CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Nature Precedings

Technical Reports in Taxonomy 02-02

taxonomy.zoology.gla.ac.uk/publications/tech-reports/

Treemap Versus BPA (Again): A Response To Dowling

Roderic D. M. Page* and Michael A. Charleston†

**Division of Environmental and Evolutionary Biology, Institute of Biomedical and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK; and*

†Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

Address for correspondence
Roderic D. M. Page
DEEB, IBLS
Graham Kerr Building
University of Glasgow
Glasgow G12 8QQ, UK
email: r.page@bio.gla.ac.uk
fax: +44 141 330 2792

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

Abstract

TreeMap is a computer program for analysing host-parasite cospeciation. We respond to Dowling's (*Cladistics*, 18: 416-435) recent comparison of TreeMap and Brooks Parsimony Analysis (BPA) by showing that Dowling's comparison suffers from several mistakes and flaws. We discuss the problems with both BPA and TreeMap, and show that BPA incorrectly counts the true number coevolutionary events more often than TreeMap 1. We also discuss the two main limitations of TreeMap 1 correctly identified by Dowling, namely its inability to handle widespread parasites, and its coarse optimality criterion (the number of cospeciation events). We suggest a simple fix for widespread parasites. The newly released TreeMap 2 uses a more sensitive optimality criterion than TreeMap 1, addressing Dowling's second concern.

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

Table of Contents

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

Introduction

The extent to which cospeciation, host switching, and other events have structured the evolution of host-parasite assemblages is a key question in phylogenetic analysis of coevolution (Brooks and McLennan, 1993; Page, 2002). In order to obtain accurate estimates of the relative frequency of these events, we need methods for accurately reconstructing the history of host-parasite associations. A number of methods have been proposed (Charleston and Perkins, 2002; Hoberg *et al*., 1997; Huelsenbeck *et al*., 2002; Ronquist, 2002), at least some of which have been implemented in computer programs. Which, if any, method is best has been the subject of some controversy. So, how does a researcher interested in cospeciation choose which method to use? In the first attempt to directly address this question, Dowling (2002) compares the performance of Brooks Parsimony Analysis (BPA) (Brooks, 1981) and the method described by Page (1994b) and implemented in the program TreeMap 1 (taxonomy.zoology.gla.ac.uk/rod/treemap.html). Dowling concludes that BPA is best.

In this paper we argue that there are serious flaws in Dowling's paper: the analyses have numerous mistakes, and his discussion of the two methods is confused. Furthermore, some of Dowling's criticisms are made obsolete by the advent of TreeMap 2 (evolve.zoo.ox.ac.uk/software/treemap) which implements jungles (Charleston, 1998; Charleston and Perkins, 2002). However, it would be both unhelpful and untrue to simply dismiss his results as being due to the use of an old version of TreeMap. Dowling's results highlight the problematic nature of widespread parasites. We suggest a simple "fix" that greatly improves the performance of TreeMap 1 and 2 when analysing widespread parasites.

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

Problems with Dowling's study

In discussing BPA, Dowling doesn't clearly distinguish between two uses of BPA: inferring the history of a given pair of host and parasite cladograms, and inferring host phylogeny from a given parasite phylogeny. His study addresses the first use, and therefore both host and parasite cladograms are given. At no time in either the BPA or TreeMap analysis can the host or parasite tree topologies change. Despite this, Dowling repeatedly assures the reader that the problem of "ghost taxa" (discussed below) is not a problem:

> "…ghost characters only show up when a host switch has occurred and do not provide any support for groupings that were not already supported in the tree" (p. 420)

> "This problem is also readily recognized and does not appear to affect the overall structure of the BPA tree as well" (p. 421)

> "Remember, ghost taxa in no way affect the tree topology, they act only

> "…it can be overlooked as a flaw in the methodology that has no effect on the actual structure of the BPA tree" (p. 431)

> "…ghost taxa that BPA mistakenly produces do not cause any topological changes in the tree" (p. 431).

Asserting the same thing many times does not make it true. Ghost taxa are indeed major a problem when interpreting the evolutionary history of a parasite and its host (see below). The case where the user might infer a host tree from a parasite tree is not relevant to the question at hand.

Dowling also seems unclear as to whether TreeMap can handle host switches: "[a] second criticism of TreeMap is its tendency to underestimate host switching (obviously since it does not incorporate host switching at all)..." (p. 423). This rather extraordinary statement is plainly false. TreeMap does incorporate host switching — if it didn't then Dowling couldn't have undertaken the study he describes. It seems Dowling has confused reconciled trees (Page, 1994a) with the algorithm implemented in TreeMap (perhaps because the default reconstruction produced by TreeMap is a reconciled tree).

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

Problematic examples

Dowling's head-to-head evaluation of BPA and TreeMap 1 contains several problematic examples. For trials 1, 45, 57, and 58 there are inconsistencies between the text, the tables, and the reconstructions depicted in the appendix, such that at least one of these descriptions is erroneous. In two cases (trials 45 and 49) Dowling has arbitrarily chosen on TreeMap 1 reconstruction from the multiple, equally parsimonious reconstructions obtained by that program. In both cases, the reconstruction chosen by Dowling differs from the actual history, when in fact that

history was among the optimal solutions found by the program. Hence, these examples should not be counted against TreeMap. There are two trials (54 and 56) which are particularly interesting and which we discuss in more detail.

