference process such as nucleotide homology assessment, model selection, tree versus phylogeny estimation, and the fact that methods differ critically in more ways than operational time. The author's preference is clearly for speed: He states (p. 161) that if a method takes longer than 14 hr, he will probably not choose it for an analysis. Similarly, on selecting a support measure (p. 163) he writes, "I tend to go with the fastest, easiest method." To a systematist, speed of the software is considered a poor primary criterion with which to

choose an inference strategy. And although Hall's caveats do describe many of the problems that can come with the phylogenetic inference process, his position on these issues sends mixed messages to the reader. He seems to imply that the existing differences of opinion about the efficacy of various phylogenetic approaches are largely a matter of personal preference, and he seems to presume that most methods will bring us to the right answer for our data sets—hence, his preference for speedy algorithms.

I find it unsettling that the sequence data used throughout the book come from bacteria, mostly *Escherichia* and *Shigella*. Since only bifurcating topologies are discussed as "phylogenetic trees," the reticulate evolution of these taxa is completely overlooked. Yes, one can rapidly produce a tree for a set of *Escherichia* sequences, but clearly this is not the same as a phylogeny of those organisms which experience frequent lateral gene transfer. The absence in this book of network approaches, of methods to search for recombination in bacterial sequence data, or even of the use of a freeware program like SplitsTree (Huson and Bryant 2006) to identify character conflict, suggests that the production of a bifurcating tree shape for publication is the priority here, and not the production of a phylogeny estimation.

Perhaps Hall's view of phylogenetics is best revealed in his list of tasks (p. 3) to be performed: 1) identify sequences to be included; 2) align sequences; 3) estimate the tree by one of several methods; and 4) draw the tree and present it to the intended audience. One might reasonably ask why there is no data quality assessment, no sensitivity analysis, no inquiry about quality of the estimation itself, no testing of alternative hypotheses, no exploration of whether the data is evolving in a tree-like fashion. In this regard, the "Learn More" sections, which are set aside in colored boxes throughout the book, often appear more like "Read If Interested" sections than as the essential background information that is required for quality phylogenetic inference.

Is it Hall's failing or the reader's if the background knowledge underlying phylogenetic methods seems tangential here? As stated in Chapter 1, the book is not intended to be used as a primary text in phylogenetics courses. But that leaves me wondering, "So what is it intended for?" I find this aspect of the book most puzzling. The text is clearly not sufficient as an introduction to phylogenetic inference or to rapid, *quality* phylogeny estimations for publication. A tutorial is a better description of its potential use, and if the book was billed as such I might be more comfortable with it. Then again, many of the programs described have useful tutorials already available on their websites or in their manuals, which makes the tutorial nature of this book somewhat redundant. That brings us back to the book as a "How-To Manual" for producing phylogenetic trees, as the subtitle states. Regrettably, it is clearly insufficient for that purpose.

I want to like *Phylogenetic Trees Made Easy*, but in the end I cannot bring myself to recommend it. Is it accurate? Often, but frequently not. Many of the proper caveats are made, but they can seem tangential to the drive of this book, which is to create easy trees to publish as phylogeny estimations. Is it useful in the classroom? Not particularly or uniquely so. Does it serve its purpose? If that purpose was clearly discernible, perhaps an answer could be provided. I empathize with the author's evident desire to provide a quality introduction to the phylogenetic inference process for those biologists who have little training in systematics. Such an aim is worthy, to be sure, and some day it may be achieved. I would not, however, advise such students and colleagues to learn and conduct phylogenetic estimations based on the recommendations of this manual.

—Scot A. Kelchner, Department of Biological Sciences, Idaho State University, Pocatello, Idaho 83209-8007 U.S.A.

LITERATURE CITED

- Huson, D.H., and D. Bryant. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23: 254–267.
Nei, M., and S. Kumar. 2000. *Molecular Evolution and Phylogenetics*. Oxford: Oxford University Press.

Flowering Plants of Kerala: A Handbook by T. S. Nayar, A. Rasiya Beegam, N. Mohanan, G. Rajkumar, and M. Sibi. 2006. 1069 pp. ISBN 81-900397-6-8. \$95.00 (hbk). Tropical Botanic Garden and Research Institute, Palode, Thiruvananthapuram - 695 562, Kerala, India.

The flora of the Indian subcontinent should be of interest not only to Indian botanists but to all those plant scientists interested in historical biogeography. Once part of Gondwana, India moved north across the equator and finally docked with the Eurasian Plate some 50 million