WILEY-BLACKWELL

Cladistics 25 (2009) 515–527

Cladistics

10.1111/j.1096-0031.2009.00258.x

# Unstable taxa in cladistic analysis: identification and the assessment of relevant characters

Diego Pol* and Ignacio H. Escapa

*CONICET, Museo Paleontológico Egidio Feruglio, Avenida Fontana 140, Trelew U9100GYO, Chubut, Argentina*

Accepted 4 March 2009

**Abstract**

A common problem in phylogenetic analysis is the presence of unstable taxa that are depicted in multiple positions in optimal topologies. These uncertainties are reflected in strict consensus trees with polytomies that hamper the interpretation of the phylogenetic results. We propose a protocol for detecting unstable branches (either terminal taxa or clades) and identifying particular characters related to their instability in cladistic analysis. This procedure is based on an iterative evaluation of the agreement of triplets among the optimal topologies (i.e. most-parsimonious trees, MPTs) and examination of character optimizations on these trees. Different types of characters underlying the unstable behaviour of taxa are detected: those with conflicting scorings that support alternative positions of problematic taxa and those with missing data in the unstable taxa that could reduce their instability if they are scored. The entire process is automated through a TNT script that provides a list of characters related to the instability of each unstable taxon. The outcome of this procedure can be used as a guide for further research efforts focused on the revision or addition of (morphological or molecular) phylogenetic data for elucidating the affinities of unstable taxa.



The benefits of extensive taxon sampling on phylogenetic estimates have been increasingly appreciated in recent years. Empirical studies have found that extensive taxon sampling can drastically affect the relationships inferred for a given group and that the exclusion of some ingroup taxa can affect the relationships inferred for the remaining taxa (Arnold, 1981; Gauthier et al., 1988; Donoghue et al., 1989; Wilkinson and Benton, 1995). In many cases fossil taxa are among the most influential, despite being much more incomplete than other (e.g. extant) ingroup taxa (Gauthier et al., 1988; Donoghue et al., 1989; Novacek, 1992; Smith, 1994; Santini and Tyler, 2004; Wheeler et al., 2004). A similar conclusion has also been reached in recent simulation studies, which found that increasing taxon sampling usually improves the performance of most phylogenetic methods (Hillis, 1996, 1998; Poe, 1998; Rannala et al., 1998; Pollock et al., 2002; Zwickl and Hillis, 2002; Hillis et al., 2003). Extending the taxon sampling regime of a phylogenetic analysis is also a desirable general goal given that it increases the empirical content and testability of alternative phylogenetic hypotheses (Grant and Kluge, 2003).

The increasing awareness of this issue has led to a current trend toward conducting taxonomically more inclusive analyses, even though they may include problematic taxa that can generate multiple most-parsimonious trees (MPTs) and a poorly resolved strict consensus. This trend goes against the (unjustified) practice of excluding *a priori* suspected problematic taxa of the phylogenetic analysis solely to gain resolution in the strict consensus, such as the commonly used exclusion of fossil taxa with abundant missing entries (Gauthier, 1986; Rowe, 1988; Benton, 1990; Grande and Bemis, 1998).

Recent advances in algorithmic and software development have allowed the analysis of large and complex phylogenetic datasets (Goloboff, 1999; Nixon, 1999) and the increase in memory management capabilities of

*Corresponding author:
*E-mail address:* dpol@mef.org.ar

current phylogenetic software (Swofford, 2002; Goloboff et al., 2008a,b) allows us to store, compare, and analyse tens of thousands of trees in an efficient manner. Within this context, methods that identify unstable or problematic taxa and retrieve common phylogenetic information from numerous trees (such as maximum agreement subtrees or reduced consensus; Gordon, 1979; Wilkinson, 1994) are becoming increasingly needed. These procedures are conducted *a posteriori* of a phylogenetic analysis including all the taxa and therefore compare sets of phylogenetic trees inferred from all the available data (i.e. the information of all taxa (including those of uncertain position) tests the interrelationships of the analysed group). Much of previous work on these *a posteriori* procedures has focused on producing trees that summarize common phylogenetic information that is obscured in the strict consensus because of the presence of unstable taxa (Wilkinson, 1994, 1995; Thorley et al., 1998). However, from an empirical point of view, the most interesting (although less discussed) aspect of this approach is the identification of unstable branches and the exploration of the underlying causes of such problematic behaviour (Kearney, 2002). Interestingly, this idea was implicit in the earliest development of one of these methods, which aimed to isolate problematic taxa for further study (Gordon, 1979).