Trial 54 is essentially unrecoverable. Dowling postulates that the parasite lineage was on the ancestor of the entire host clade, was subsequently lost from the lineage leading to hosts ABCD, then later recolonised that lineage. TreeMap 1 recovers the host switch, and places the root of the parasite tree on the host lineage leading to EFGHJ (Fig. 1). Because the parasite tree input into TreeMap 1 will not have the extinct lineage postulated by Dowling, the program cannot reconstruct this earlier history. Note that Dowling scores 8 cospeciation events for this artificial history, when in fact there are only six. Perhaps the basal "node" of the parasite tree has been counted as both a cospeciation and a sorting event. Hence, the value of six cospeciations reported by both BPA and TreeMap is correct.

Trial 56 is interesting in that neither BPA nor TreeMap 1 reconstruct it correctly. The problem is that the actual history is not recoverable under either method. In trial 56 the ancestor of the parasite clade V+VI switched from the ancestor of hosts ABCD to the ancestor of hosts EFGHIJ. Because a descendant of this parasite lineage does not infect hosts EF, we must postulate a sorting event early in the history of these parasites. The TreeMap 1 reconstruction has one fewer sorting events than the actual history because it is more parsimonious to postulate that the ancestor of parasites V+VI landed on most recent the ancestor of their hosts G and J (Fig. 2).

In this case the artificial history is not recoverable, based on the information to hand. It is worth noting that the actual history could be recovered using the jungle method implemented in TreeMap 2, if we had information on the relative ages of the host lineages. For example, if we knew that common ancestor of hosts GHJ existed

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

later in time than the ancestor of ABCD, then the switch found by TreeMap 1 would not be feasible (switch 2 in Fig. 3). The switch would have to be made earlier in time, onto the ancestor of EFGHIJ (switch 1 in Fig. 3).

The problem with BPA

As Page (1994b) and others have argued, BPA can overestimate the number of host switches due to the non independence of the characters derived from the parasite tree. Dowling acknowledges this, but as we discussed above, chooses to dismiss it as a minor annoyance due to "ghost taxa." We beg to differ. Any reasonable method should count events correctly, rather than require the user to go through each homoplasious character reconstruction *a posteriori*, checking whether it is erroneous or not. Because of this problem, we cannot immediately use the counts of the different events found by BPA as reliable estimates of the true number of events.

In small cases like the artificial examples presented by Dowling, this might not seem too difficult. However, Dowling himself did not attempt to go through his 62 trials and correct the counts. The values he reports in his study include ghost taxa, and hence are in many cases not the actual number of events. This makes it difficult to determine the success rate for BPA, because we know *a priori* that many numbers in table 5 will not be correct. Moreover, any attempt to use statistical methods to assess whether the congruence between host and parasite trees is due to chance (e.g., Page, 1994b; Siddall, 1996) that relies on BPA will generate spurious distributions of the fit between host and parasite tree. Given that these distributions may comprise thousands of trees, manually checking them is impractical.

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

We can get some inkling of whether a BPA reconstruction contains "ghost taxa" by counting the number of events. In the case of parasites restricted to a single host, each internal node in a parasite tree will belong to one of three categories: cospeciation (C), duplication (D), or host switch (H). Hence for a binary parasite tree with no widespread parasites the total number of C + D + H events is $n - 1$, where n is the number of parasite taxa. TreeMap 1 reconstructions always satisfy this requirement, but BPA reconstructions need not. This is a direct consequence of the "ghost taxa". In Dowling's table 5 there are 40 reconstructions (numbers 19-58) that do not involve widespread taxa. In six cases the total of C + D + H events for BPA exceeds the possible value (ignoring trial 58 where BPA has too few events because Table 5 and 6e incorrectly lists 8 cospeciation events, when there are in fact only 7). However, this by itself will not uncover all the erroneous reconstructions. Of the 40 reconstructions being considered, Dowling states that TreeMap 1 recovered numbers 19-44 correctly. Of the remaining reconstructions (45-58), the only reconstructions where Dowling's table 5 shows that TreeMap failed are those we have shown were flawed (i.e., mistakes in scoring, or unrecoverable). Hence, TreeMap performs very well for these trials. BPA, in contrast, deviated from the number of events in 18 of the 40 reconstructions.

The problem with TreeMap

The major failing of TreeMap 1 identified by Dowling concerns those trials (1-18 and 59-62) that involve widespread parasites. TreeMap 1 consistently requires large numbers of sorting events, when none (1-18) or few (59-62) are implied by the actual history. This is a consequence of how TreeMap 1 treats widespread parasites.

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

As Page (1994b, p. 162) discussed, TreeMap 1 interprets a widespread parasite as representing a larger, unresolved clade. The range of the parasite is mapped onto the node in the host tree that is the most recent common ancestor of all the hosts infested by the parasite. As Page (1994b) noted, although this is computationally very straightforward, it need not be the most parsimonious interpretation of their distribution. In particular, it leads to erroneous reconstructions if applied to a parasite species that has increased its host range through host switching, as found by Dowling.

