

## Seeing the Forest for the Trees

ANYA PLUTYNSKI

*Department of Philosophy*

*University of Utah*

*260 S. Central Campus Drive*

*Salt Lake City, UT 84112*

*USA*

A review of Roderic Page, *Tangled Trees: Phylogeny, Cospeciation and Coevolution*, University of Chicago Press, Chicago, IL, 2002, 378 pp. (Hb) ISBN 0-226-64466-9, US\$75.00.

Roderic Page's new book, *Tangled Trees: Phylogeny, Cospeciation and Coevolution* (2003), is a worthwhile read for anyone interested in either methodological issues in systematics, or how organisms shape one another's selective environments. "Cospeciation," for the uninitiated, is the concurrent speciation of two or more lineages that are ecologically associated (e.g. host-parasite associations, as well as mutualistic or symbiotic associations). "Coevolution," in contrast, is the reciprocal adaptation of hosts and parasite taxa. The main focus of Page's book is thus when, how and why the branching process of host taxa mirrors that of parasite taxa. "Parasite" here is broadly conceived to be anything from a louse to a virus to a retrotransposon, and "host" may be anything from a genome to a whale.

The book is roughly divided into two main sections. The first half, "theoretical considerations" provides a overview of several methodologies for tree reconciliation; the problem of determining the extent of match between host and parasite phylogenies. This becomes very quickly quite technical; one wishes that the editor had devoted more time in the introduction to a critical discussion of the merits and disadvantages of each approach. As it is, a reader unfamiliar with cost matrixes, the ins and outs of PAUP, or the significance of incorporating branch length into one's tree, will soon become a bit lost. I will make some attempt to clarify the main issues at stake below. The second half of the book is a nice overview of a few of the most well studied empirical examples of coevolution and cospeciation. For those who enjoy learning about organismal adaptation, these last chapters will be of greatest interest. Especially compelling (and quite beautiful) are the illustrated examples of

the specific shapes of avian wing versus body lice and the striking correlation between head groove dimensions of mammalian louses and hair dimensions of their hosts.

Biologists have long recognized how parasite phylogeny may be a clue to host phylogeny, by analogy with looking to common morphology or behavior (Hennig 1966). Obligate species-specific parasites, not unlike genes, can be used as clues to the history of an organism's phylogeny. However, the use of parasite data to reconstruct phylogeny has not only some of the advantages, but also some of the limitations of the use of genomic data. One cannot assume that parasite trees faithfully track host trees, just as one cannot assume that gene trees faithfully track species trees (Nichols 2001); similar patterns of lateral transfer occur for both. Apart from cospeciation, parasites may speciate independently of hosts (called duplication), go extinct, or not speciate despite host speciation (called missing the boat). The most difficult problem that the first half of *Tangled Trees* is meant to solve, however, is incorporating what is called "host-switching" – parasites moving to a new host – into tree reconciliation.

I will confine the bulk of my main comments to these methodological sections, as these would seem to be of greatest interest to philosophers (especially those who have followed the debates between pattern cladists and everyone else over the past few years). Tree reconciliation, at first glance, seems very simple: take phylogenies for host and parasites and compare them. Alas, it is not nearly so simple! There are three main methods of tree based reconciliation, Brooks Parsimony Analysis, Tree Map, and Tangled Trees. Each is a different tool, starting with different assumptions, and each suffers from different but related difficulties. The central problem of all three, however, is that they all treat the trees themselves as data. Thus, as Page puts it, "the study of cospeciation stands or falls on the reliability of our estimates of phylogeny." Ronquist's apparently modest statement that, "The host and parasite phylogenies are usually considered known without error," turns out to commit him to a rather brave assumption. It's almost never the case that even the best resolved tree is believed to be the true tree; they are always spoken of as "estimates." To be fair, each tree reconciliation approach attempts to accommodate this fact; updated BPA attempts to resolve this on the one hand by replacing a single input tree with a set of weighted trees, and for TreeMap, one can also delimit a confidence set of trees and take all comparisons between host and parasite sets, though this too becomes quickly difficult to manage (and has some technical problems, see Felsenstein 1978). Thus, it should be noted that the treatment of trees as data ought to make one at least mildly skeptical of tree reconciliation at the outset.

This concern aside, I turn to a discussion of two of the several methods. The first theoretical paper discusses Brooks Parsimony Analysis, or BPA, one of the first attempts to treat tree reconciliation in a quantitative way. This chapter serves as a nice illustration of how, at least in the case of host-parasite coevolution, philosophical critiques of strictly pattern-based methods in phylogenetics have had genuine influence on the practice of science. Ronquist's *Modified Brooks Parsimony* is an attempt to incorporate event-based thinking into a formerly strictly pattern-based method of tree reconciliation. The original BPA scored the presence or absence of a parasite as a set of binary characters. The parasite trees were treated as character-state trees, which were then used to reconstruct host phylogeny. This method came under fire, however, when it was pointed out that the meaning of such concepts as homology and homoplasy were obscured by using parasite presence and absence as characters on a tree. BPA was thus criticized for treating phylogenetic reconstruction as exclusively a problem of identifying patterns, instead of reconstructing a historical process. Modified BPA instead involves assigning a cost to each of a set of possible association events – cospeciation, duplication, extinction, and so on. The cost is inversely related to the likelihood of the event on one's model. Thus the most "parsimonious reconstruction" – the one with the lowest cost – will be the most likely explanation of the observed data. In Modified Brooks Parsimony, one is seeking the minimum cost, or most parsimonious tree. The advantage of this approach is that one makes explicit at the outset the cost assigned to each type of event. The disadvantage, as mentioned above, is that trees themselves are treated as data.

