

Book Review

Syst. Biol. 51(3):536–538, 2002
DOI: 10.1080/10635150290069959

Molecular Evolution: A Phylogenetic Approach.—Roderic D. M. Page and Edward C. Holmes. 1998. Blackwell Science, Oxford, UK. 346 pp. ISBN 0-86542-889-1. \$49.95.

While most students of evolutionary biology are intimately familiar with the neo-Darwinian synthesis of the 1930s and '40s, a more recent evolutionary synthesis has perhaps gone unnoticed. The International Congress of Systematics and Evolutionary Biology, which convened at the University of Maryland in the summer of 1989, was a watershed, marking a conceptual and technical revolution in evolutionary biology. The excitement was palpable and it was clear to all that something very special was happening in the way of a major integration of divergent fields. We were witnessing a confluence of molecular genetics, phylogenetics, and developmental biology. The time had come when scientists from these different fields were sharing ideas and methodologies. Perhaps foremost in enabling this synthesis of disciplines was the development of molecular methods and computer technology, finally made sufficiently user-friendly for evolutionary biologists to participate alongside the princely realm of molecular medical research. To a degree, this was an example of the tail wagging the dog—technology was paving the way for a codification of molecular evolutionary biology: a synthesis of molecular evolution, molecular phylogenetics, and evolutionary developmental biology.

Although molecular biology and phylogenetics have been coalescing in unprecedented fashion during the last decade, the classroom training of systematics students has, perhaps not surprisingly, taken some time to catch up. Recently, however, an introductory, conceptually comprehensive textbook has come onto the market to effectively remedy this situation. *Molecular Evolution: A Phylogenetic Approach* is an attractively produced new book intended to serve as an introduction to the issues of molecular evolution and how they influence the study of the phylogenetic and populational histories of organisms. Some books have been written recently as introductions to molecular phylogenetic analysis and molecular techniques (e.g., Hillis et al., 1996; Hall, 2001), whereas others have focused on molecular evolution as a subdiscipline (e.g., Nei, 1987; Gillespie, 1991; Li, 1997) or on the useful applications (Harvey et al., 1996; Avise, 2000) or statistical methods (Nei and Kumar, 2000) of molecular phylogenetics. Page and Holmes, however, are unique in combining molecular phylogenetics and evolution with an up-to-date explanation of the aspects of basic genetics and molecular biology that are important background for new students of molecular systematists.

Chapter 1, "The Archaeology of the Genome," provides a series of examples of important scientific contributions made by molecular phylogenetics in recent

years, including an understanding of the early evolution of prokaryotes and eukaryotes, evolutionary relationships of humans among the primates, and tracking the source of HIV-1 in a series of Florida AIDS cases to a mutually shared dentist. The coverage is meant to provide only a taste of the field and why it is important and exciting. The examples are nicely explained for a general readership and serve their purpose without attempting to provide a detailed historical background. An absence of detail is especially noticeable, however, in the brief description of the development of molecular evolution as a discipline. The authors bypass the early insights of Zuckerkandl and Pauling (1965) and Cavalli-Sforza and Edwards (1967), although the former are cited elsewhere in the text. Thus, readers looking for a historical perspective on the field should look elsewhere. The text excels in other aspects, such as lucid explanation of concepts. It is also in the first chapter that the pattern of having a short section "For Further Reading," is established. This section cites several readings for the topics covered in the chapter. Literature is otherwise not cited directly in the text and all references are combined at the end of the book. Although this approach makes for smoother reading and is space-saving, our students (a test case for the book) actually found it somewhat annoying, as it was difficult to backtrack to other literature dealing with a specific topic.

Chapter 2, "Trees," is a relatively short and concise introduction to the concepts and terminology of phylogenetic trees, both cladistic and distance-based. Although elementary in general, it introduces more complex concepts such as splits, additive and ultrametric distances, gene trees versus species trees, lineage sorting and coalescence, and consensus techniques. Beginning readers may find some of these concepts difficult to grasp so early in the text, but they are introduced at this point for reference in subsequent discussions in later chapters. On the whole, the writing is clear and the diagrammatic illustrations are excellent (in particular, the one distinguishing parallel evolution, convergence, and divergence—Fig. 2.1.2). The portrayal of the time dimension in some figures may be confusing to some readers (sometimes it merely means the *direction* of time, whereas at others it refers to the actual *amount* of time elapsed since divergence), but our students quickly figured out what was intended. The treatment of trees on the whole is refreshingly eclectic and free of methodological posturing.