Although this is indeed a serious flaw in TreeMap, we can suggest a simple way to improve how the program handles widespread parasites, and that is to create additional "dummy" lineages for each occurrence of the parasite on a different host. For example, in trial 1 we can split parasite II into two sister lineages, II 1 on host B and II 2 on host D (Fig. 4a). The optimal reconstruction for this tanglegram requires a single host switch from host B to D (Fig. 4b), which is the actual history. Creating additional "dummy" lineages is not an elegant solution, particularly if the parasite has more than two hosts, in which case we would potentially need to arbitrarily resolve the tree relating each dummy lineage. However, in practice (particularly in the case of molecular phylogenetic studies) there may be little or no need to do this. If parasite individuals have been sequenced from throughout the host range of a widespread parasite species, and those individuals are the terminals in the phylogenetic analysis, then each parasite will have a single host. An excellent example of this kind of study is Johnson et al.'s (2002) investigation of intraspecific genetic differentiation in dove lice.

Numbers of reconstructions

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

Dowling reports analyses of three empirical studies (Hafner *et al*., 1994; Hugot, 1999; Paterson *et al*., 2000), and in each case TreeMap 1 finds multiple reconstructions (more than 1000 in one case). In part this is a consequence of the optimality criterion used. TreeMap 1 scores each reconstruction solely by the number of cospeciation events, which will range from 1 to $n - 1$, where n is the number of parasites. It ignores the other events when scoring reconstructions. BPA scores all events (although we argue that it need not do this correctly). Because there may be many reconstructions with the same number of cospeciation events, TreeMap 1 can yield multiple solutions. In practice, some users have looked at the numbers of duplications, host switches and sorting events to help chose among these reconstructions (e.g., Hugot, 1999; Siddall, 1997). However, we agree with Dowling (p. 431) that TreeMap 1 can force the user to have to trawl through large numbers of reconstructions to find the most appropriate reconstruction. A further reason TreeMap 1 can find multiple reconstructions is that TreeMap 1 does not guarantee that reconstructions involving more than one host switch are feasible. This problem was first noticed by Ronquist (1995), and is discussed in detail in Page and Charleston (1997) and Charleston (1998).

TreeMap 2 avoids this problem by using the jungle method to ensure that all solutions are feasible. It searches for all feasible reconstructions within bounds set by the user (for example, the user can specify the maximum number of host switches any reconstruction can have) and then filters the solutions to remove any that are definitely non-optimal for any set of costs. To evaluate individual reconstructions the user can specify costs for each event (duplication, host switch and sorting events). In this way the user can still explore alternative reconstructions, but not be swamped with many similar, but non optimal solutions.

To illustrate this point, we reanalysed the example shown in Dowling's figure 21, which is based on Hugot's (1999) study of primate pinworms (Fig. 5). We used TreeMap 2 to search for reconstructions over a range of host switching parameters (0-7) using the default costs. We found 32 reconstructions, of which 9 had the best score of 28 non codivergence events. These reconstructions all had seven cospeciation events, and required 5-6 switches, 3-4 duplications, and 4-5 sorting events. One of these reconstructions is shown in Fig. 6.

Note that the values reported in Dowling's table 7 for BPA for Hugot's data set cannot be correct. For the parasite tree C + D + H must equal 16, whereas for BPA it equals 15. Dowling also reports 0 sorting events, which seems very unlikely. Using TreeMap 2 we searched for optimal, feasible reconstructions that had no sorting events, and found 9 reconstructions, all of which required two duplications and 9-10 host switches. One of these reconstructions is shown in Fig. 7.

Multiple reconstructions in BPA

BPA's perceived advantage over TreeMap 1 of finding a single reconstruction is in large part due to Dowling's use of DELTRAN optimization (Swofford and Maddison, 1987) to map parasite "characters" onto the host tree (p. 424). This means that he will only recover a single reconstruction for a given data set. Given that it is possible to have multiple, equally parsimonious reconstructions for homoplasious binary characters on a tree (Swofford and Maddison, 1987), we might ask why impose this constraint on BPA?

Although exact calculations are hampered by the fact that BPA codes are not independent (never mind the issue of manually adjusting the mapping afterwards) we can readily discover multiple reconstructions using the program MacClade (Maddison

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

and Maddison, 1992). For each of the three empirical examples considered by Dowling we used TreeMap 1 to create BPA matrices and used MacClade to compute the number of most parsimonious reconstructions (MRPs) for each character when mapped onto the host tree. For Paterson et al.'s (2000) seabird lice two characters have two equally parsimonious reconstructions each, so there are $2 \times 2 = 4$ possible reconstructions. For Hugot's primate pinworms, there are $2 \times 2 \times 2 \times 2 \times 4 = 32$ different reconstructions, and for the gopher and louse example $2 \times 2 \times 2 \times 2 \times 3 \times 3 = 72$ reconstructions. Clearly, multiple solutions are possible. Dowling does not provide any justification for limiting the set of possible reconstructions by enforcing DELTRAN optimization, a procedure that dates from Wiley (1987).

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

Experimental design

A weakness in Dowling's experimental design is his method of generating host-parasite phylogenies and associations. Dowling generated his 62 scenarios by hand, rather than by computer simulation. The advantages of simulations are that they make explicit the assumptions employed to generate the trees and associations, they enable us to learn under what conditions a method might fail, and how frequent those failures might be. In the absence of explicit rules for generating the artificial histories, it is difficult to generalise Dowling's results. For example, whether one favours BPA or TreeMap based on Dowling's results depends very much on how common widespread parasites are, which is a function of the probabilities of parasite speciation and host switching. If scenarios are generated manually there is no guarantee that they have properly sampled the space of possible scenarios. The importance of good design is highlighted by trials 54 and 56 (see above), which are unrecoverable by

either BPA or TreeMap. It is of interest under what conditions such cases might arise, as this sets limits on our ability to accurately infer the history of a host-parasite association in practice. Lastly, simulations can be used to establish the efficacy of any proposed statistical test of cospeciation. For example, Legendre et al. (2002) used extensive simulations to establish the type I error rate and statistical power of their matrix permutation test of cospeciation.