The identification of problematic or unstable taxa can be challenging given the complex nature of topological comparisons. Several algorithms that produce summary trees ignoring the alternative positions of some unstable taxa have been proposed although they suffer from some undesirable properties. In the present paper, we propose a method to quickly discover the most unstable taxa on a set of trees, which is the focus of the first section. Once the problematic taxa are identified, we propose a protocol to evaluate the underlying cause of their phylogenetic instability (i.e. lack of information or character conflict; see Kearney, 2002) and identify the particular characters involved in this problem. This final phase, treated in the second section, can provide useful data for guiding further research effort on particularly problematic organisms. These procedures are implemented in a single script for TNT (Goloboff et al., 2008a,b) that automates this protocol given a dataset and a set of MPTs. Lastly, we analyse some empirical cases demonstrating the use of this approach.

## Quantifying topological stability

The task discussed in this section is the identification of unstable taxa (or even clades) from a collection of trees (e.g. MPTs) that are responsible for the lack of resolution in the strict consensus. Although clear sample cases in which a single taxon takes multiple positions (i.e. "wildcard" taxon; Nixon and Wheeler, 1993) are easy to identify, more complex topological variations that show taxa with different degrees of stability are commonly found in real datasets. As noted by Thorley and Wilkinson (1999), measures based on the congruence of triplets quantify the stability of taxa and thus are potentially useful for analysing complex topological variations involving multiple taxa. Here we use a modification of one of such measures, the leaf positional congruence (PC) index (Estabrook, 1992), to evaluate the stability of taxa involved in polytomies, such that unstable branches are quickly identified (concomitantly producing informative reduced consensuses in an efficient way).

### *Positional congruence*

The only method proposed for quantifying the stability of a taxon among different topologies is the leaf PC index proposed by Estabrook (1992). This measure was originally developed for evaluating the stability of a leaf (terminal taxon) in two unrooted phylogenetic trees coming from different datasets, although the same measure can be extended to multiple rooted trees from the same dataset (see Thorley and Wilkinson, 2000). For unrooted trees PC is based on a comparison of unrooted quartets whereas for rooted trees it is based on rooted triplets, given that these are the smallest units of grouping information for each kind of tree (Estabrook et al., 1985; Thorley, 2000). Here we deal only with rooted topologies and therefore will restrict our discussion to rooted triplets. The PC of taxon $x$ ($PC_x$) depends on the number of triplets that include $x$ and display the same grouping information in all topologies. $PC_x$ is obtained by dividing this quantity by the total number of triplets that contain $x$ and therefore this index varies between 0 and 1 (see Fig. 1).

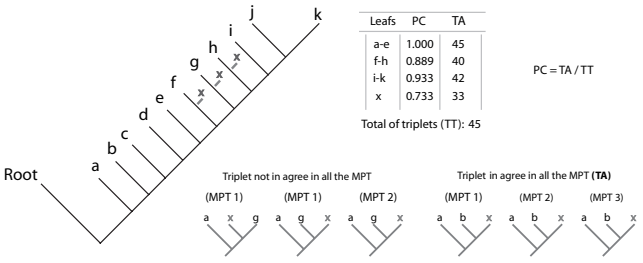


Fig. 1. Sample case of the positional congruence measure (PC). In this example there are three different most-parsimonious trees (MPTs) that only differ in the position of the unstable taxon $x$ (alternative positions shown in grey). The table on the right shows the PC values for each of the terminal taxa (leaves). Note that the PC value of each taxon is the number of triplets (that include such taxon) that have the same topology in all the analysed trees (TA) divided by the total number of triplets that include such a taxon (TT). Examples of triplets are shown at the bottom for cases in which a given triplet does not agree in all the MPTs (bottom left) and cases in which a given triplet agrees in all the MPTs (bottom right).

Although this measure has not been widely used, it has provided the basis for the development of leaf stability, a support measure that accommodates the stability of taxa among trees derived from bootstrapping (Thorley and Wilkinson, 1999; Thorley, 2000).