Chapter 3, "Genes: Organisation, Function and Evolution," consists of a broad survey of basic genetics, including outlines of the processes of replication, transcription, translation, and gene expression. For systematists, the most useful review pertains to the types, structures, and functions of various genes and DNA regions and the implications of the differences among these for genetic analysis and, in particular, phylogenetic analysis. A huge range of topics are discussed clearly with

illuminating examples and clean, simple illustrations. Two exceptions to the latter are Figure B3.1, which uses a phylogenetic tree to compare competing hypotheses for the origin of introns, and Figure 3.10, which contains a series of diagrams comparing mitosis and meiosis. In the former, interpretation of the tree would have been easier (especially for novices) if a long branch to the eukaryotes were shown, onto which the serial "colonization" by mitochondrial and chloroplast genomes could be indicated by the use of arrows. Our lab group stared at this diagram for quite some time before realizing that the diagram meant to show exactly what we thought it should. Because the mitochondrial and chloroplast genomes (and introns) play major roles in phylogenetic study, this figure is an important introductory one and should be clearer. The mitosis and meiosis diagrams were intended to be simple (which they are) but seem to be showing chromosomal pairing and duplication in mitosis identical to that in meiosis. Thus, although the text is clear on this matter, the diagram may confuse some readers about the significance of chromosome pairing for recombination during meiosis. Surely there is another way to show the patterns simply. On the whole, however, Chapter 3 provides a rapid, thorough and up-to-date survey of basic genetic concepts and should be useful to many students as a reference.

Chapter 4, "Genes in Populations," is likewise concise, extremely comprehensive, and current. Especially excellent were the series of clear diagrams showing populational and historical genetic processes (Figs. 4.3, 4.6, 4.7, 4.10, and 4.16). The use of consistent symbols and diagrammatic style to describe types of selection, genetic drift, population bottlenecks, gene flow, and gene genealogies links these concepts in a clear, meaningful way. A large amount of material is covered in the 65 pages of this chapter and may be a daunting assignment for readers new to population genetics to comprehend it all without flipping pages back and forth. The chapter should provide an excellent review for many readers and contains some interesting, relatively current examples of studies from conservation genetics (genetic variation in cheetahs and lions), biological invasions (hybrid zones in fire ants), and human prehistory (genetic evidence for the out-of-Africa and multiregional models of modern human origins). An occasional small design glitch intrudes, such as a comparison of RST and FST population subdivision measures discussed eight pages before *F*-statistics are even explained. In general, however, the two chapters on basic genetic principles are extremely well laid-out and clearly developed. Such background should be essential reading for molecular systematists.

Chapter 5, "Measuring Genetic Change," considers the important issues of homology determination (especially in sequence alignment) and the correction of evolutionary distances (divergence estimates) for multiple substitutions at the same site and for various differences in rates of substitutional change across sites and among substitutional classes. The discussion of alignment issues is generally lucid, but the dot plots (Figs. 5.4–5.6) are not very useful in explaining how optimal alignments are identified. Also, the otherwise clear and helpful Figure 5.8 has a bold line missing at the bottom of the *Sum of pairs alignment* diagram, which can lead to confusion. Five of the common models of nucleotide substitution are described and compared: Jukes–Cantor, Kimura 2-parameter, Felsenstein 1981, HKY85, and the

general time-reversible (here called general reversible). The question of which model to use for a given data set is dealt with in a likelihood framework. The issues of compensatory change in rRNA stems, differences in base compensation across taxa, and differences in substitutional rates across sites are all well covered, as are methods used to compensate for these data biases. The final portion of the chapter deals with the methods for measuring evolutionary change on trees (estimating branch lengths) and discusses the methods and problems associated with reconstructing ancestral states. A particularly good text box (5.3) explains why consensus sequences cannot be assumed to be estimates of ancestral sequences. On the whole, the chapter lays a strong foundation for understanding the underlying processes of change in sequence data as well as the challenges these processes present for estimating phylogenies.