Summary

Dowling's work is a first attempt at evaluating some of the available methods for analysing host-parasite cospeciation, but it is has some flaws. There is clearly scope for more extensive simulations that evaluate a wider range of coevolutionary scenarios, and test a broader range of methods. Simulations, together with the growing body of empirical studies (Page, 2002) will provide an ongoing challenge to methodologists to develop tools adequate to the task.

No current method of reconstructing host-parasite coevolutionary history is "perfect." Existing methods simplify the range of possible events, and researchers in this field must still struggle with incomplete solutions to this problem (Paterson and Banks, 2001). However, recent developments in both parsimony and Bayesian methods (see chapters in Page, 2002) offer considerable promise. The field has moved beyond the rather tired "TreeMap versus BPA" debate. Furthermore, TreeMap 1 has been recently rendered obsolete by the availability of TreeMap 2. The new version of TreeMap offers much more sophisticated algorithms for computing and displaying reconstructions of host-parasite evolution.

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

Acknowledgements

We thank James Cotton, Martyn Kennedy, Vince Smith, and Arnold Kluge for comments. The development of TreeMap 2 was supported in part by NERC grant GR3/1A095 to RDMP. MAC is supported by the Royal Society.

References

Brooks, D. R., 1981. Hennig's parasitological method: a proposed solution. Syst. Zool., 30, 229-249.

Brooks, D. R., McLennan, D. A., 1993. Parascript: Parasites and the language of evolution, Smithsonian Institution Press, Washington.

Charleston, M. A., 1998. Jungles: a new solution to the host/parasite phylogeny reconciliation problem. Math. Biosci., 149, 191-223.

Charleston, M. A., Perkins, S. L., 2002. Lizards, malaria, and jungles in the Caribbean, in: R. D. M. Page (Ed.), Tangled trees: phylogeny, cospeciation and coevolution, University of Chicago Press, Chicago, pp. 65-92.

Dowling, A. P. G., 2002. Testing the accuracy of TreeMap and Brooks parsimony analysis of coevolutionary patterns using artificial associations. Cladistics, 18, 416-435.

Hafner, M. S., Sudman, P. D., Villablanca, F. X., Spradling, T. A., Demastes, J. W., Nadler, S. A., 1994. Disparate rates of molecular evolution in cospeciating hosts and parasites. Science, 265, 1087-1090.

Hoberg, E. P., Brooks, D. R., Seigel-Causey, D., 1997. Host-parasite co-speciation: history, principles, and prospects, in: D. H. Clayton, J. Moore (Eds), Host-

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

Parasite Evolution: General Principles and Avian Models, Oxford University Press, Oxford, pp. 212-235.

Huelsenbeck, J. P., Rannala, B., Larget, B., 2002. A statistical perspective for reconstructing the history of host-parasite associations, in: R. D. M. Page (Ed.), Tangled trees: phylogeny, cospeciation and coevolution, University of Chicago Press, Chicago, pp. 93-119.

Hugot, J.-P., 1999. Primates and their pinworm parasites: the Cameron hypothesis revisited. Syst. Biol., 48, 523-546.

Johnson, K. P., Williams, B. L., Drown, D. M., Adams, R. J., Clayton, D. H., 2002. The population genetics of host specificity: genetic differentiation in dove lice. Mol. Ecol., 11, 25-38.

Legendre, P., Desdevises, Y., Bazin, E., 2002. A statistical test for host-parasite cospeciation. Syst. Biol., 51, 217-234.

Maddison, W. P., Maddison, D. R., 1992. MacClade: Analysis of phylogeny and character evolution. Version 3.0, Sinauer Associates, Sunderland, Massachusetts.

Page, R. D. M., 1994a. Maps between trees and cladistic analysis of historical associations among genes, organisms, and areas. Syst. Biol., 43, 58-77.

Page, R. D. M., 1994b. Parallel phylogenies: reconstructing the history of host-parasite assemblages. Cladistics, 10, 155-173.

Page, R. D. M., Ed. 2002. Tangled trees: phylogeny, cospeciation and coevolution. Chicago: University of Chicago Press.

Page, R. D. M., Charleston, M. A., 1997. Reconciled trees and incongruent gene and species trees, in: B. Mirkin, F. R. McMorris, F. S. Roberts, A. Rzhetsky (Eds),

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

Mathematical Hierarchies in Biology , Vol. 37, American Mathematical Society, Providence, Rhode Island, pp. 57-70.

Paterson, A. M., Banks, J., 2001. Analytical approaches to measuring cospeciation of host and parasites: through a glass, darkly. Int. J. Parasitol., 31, 1012-1022.

Paterson, A. M., Wallis, G. P., Wallis, L. J., Gray, R. D., 2000. Seabird and louse coevolution: complex histories revealed by 12S rRNA sequences and reconciliation analysis. Syst. Biol., 49, 383-399.

Ronquist, F., 1995. Reconstructing the history of host-parasite associations using generalised parsimony. Cladistics, 11, 73-89.