The PC index presents two main problems that make its use for our purposes difficult. The first is that an exceedingly large number of triplets need to be evaluated to obtain $PC_x$, most of which can be irrelevant to the problem at hand. This property is relatively unproblematic for small datasets, but creates an unnecessarily high computational cost for medium to large data matrices. For instance, if taxon $x$ is included in a single polytomy of degree 4 (i.e. with four descendant branches) in the consensus tree of a dataset of 500 taxa, over 120 000 triplets are evaluated for calculating $PC_x$. Note that only three of them are relevant for detecting the topological variations related to $x$ (i.e. triplets involving $x$ and the three other descendants of the polytomy). The second and more important problem is that in many cases the branches involved in polytomies are not terminal taxa but internal branches leading to clades. In such cases, it would be desirable to measure the stability of the whole clade. However, PC focuses exclusively on the stability of terminal taxa (leaves) and the PC values of the clade members may not reflect the stability of the entire clade. This occurs when, for example, the clade stemming from a polytomy contains one or more polytomies within it.

*Positional congruence (reduced)*

Given that unstable taxa (or clades) are necessarily located in polytomies of the strict consensus, we need only to focus on measuring the relative stability of direct descendants of such polytomies. Evaluating the stability branches not related to a polytomy is irrelevant for our purposes and increases the computational cost with no benefits. Therefore, a modification of PC is introduced (PCR) restricting the scope of triplet comparisons, only evaluating the agreement of triplets comprising the direct descendants of a given polytomy (whether they are terminal taxa or an internal node representing entire clades; see Fig. 2). The PCR is calculated for the descendants of all the polytomies present in the strict consensus by evaluating the agreement of triplets in all the MPTs (as in PC).

PCR drastically reduces the number of triplet comparisons needed to measure the stability of a given taxon (or clade). In the case mentioned above of the polytomy (degree 4), only three triplets would be evaluated (irrespective of the number of taxa present in the dataset). Thus, the number of triplets ($n_{t(x)}$) evaluated for calculating the PCR of taxon $x$ ($PCR_x$) depends only on the degree of the polytomy ($d$) from which $x$ descends, following the formula:

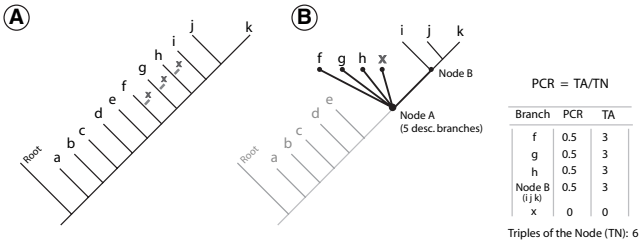


Fig. 2. Sample case of the positional congruence (reduced) measure (PCR). This example is the same as that in Fig. 1. (A) Three different most parsimonious trees (MPTs) that only differ in the position of the unstable taxon $x$ (alternative positions shown in grey). (B) Strict consensus of three MPTs showing a polytomy of degree 5 (Node A). Note that the PCR is only calculated for the five descendant branches of Node A (f, g, h, x, and node B). The table on the right gives the PCR values; note that the unstable taxon has the lowest value of the descendants of the polytomy and that the number of triplets revised for that node (TN) is low.

$$n_{t(x)} = (d-1)!/(2 \times (d-3)!),$$

whereas the number of triplets evaluated in the calculation of $PC_x$ follows the same formula but $d$ is the number of taxa of the entire tree. If the strict consensus contains multiple polytomies, the PCR is calculated for the direct descendants of each polytomy and the number of triplets to evaluate is simply the sum of the triplet comparisons made for each polytomy. Only when the strict consensus is completely collapsed into a single polytomy is $PCR_x$ identical to $PC_x$ and they have equal computational cost.

As mentioned above, PCR also allows us to evaluate the stability of an entire clade that descends from a polytomy. When the descendants of a polytomy include both terminal taxa and an entire clade, the PCR measures the relative stability of all these branches. This allows us to distinguish cases in which a wildcard terminal taxon causes the polytomy from cases in which an entire descendant clade takes alternative positions among the MPTs (see Empirical Cases below).

The outcome of calculating the PCR for a collection of trees is an index that represents the stability of all the branches that directly descend from a polytomy in the strict consensus (Fig. 2). The PCR measures the stability of a taxon (or clade) with respect to other descendants of the same polytomy. This property is ideal for our purposes given that it helps to identify the most unstable branches within the context of a given polytomy, but makes the PCR values only comparable among the descendants of the same polytomy. For instance, when a single problematic taxon takes two alternative positions in a set of MPTs that are otherwise identical to each other (e.g. taxon $x$ in Fig. 2), the $PCR_x$ will be equal to 0, irrespective of the distance between these two alternative topological positions (and therefore the degree of the polytomy generated in the strict consensus).