The second paper, for which the book is named, advances a methodology called "Jungles." The "jungle" is a representation of all possible mappings of the parasite onto a host tree. The input data is a tanglegram – parasite and host trees set opposite one another. Then, different combinations of host-parasite association are posited, and each possible association is assigned a cost on the basis of anticipated costs of various coevolutionary events – cospeciation, duplication, and lineage sorting, and host-switching. As far as I can determine, the major advantage of this approach is also its major limitation – there can be a "large number of solutions" – or one creates quite a few of these jungles. It's never made quite clear how we are to choose between them. A further difficulty is that some of the solutions involve radical shifts in host – from one branch to a quite distant one. It's unclear how often this occurs in nature. Again, this method too suffers from treating the original trees in the tanglegram as data.

Huelsenback et al.'s paper, the third in the collection, is refreshingly clear. They defend a Bayesian approach that uses sequence data directly to

determine extent of cospeciation. While far more quantitatively demanding, it seems that this approach is the methodologically speaking most acceptable. First, this approach does not treat trees themselves as data. Huelsenback et al. accommodate the uncertainty of host and parasite phylogenies by systematically testing to what extent the topologies of two trees compare to the null model of no association, whether the phylogenies are more similar than would be expected if independent processes produced the two, and whether branching times are sufficiently similar to infer strict cospeciation. This approach requires sequence data, since in order to compare rates of change, one needs a measure – the molecular clock – of the relative “length” of branches, or time since most recent node. While the Bayesian approach has advantages over TreeMap, BPA, and Jungles in that uncertainty of mapping is known and one does not assume trees are known without error, the model of host switching that it employs is rather simplistic (it assumes a constant rate of host switching and extinction of any prior host).

The last nine chapters of the book move away from technical issues in systematics toward such concrete issues as the effect of population structure on rates of speciation, how often host switching actually takes place, and the ecological conditions necessary for cospeciation. These last chapters are fine examples of how to develop and test hypotheses of adaptation. The high host specificity of some parasites give ample evidence that organisms shape the selective conditions of their ecological associates. The chapter on pocket gopher-chewing lice associations is also a very nice case study in the history of the development of allozyme and then DNA-based methods for reconstructing phylogeny.

Why is all this interesting to philosophers? Or, how are we to see the forest for the trees? There are three reasons. First, host-parasite, or mutualistic and symbiotic associations occur frequently in nature, and are excellent examples of adaptation in action. So, the study of historical associations between lineages is a compelling example of how to generate and test those infamous “adaptationist” hypotheses. Moreover, it is relevant to the history of species diversification more generally, and has potentially wide applications – not only in evolutionary biology, but also in medicine.

Second, aside from the intrinsic interest of these host-parasite associations, philosophers can find here an illustration of a lively and active debate over the major methodologies for reconciling parasite-host phylogenies. Not for the faint of heart (or those with severe allergies to Bayesianism), the first five chapters of this book provide an overview of the main principles and aims of these competing methodologies, along with arguments by their proponents as to the advantages of one or the other. It seems to be a live question which particular approach is preferable. Those with an interest in scientific

controversy will find ample grist for the mill, and perhaps philosophers can contribute to this debate by helping to clarify some of the assumptions and limitations of these different methods of studying historical associations.

Third, the book as a whole is a fine example of how systematists, population geneticists and ecologists can shed light, from different angles, on the same suite of problems. Coevolution and cospeciation is an area of inquiry in biology that draws upon expertise in several different subdisciplines. Thus it serves as a case in point for a live issue today in philosophy of biology – what it means to be a pluralist about methodology in the sciences and why or whether it may be a good thing. Speciation, and cospeciation in particular, seems to be one of the most active cases of interdisciplinarity in biology. So, for those with an interest in the question of how explanations, methods, and concepts from different sub-disciplines in a science can severally contribute to understanding, this is a fascinating case in point.

### References

- Felsenstein: 1978, 'The Number of Evolutionary Trees', *Systematic Zoology* **27**, 27–33.  
Hennig: 1966, *Phylogenetic Systematics*, University of Illinois Press, Urbana.  
Nichols, R.: 2001, 'Gene Trees and Species Trees Are Not the Same', *Trends in Ecology and Evolution* (July) **16**(7), 358–364.  
Page, R.: 2003, *Tangled Trees: Phylogeny, Cospeciation and Coevolution*, University of Chicago Press, Chicago.