Ronquist, F., 2002. Parsimony analysis of coevolving species associations, in: R. D. M. Page (Ed.), Tangled trees: phylogeny, cospeciation and coevolution, University of Chicago Press, Chicago, pp. 22-64.

Siddall, M. E., 1996. Phylogenetic covariance probability: confidence and historical associations. Syst. Biol., 45, 48-66.

Siddall, M. E., 1997. The AIDS pandemic is new, but is HIV *not* new? Cladistics, 13, 267-273.

Swofford, D. L., Maddison, W. R., 1987. Reconstructing ancestral character states under Wagner parsimony. Math Biosci, 87, 199-229.

Wiley, E. O., 1987. Methods in vicariance biogeography, in: P. Hovenkamp (Ed.), Systematics and evolution: a matter of biodiversity, Utrecht University, Utrecht, pp. 283-306.

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

Figure captions

Fig. 1 Comparison of the TreeMap 1 reconstruction for trial 54 with the actual history. The actual history cannot be recovered, as there is no information in the parasite tree that allows us to infer that the parasite lineage was originally on the ancestor of all the extant hosts. Key to symbols: (●) cospeciation event, (➔) host switch.

Fig. 2 Comparison of the TreeMap 1 reconstruction for trial 56 with the actual history. The reconstructions differ in which lineage the host switch landed on. In the actual history the switch landed on the common ancestor of hosts E, F, G, H, and J, whereas TreeMap 1 reconstructs a more parsimonious history that saves one sorting event by having the switching parasite lineage land on the ancestor of hosts G, H, and J. Key to symbols: (●) cospeciation event, (➔) host switch, (⊕) sorting event.

Fig. 3 Tanglegram for trial 56, showing a time scale (in arbitrary units) and two alternative host switches (numbered 1 and 2). Given the relative ages of nodes in the host tree, host switch 1 is feasible as the source and target host lineages are contemporaneous. However, switch 2 is not feasible as the target lineage (the common ancestor of hosts G, H, and J) was not extant at the time of the proposed switch.

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

Fig. 4 (a) Tanglegram for trial 1, modified so that the widespread parasite species II is split into two taxa corresponding to its two hosts. (b) The optimal reconstruction for this tanglegram. .Key to symbols: (●) cospeciation event, (➔) host switch.

Fig. 5 Tanglegram for primates and pinworms (after Dowling, 2002, fig. 21).

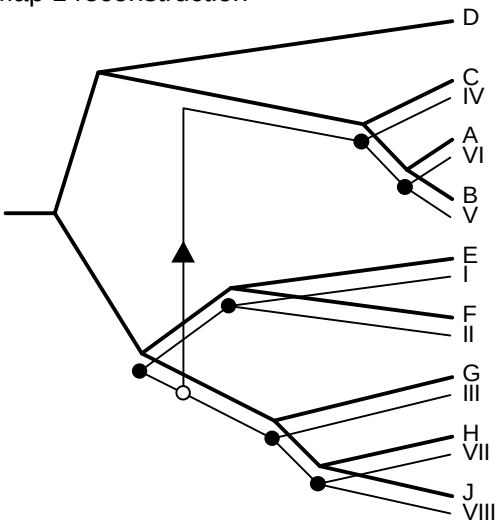
Fig. 6 An optimal reconstruction for the host and parasite trees shown in Fig. 5. This reconstruction requires 7 cospeciation events (●), 3 duplications (□), 6 switches (➔), and 4 sorting events (⊕).

Fig. 7 A reconstruction for the host and parasite trees shown in Fig. 5 that has no sorting events. It requires 9 host switches and two duplications, and has only 5 cospeciation events (symbols as in Fig. 6).