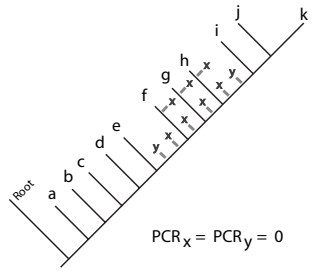


Fig. 3. Sample case showing properties of the PCR measure. The unstable taxon *x* takes seven alternative positions (shown in grey) and the unstable taxon *y* takes only two alternative positions (shown in grey) in an otherwise stable topology. The instability of *x* and *y* creates a polytomy (of degree 6) in a strict consensus.

Another desirable property of PCR is that it measures the stability of a taxon (or clade) in terms of how much its alternative positions are responsible for creating the analysed polytomy in the strict consensus. An example of this property is shown in Fig. 3 in which there are two unstable taxa (*x* and *y*). One of them (*x*) takes multiple alternative positions, whereas the other (*y*) takes only two alternative positions (Fig. 3). However, the PCR values of taxa *x* and *y* are equal, as the alternative positions of the two taxa encompass the same range of internal nodes and therefore their instabilities contribute in the same way to create the polytomy in the strict consensus tree. Although it could be argued that taxon *x* is more unstable than taxon *y* (as it takes more alternative positions), PCR quantifies how much influence each taxon has in creating a given polytomy in the strict consensus. This fits our purposes, as we are interested in identifying the unstable taxa that diminish the consensus resolution.

In simple cases, such as when there is a single unstable taxon (or clade) responsible for the strict consensus polytomy, PCR (as well as other methods) can easily identify it and produce at the same time a resolved and informative reduced consensus. However, many empirical cases show polytomies generated by more complex topological variations with several problematic entities (taxa or clades). Many of the direct descendants of a polytomy can take partially or totally overlapping ranges of alternative positions in the MPTs and have different degrees of stability. These "complex" polytomies are much harder to analyse and require the iterative calculation of PCR after some of the taxa or clades have been pruned from the MPTs.

### *Iterative PCR*

In order to quickly identify the unstable taxa (or clades) we implemented an iterative procedure that progressively selects and prunes problematic branches from the set of MPTs based on the PCR values. The first iteration calculates the PCR values of the descendants of all polytomies of the strict consensus. Then, for each polytomy, it selects the descendant branch(es) with the lowest PCR value and prunes it(them) from the set of MPTs. The set of pruned MPTs are taken as the input of the next iteration, from which a new strict (reduced) consensus is produced and a new set of PCR values are calculated (for the descendants of each of its polytomies). A recalculation of the PCR values during each iteration is necessary because the PCR values obtained in the previous iteration are influenced by the alternative positions of the most unstable descendant identified in the previous iteration (i.e. the branch that was pruned from the trees). Therefore, taking the ranking of PCR values of the first iteration as a guide to rank the stability of all descendants of a given polytomy can be misleading.

This iterative procedure is stopped when all the unstable taxa have been identified (and pruned from the tree) or when two alternative situations are reached, both involving the detection of cases in which all descendants of a polytomy are equally unstable. These stopping rules distinguish this procedure from other methods that also identify unstable taxa and produce a reduced consensus, some of which pruned the trees excessively for our purposes (e.g. maximum agreement subtrees; see below). The first stopping rule is reached when the set of pruned MPTs (after a given iteration) produces a strict consensus in which all polytomies are of degree 3 (trichotomies). In this kind of polytomy it is not possible to identify which of the descendants is the most unstable and therefore they cannot be resolved without taking arbitrary decisions. The second stopping rule is reached when all the descendants of a given polytomy have a PCR value equal to zero. In these cases there is not a single triplet that is supported by all the MPTs, and therefore all the descendants are similarly unstable and will be considered as such. This case is frequently found in polytomies of low degree (e.g. 4–6; see Empirical Cases below).

