

Seeing the Forest for the (Gene) Trees

Author: NEPOKROEFF, MOLLY

Source: BioScience, 52(6) : 531-534

Published By: American Institute of Biological Sciences

URL: [https://doi.org/10.1641/0006-3568\(2002\)052\[0531:STFFTG\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2002)052[0531:STFFTG]2.0.CO;2)

The BioOne Digital Library (<https://bioone.org/>) provides worldwide distribution for more than 580 journals and eBooks from BioOne's community of over 150 nonprofit societies, research institutions, and university presses in the biological, ecological, and environmental sciences. The BioOne Digital Library encompasses the flagship aggregation BioOne Complete (<https://bioone.org/subscribe>), the BioOne Complete Archive (<https://bioone.org/archive>), and the BioOne eBooks program offerings ESA eBook Collection (<https://bioone.org/esa-ebooks>) and CSIRO Publishing BioSelect Collection (<https://bioone.org/csiro-ebooks>).

Your use of this PDF, the BioOne Digital Library, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Digital Library content is strictly limited to personal, educational, and non-commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne is an innovative nonprofit that sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Seeing the Forest for the (Gene) Trees

Phylogenetic Trees Made Easy: A How-To Manual for Molecular Biologists. Barry G. Hall. Sinauer Associates, Sunderland, MA, 2001. 179 pp., illus. \$24.95 (ISBN 0878933115 paperback).

In recent times, the importance of "tree thinking," that is, incorporating the historical perspective provided by phylogeny, in the study of diverse biological disciplines has been both acknowledged and celebrated (Harvey et al. 1996). Increasingly, phylogenetic trees are used as tools for understanding biological processes, and in molecular biology for understanding gene and protein function. Some examples of new uses for phylogenetic trees in molecular biology and medical sciences include reconstruction of ancestral gene and protein sequences (Chang and Donoghue 2000), prediction of protein secondary structure (Goldman et al. 1996), and "applied phylogenetics," for example, origin and evolution of HIV (Korber et al. 2000),

prediction of influenza lineages (Bush et al. 1999) or antibiotic resistance, and *in vitro* evolution of pharmaceuticals and evolutionary methods in drug discovery (Hillis 1999). The relevance of an evolutionary framework for the interpretation of gene function is exemplified by the fact that information on new gene and protein sequences in national databases is far surpassing knowledge of gene function. By comparing new sequences with related sequences of known function, a phylogenetic approach can provide a powerful tool for assigning new sequence function.

In his new book, *Phylogenetic Trees Made Easy: A How-To Manual for Molecular Biologists*, Barry Hall provides a primer for the molecular biologist who is interested in reconstructing phylogenetic trees to address new problems in gene relationship, function, and molecular evolution and to serve as a framework for reconstructing ancestral DNA and protein sequences. The role of this book is to give enough background such that a mol-

ecular biologist can find and use the appropriate and most widely used computer software for analyzing evolution of genes and sequences. While this audience may at first appear to exclude more "organismally oriented" scientists (e.g., systematists), knowledge of molecular evolution has had a reciprocal illuminatory effect on phylogenetics, that is, it has improved phylogenetic methods by suggesting which properties of molecules should be incorporated in analyses. Indeed, the integration of molecular evolutionary theory, together with information on phylogenetic relationships, is revolutionizing systematics, development, ecology, and many other fields of biology. Thus, I would recommend this book not only to molecular biologists but also to systematists and other evolutionary biologists.

The author's background and research interests lend a particularly interesting and timely slant to this book. Professor Hall is a faculty member at the University of Rochester Department of Biology,



Envisioning Science
The Design and Craft of the Science Image
Felice Frankel
A guide to creating dynamic science photographs—from physics, chemistry, and biology to materials science and nanotechnology.
"A true teacher's gift—a master photographer of the art of science teaches her craft, with patience, graphic detail, and feeling to all of us who need to visualize and represent this world."
— Roald Hoffmann, Cornell University, Nobel Laureate in Chemistry (1981)
320 pp., color throughout \$55

The Earth's Biosphere
Evolution, Dynamics, and Change
Vaclav Smil
"I know of no comparable work where the mastery of the technological, scientific, and historical material relevant to the subject is so seamlessly presented."
— John Katzenberger, Director, Aspen Global Change Institute
360 pp., 138 illus. \$32.95