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

TreeMap 1 reconstruction



Actual history

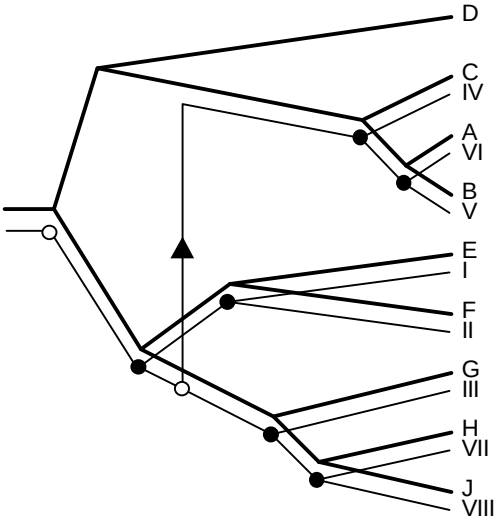


Fig. 1

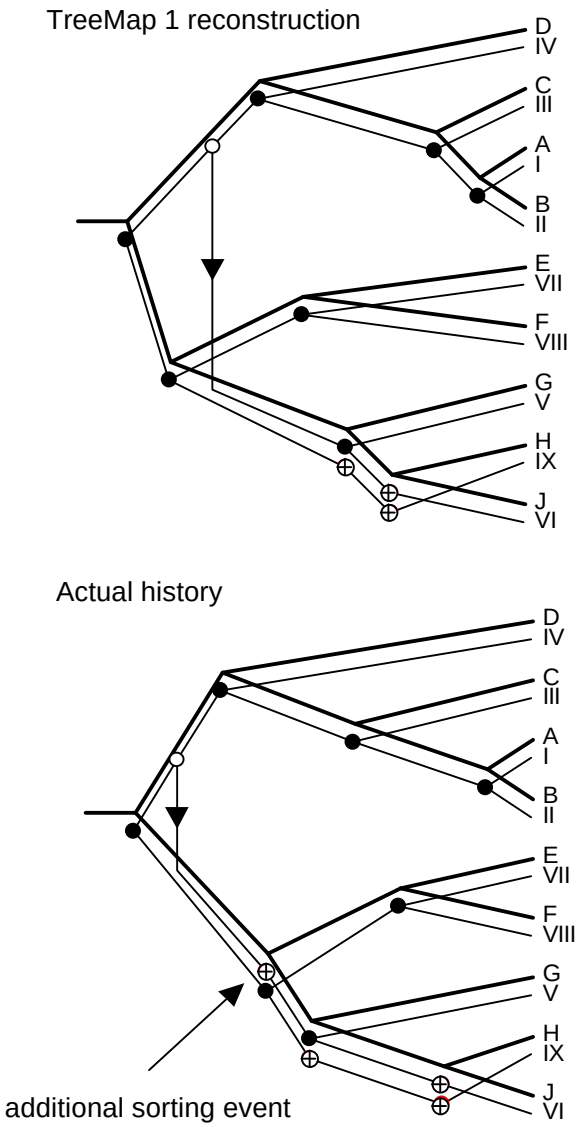


Fig. 2

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