Chapter 6, "Inferring Molecular Phylogeny," is in many ways the heart of the book. Especially welcome is the broad, relatively unbiased coverage of phylogenetic methodology. The authors also provide perhaps the most approachable introductory survey yet available for understanding which methods work best under which circumstances, and why. There are none of the overblown claims for the philosophical superiority of parsimony or the statistical superiority of maximum likelihood methods, which often mar other texts. The explanations of spectral analysis and split decomposition, relatively new and promising methods for assessing conflict and data bias, are the clearest we have seen anywhere. And the treatment of data exploration methods (section 6.7, *Have we got the true tree?*) is excellent. Complicated and controversial issues such as long-branch attraction are dealt with clearly and realistically, and the short summary of "What can go wrong?" (p. 225) informs us of the reasons that phylogeny estimation may fail to provide the correct tree. Readers will take away a balanced view of the methodological options available and the appropriate applications, strengths, and weaknesses of each method.

Chapter 7, "Models of Molecular Evolution," at first appears to veer off track, especially when a historical treatment of the classical and balance schools of population genetics develops, at some length, into a discussion of the controversies between the neutralist, selectionist, and nearly-neutralist perspectives. However, as the chapter unfolds, treatments of functional constraints on nucleotides, the evolution of base composition and codon usage, and tests of the molecular clock are introduced with clear relevance to phylogeny. The discussion of lineage-specific rate variation and its possible causes is especially insightful. Although we found this chapter, at least in part, to be the most divergent from our interests in phylogeny, we probably learned new perspectives from these diversions. Those issues that first seemed relatively unrelated to phylogeny estimation in the end served to improve our understanding of molecular data.

The eighth and final chapter, "Applications of Molecular Phylogenetics," should make exciting reading for students, and provides some excellent examples of the use of molecular phylogenies for answering questions in other evolutionary disciplines. Examples include reconstructing histories of gene families, investigating coevolution between parasites and their hosts, tracing the spread of infectious diseases between hosts, and estimating the ages (times of origin) of taxonomic groups

by integrating patterns and rates of molecular evolution with fossil data. Several of these areas have received significant original contributions from the two authors, so it is not unexpected that they would be well formulated. Other issues are included in this chapter that, although well presented, seem misplaced in a chapter on applications because they deal more directly with approaches to phylogenetic inference. For instance, most of the discussion on separate versus combined analyses of different data sources (briefly introduced in a text box in Chapter 6) comes up in this chapter—largely because molecular data are being “applied” to help solve problems originally approached with comparative morphology or fossils, for instance. The chapter also features sections on the differences between gene trees and species trees, including an excellent text box describing how gene duplications can actually help root species trees. In their attempt to integrate these issues into the more general discussion of applications, the authors have brought them up in an unexpected part of the book. Yet, taken as a whole, the chapter makes a strong case for the powerful predictive use of phylogenies in other areas of science.

Page and Holmes are to be congratulated for an amazingly well-designed, compact, affordable, and highly readable book that fills an important gap in the classroom. Given some of the complex discussions of models of molecular evolution and methods of phylogenetic inference, the text would likely prove difficult reading for most introductory classes in systematics, particularly for students with weak backgrounds in genetics or statistics. For students with strong interests in phylogenetics and

molecular evolution, this is currently the best place to start.

REFERENCES

- AVISE, J. C. 1994. Molecular markers, natural history and evolution. Chapman and Hall, New York. 511 pp.
- CAVALLI-SFORZA, L. L., AND A. W. F. EDWARDS. 1967. Phylogenetic analysis: Models and estimation procedures. *Evolution* 21:550–570.
- GILLESPIE, J. H. 1991. The causes of molecular evolution. Oxford Univ. Press, Oxford, UK.
- HARVEY, P. H., A. J. LEIGH BROWN, J. MAYNARD SMITH, AND S. NEE, eds. 1996. New uses for new phylogenies. Oxford University Press, Oxford, UK. 349 pp.
- HILLIS, D. M., C. MORITZ, AND B. K. MABLE. 1996. Molecular systematics, 2nd edition. Sinauer Associates, Sunderland, Massachusetts. 655 pp.
- LI, W.-H. 1997. Molecular evolution. Sinauer Associates, Sunderland, Massachusetts.
- NEI, M. 1987. Molecular evolutionary genetics. Columbia Univ. Press, New York.
- ZUCKERKANDL, E., AND L. PAULING. 1965. Evolutionary divergence and convergence in proteins. Pages 97–166 in *Evolving genes and proteins* (V. Bryson and H. J. Vogel, eds.). Academic Press, New York.

Sydney A. Cameron and James B. Whitfield, Department of Entomology, University of Illinois, Urbana, Illinois 61801, USA.