Current Topics in Computational Molecular Biology
edited by Tao Jiang, Ying Xu, and Michael Q. Zhang
This survey of computational molecular biology combines algorithmic, statistical, database, and AI-based methods for studying biological problems.
Computational Molecular Biology series
540 pp., 100 illus. \$55

What Genes Can't Do
Lenny Moss
A historical and critical analysis of the concept of the gene that attempts to provide new perspectives and metaphors for the transformation of biology and its philosophy.
A Bradford Book
Basic Bioethics series
258 pp., 2 illus. \$34.95

Gene Regulation and Metabolism
Post-Genomic Computational Approaches
edited by Julio Collado-Vides and Ralf Hofestädt
An overview of current computational approaches to metabolism and gene regulation.
Computational Molecular Biology series
295 pp., 68 illus. \$45

Foundations of Systems Biology
edited by Hiroaki Kitano
An overview of the methodologies and techniques of the emerging field of systems biology.
320 pp., 100 illus. \$45

new from the mit press

<http://mitpress.mit.edu>
To order call 800-405-1619.
Prices subject to change without notice.

whose research interests are in predicting evolutionary potential in bacteria as an experimental model system and predicting the evolution of antibiotic resistance genes. The focus of this book is on reconstructing gene trees, as opposed to organismal trees, which at first caught me off guard as a molecular systematist. Although many of the principles of reconstructing gene trees and organismal trees are the same, there are some fundamental differences in questions and approaches. Thus, systematists may be surprised by the use of the term *homology*, which means “similarity” to many molecular biologists but has a different meaning in systematics, namely, “presence of shared, derived characteristics” or “fitting a probabilistic model of evolution.” Likewise, sections such as “Obtaining Related Sequences by a BLAST Search” will seem foreign for most systematists, who will obtain related sequences through other methods—by determining what study organisms are part of their in-group and in many cases gener-

ating the sequences (or other characters) themselves. An example of how molecular evolutionary biologists and systematists will differ in their perspective and how they come to use this book is found in this statement from section 1: “For the purposes of this book, the words ‘taxa’ and ‘sequences’ are used interchangeably. When PAUP* says ‘taxon label’ you can just as well think of that as ‘sequence name’” (p. 44). These examples underscore the major differences between the traditional approaches used by systematists and molecular evolutionary biologists and illustrate some of the biases workers in these fields have developed. However, the phylogenetic and sequence analysis tools described in this book are just as important for more “organismally oriented” students and researchers as they are for molecular biologists.

I found *Phylogenetic Trees Made Easy* to be well thought out, clear, and easy to read. The book is divided into six sections, covering the topics (1) doing

BLAST searches and creating multiple alignments of protein and DNA sequences, (2) creating phylogenetic trees, (3) presenting and printing trees, (4) fine-tuning alignments, (5) reconstructing ancestral DNA and protein sequences, and (6) dealing with some common problems. The book does an excellent job at introducing a number of the phylogenetic computer program “workhorses” (PAUP*, ClustalX, and Tree-Puzzle [originally called Puzzle] for maximum likelihood analysis of protein sequences), as well as some cutting-edge programs for phylogenetic analysis (John Huelsenbeck’s MrBayes). One of the strengths of this book is inclusion of tutorials for the topics listed above and sample data sets (available at the book’s Sinauer Web site).

One especially nice feature of the book is the “Learn More About It” boxes, which provide more information for readers who are particularly interested in the theory behind aspects of phylogenetic analysis. Also, the author has highlighted some interesting new ideas and computer programs in analysis of genes and protein evolution, including phylogenetic inference using Bayesian analysis and codon-specific models for maximum-likelihood inferences. This book may be worth its extremely reasonable retail price alone for the tutorial on reconstructing ancestral sequences using MrBayes. In addition, the book includes access to a Web site that is designed for readers and provides downloads for tutorial data sets, links to program download sites, and one program written by the author (CodonAlign).