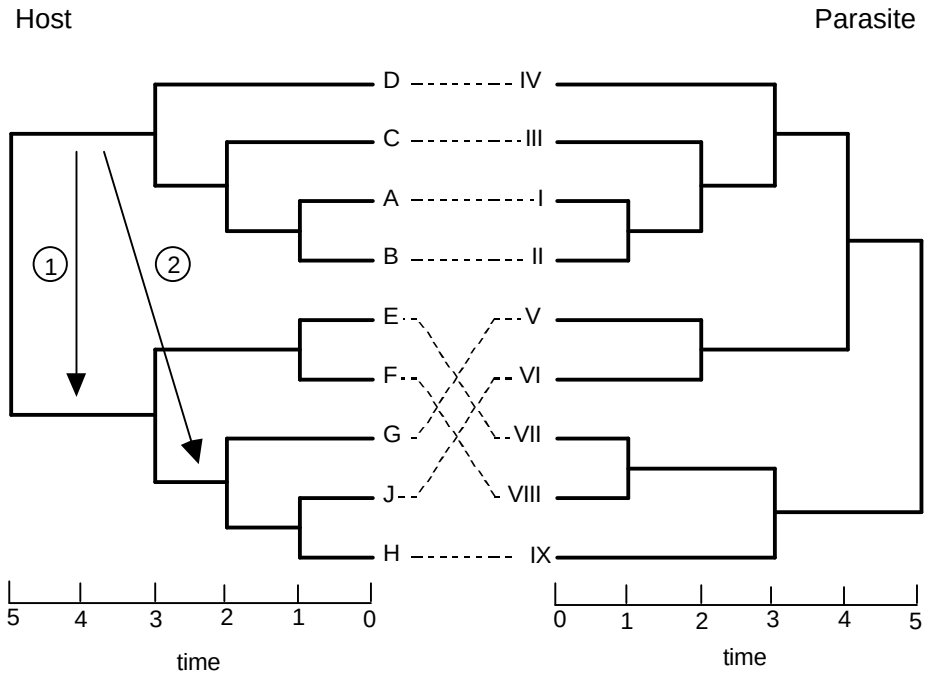


Fig. 3

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

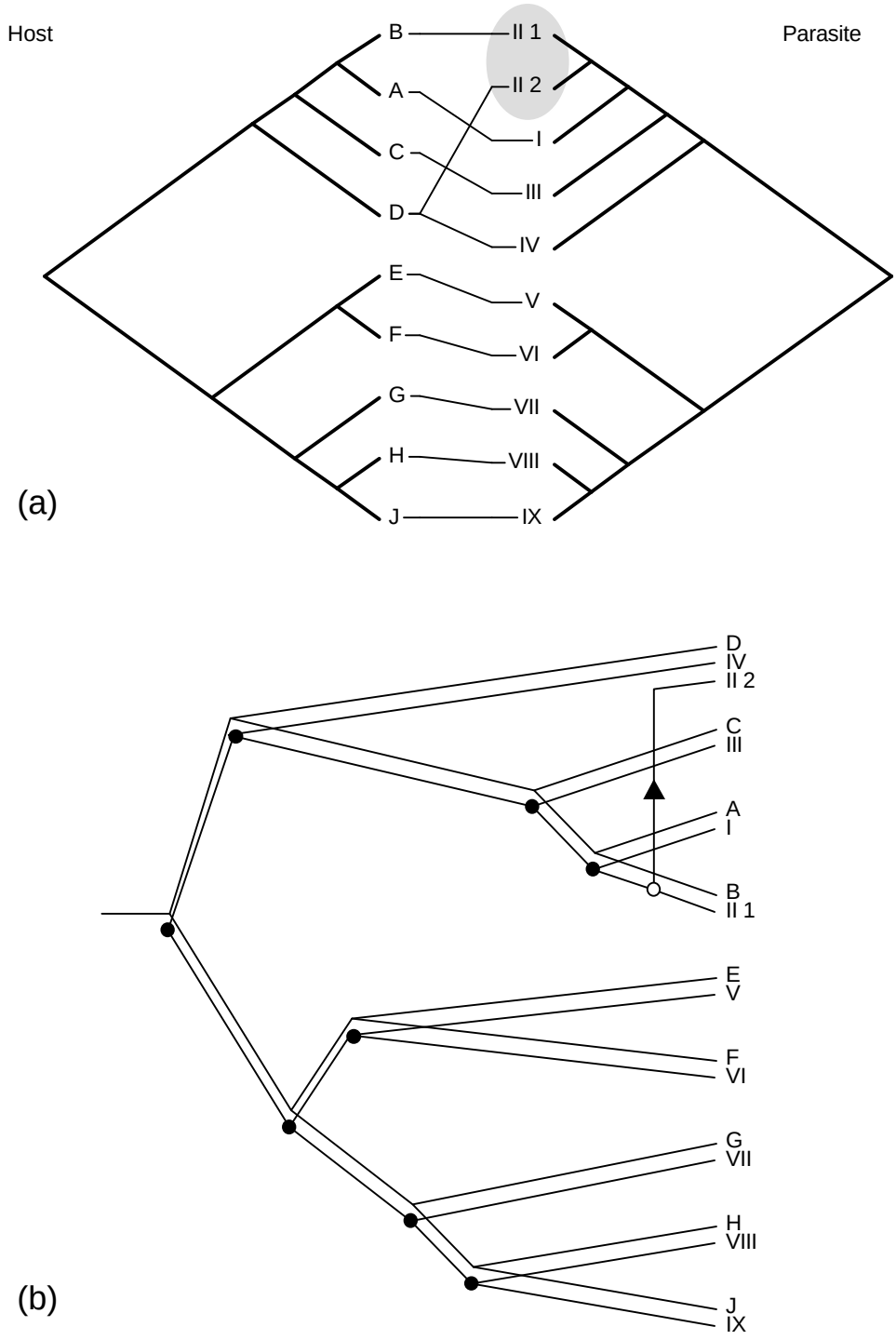


Fig. 4

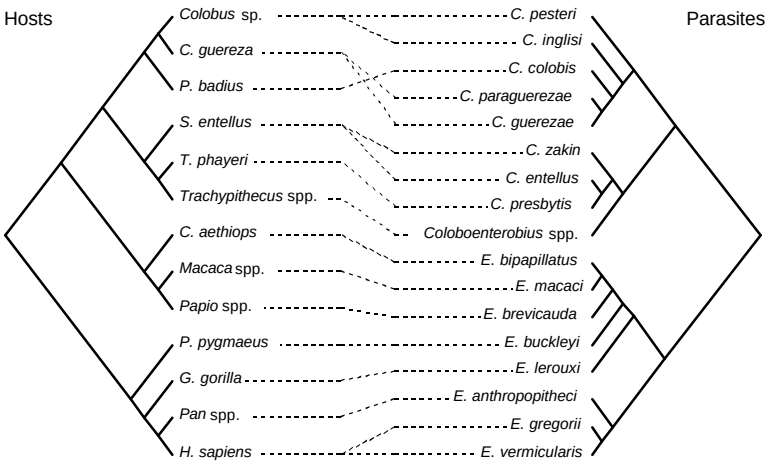


Fig. 5

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

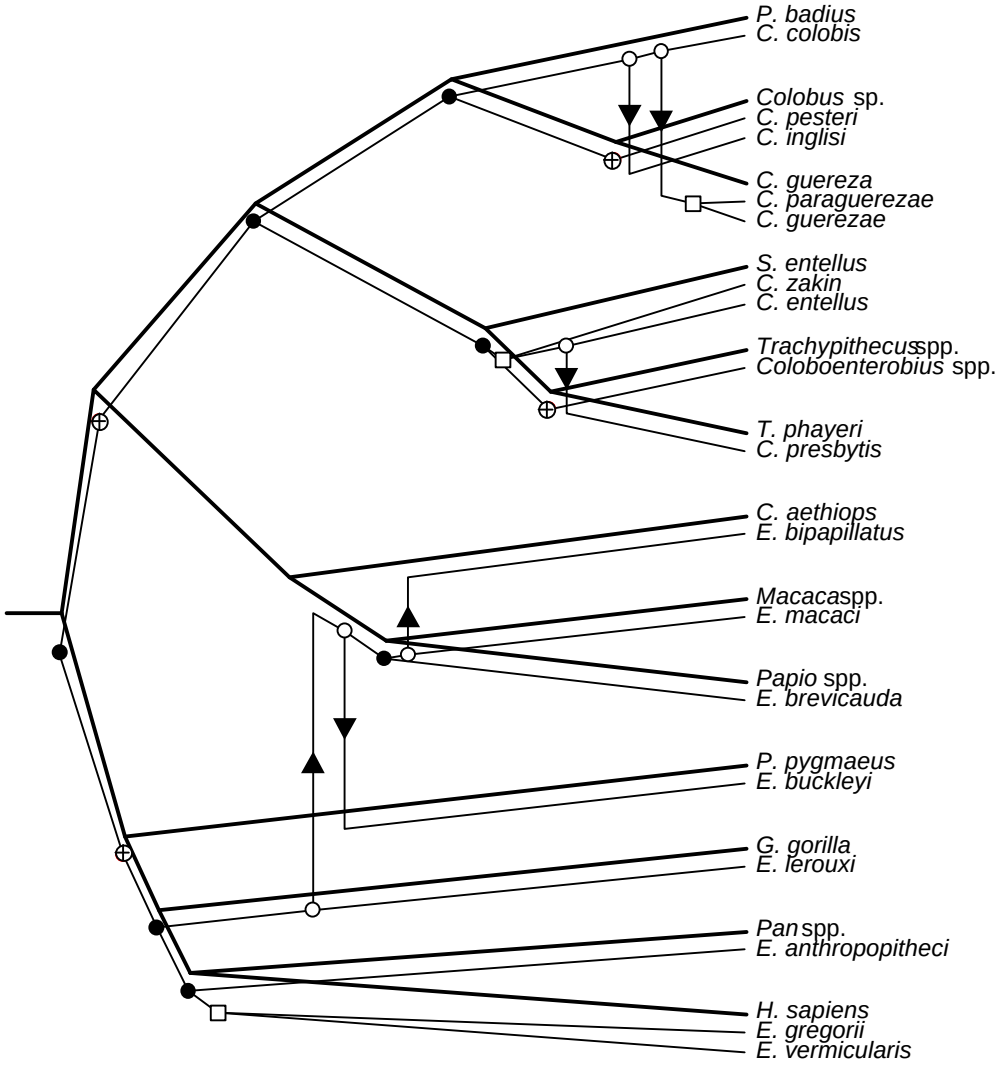


Fig. 6

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007

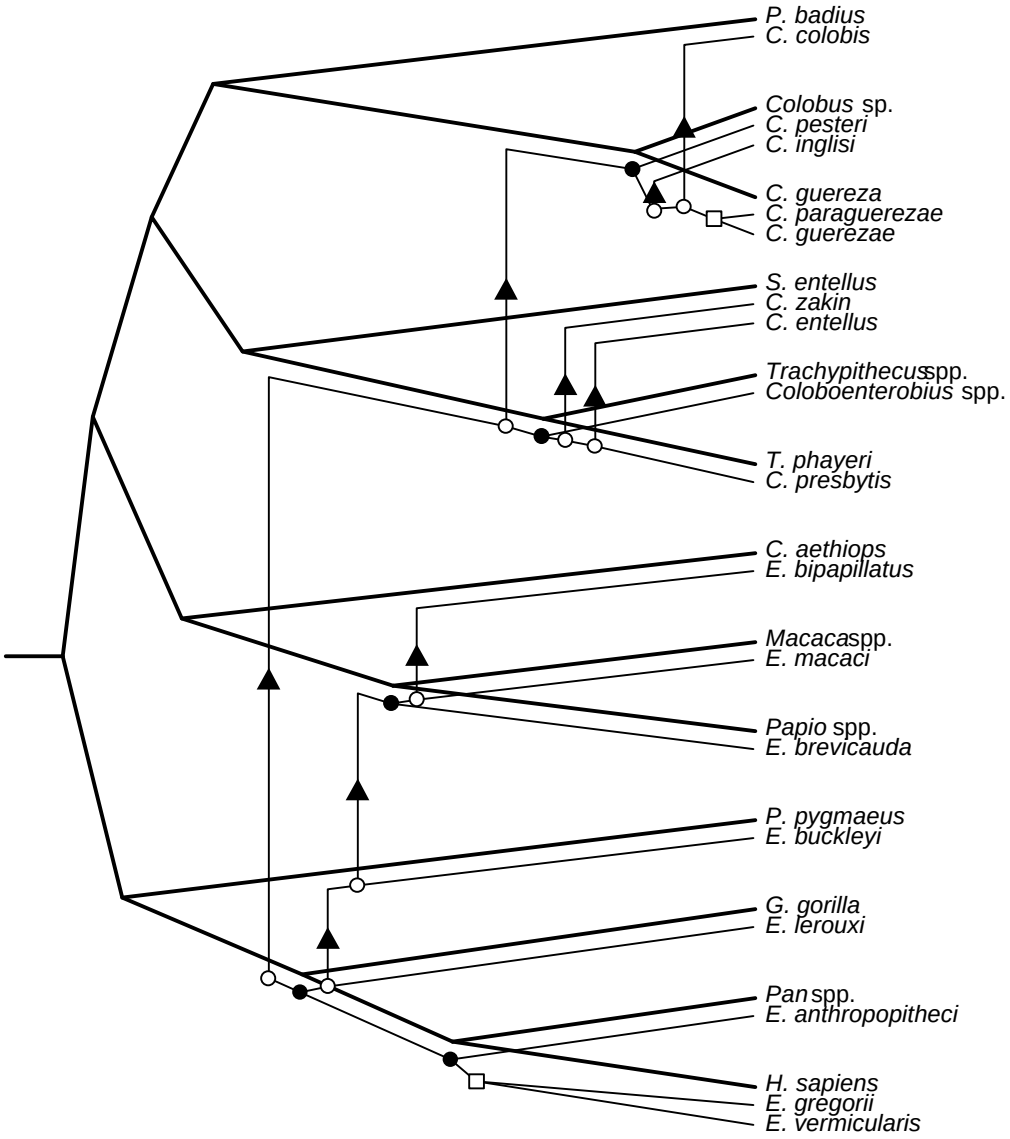


Fig. 7

Nature Precedings : doi:10.1038/npre.2007.1030.1 : Posted 18 Sep 2007