I recommended this book to students in a new course I am teaching in molecular approaches to evolution, ecology, and systematics and will most likely be including it as a supplementary text for this course in the future. I am particularly interested in adapting some of the tutorials as computer labs. From a practitioner’s standpoint, one especially helpful feature of this book is that blocks of text may be cut and pasted directly in one’s data file in the PAUP* program so that this important phylogenetic analysis program can be run in “batch” mode by “remote control” rather than through use of

LIZARD SOCIAL BEHAVIOR

edited by Stanley F. Fox,
J. Kelly McCoy, and Troy A. Baird

“This is an original, substantial, and long-needed contribution. The introduction places the subject in context and shows how lizards can provide unique information not readily available through study of other organisms. The book is logically organized, beginning with a focus on individual variation, moving to comparisons between populations, and finishing with species comparisons.”—George A. Middendorf III, Howard University
408 pages, 82 illustrations, \$89.95 hardcover



NATURE’S FLYERS

Birds, Insects, and the Biomechanics of Flight

David E. Alexander

foreword by Steven Vogel

“This book provides a useful popular-level treatment of animal flight, covering not just biomechanical but also ecological and evolutionary aspects of aerial locomotion . . . No comparable treatment of flight exists in the literature.”—Robert Dudley, University of Texas at Austin

384 pages, 97 illustrations, \$49.95 hardcover

PRINCIPLES AND PRACTICES OF UNBIASED STEREOLOGY

An Introduction for Bioscientists

Peter R. Mouton

“A valuable addition to the current literature on stereologic methods . . . Provides the reader with sufficient technical background without being overwhelming with mathematical details, and is designed for a ‘hands-on’ approach to learning stereologic methods.”—Patrick R. Hof, Mount Sinai School of Medicine

304 pages, 54 illustrations, \$34.95 paperback

The Johns Hopkins University Press

1-800-537-5487 • www.jhupbooks.com

the menus. The author also includes subtle nuances necessary for interfacing and integrating output from various programs with each other. The book has been an invaluable reference, and I find myself returning to it for ideas while using these programs.

Admittedly, the idea of a “cookbook” that one can mindlessly plug data into is not one with which many professional phylogeneticists would feel comfortable. More acceptable is the idea of looking critically at one’s data (alignments, trees), with phylogeny construction as both a first step in the process and simultaneously as a guiding framework for further analyses. While the author describes this book in the “Read Me First” section as “a ‘cookbook’ intended as a tool to aid beginners in creating phylogenetic trees,” please do not be fooled! He is clearly attuned to both the simultaneously dichotomous and integrative nature of phylogenetic analysis—for example, in the use of alignments not simply for phylogenetic analysis but also as a tool for identification of active sites in proteins. Although some readers may be biased against a cookbook for doing phylogenetic analysis, there is a genuine need to make phylogenetic analyses accessible to a growing number of researchers. Hall’s book is filled with enough cautions, caveats, and background information that the reader should be able to make important decisions regarding his or her own data analysis.

We are entering into an extraordinarily exciting time for biology. The time has come to integrate phylogenetic analysis with molecular and genome evolution and a diversity of other biological fields. From a systematist’s standpoint, it is not enough to merely use genes to examine organismal relationships. Understanding the molecular evolution of the genes used in phylogenetic analysis, as well as their function, will help bridge the understanding of relationships of organisms with the very processes at the molecular level that drive evolution. Thus, this book will not only serve as a primer for phylogenetic analysis for molecular biologists but may stimulate some new thinking about the role of gene and protein function in phylogenetics and sys-

tematics. Such thinking will undoubtedly enrich both fields.

References cited

- Bush RM, Bender CA, Subbarao K, Cox NJ, Fitch WM. 1999. Predicting the evolution of human influenza A. *Science* 286: 1921.
- Chang BSW, Donoghue MJ. 2000. Recreating ancestral proteins. *Trends in Ecology and Evolution* 15: 109–114.
- Goldman N, Thorne JL, Jones DT. 1996. Using evolutionary trees in protein secondary structure prediction and other comparative sequence analyses. *Journal of Molecular Biology* 263: 196–208.
- Harvey PH, Leigh Brown AJ, Maynard Smith J, Nee S, eds. 1996. *New Uses for New Phylogenies*. New York: Oxford University Press.
- Hillis DM. 1999. Predictive evolution. *Science* 286: 1866–1867.
- Korber B, Muldoon M, Theiler J, Gao F, Gupta R, Lapedes A, Hahn BH, Wolinsky S, Bhattacharya T. 2000. Timing the ancestor of the HIV-1 pandemic strains. *Science* 288: 1789–1796.

MOLLY NEPOKROEFF

Assistant Professor

Department of Biology

University of South Dakota

Vermillion, SD 57069