The descendant branch with the lowest PCR is selected for pruning regardless of whether it is a terminal taxon or a clade, except in one special case: when the problematic clade is excessively large. For instance, if a clade of 20 taxa is pruned in order to resolve a polytomy of grade 4, the relationships (and topological conflicts) within the 20-taxon clade would be ignored in successive iterations. Therefore, the iterative procedure proceeds with the pruning only when the number of descendants of the unstable clade is equal to or smaller than the number of descendants of the polytomy. The script can run under an alternative rule, fixing the maximum clade size allowed to be pruned (see Appendix S1).

*Alternative approaches*

The iterative PCR described above aims to identify the most unstable taxa or clades, and creates, as a byproduct, a reduced consensus with a high content of common phylogenetic information present in the MPTs. Although the PCR index is used in our procedure, other methods exist that can accomplish similar tasks but are based on maximizing different criteria (e.g. number of nodes in a reduced strict consensus, cladistic information content of reduced strict consensus, number of nodes in a dichotomous topology). Most of these methods do not fit our specific purposes and either have a higher computational cost or produce an excessive pruning of taxa.

The most similar approach to the iterative PCR procedure is the method for pruning trees implemented in TNT (Goloboff et al., 2008a,b) through the command *prunnelsen*. This command is executed on a set of trees and has several options, although some of them allow us to find the taxa that improve the resolution of the strict consensus when excluded and can produce the same reduced consensus retrieved from the iterative PCR. However, as implemented in TNT, the improvements in the strict consensus are found by exhaustively testing the exclusion of each of the descendants of a given polytomy (or all the possible combination of *k* descendants). For those polytomies caused by only one (or a few) unstable descendant branch(es) the *prunnelsen* command is extremely efficient and faster than the iterative PCR script. However, if the complexity of the problem increases (i.e. several unstable descendants in large polytomies) the iterative PCR procedure is much more efficient (see Empirical Cases below). Irrespective of implementation differences that affect this comparison, in complex cases the number of operations conducted by *prunnelsen* is much larger than those needed to calculate the PCR values, increasing their difference in computational cost as the complexity of the problem increases.

A related approach is finding the maximum agreement subtree (MAST), which is the largest subtree shared by all the analysed trees (i.e. MPTs). The MAST is obtained by pruning the minimum number of taxa that have to be removed so that the reduced MPTs become identical to each other. This approach, originally proposed by Gordon (1979, 1980), has been implemented in current phylogenetic software (e.g. PAUP*, TNT), although it has not been frequently used in empirical phylogenetic studies. In the simplest cases (i.e. when only one or a few taxa are highly unstable and the rest of the taxa are stable), the taxa pruned from the MAST are the most unstable and therefore it may produce results identical to those of the iterative PCR (and, as implemented in TNT and PAUP*, more efficiently). However, in more complex cases, the MAST can produce very different results and prune more taxa than those identified as unstable by the iterative PCR. For instance, in situations such as the stopping rules described above for iterative PCR, the MAST algorithm will continue pruning taxa until a dichotomous tree is obtained (if the MPTs are dichotomous). Such cases are undesirable for our purposes as there is no rational way of selecting which taxa will be pruned and which will be left in the MAST (or alternatively multiple MASTs are produced). Given the widespread occurrence of these cases in empirical studies, the MAST approach does not fit our specific purposes, i.e. identifying the most unstable taxa in order to later assess the causes of their instability.

Wilkinson (1994) proposed a method for constructing a strict reduced consensus (SRC) that shares some properties with the approach described above. However, instead of focusing on the agreement of triplets (three-taxon statements) involved in polytomies, the SRC is constructed using *all n*-taxon statements implied by the analysed trees (MPTs). The basic output is the SRC profile, a collection of (sometimes numerous) reduced consensus trees that summarize *all* the cladistic information present in the MPTs. This approach has a much broader objective than the iterative PCR, given that the latter creates a *single* reduced consensus as a byproduct of identifying the most unstable branches. However, a single reduced consensus (primary RC) can also be obtained through the SRC method. The primary RC is the most informative reduced consensus tree (e.g. the one with the highest CIC; see Thorley et al., 1998; Wilkinson, 2003) and is selected either from the SRC profile or from a larger pool of topologies that also contain "derivative trees" (constructed "fusing" SRC trees; see Wilkinson, 1995). As with the previous approaches, in the simplest cases, the primary RC can be identical to the reduced consensus obtained by the iterative PCR procedure and can be directly obtained from the SRC profile, as implemented in RadCon (Thorley and Page, 2000). However, in realistic scenarios the primary RC may be very difficult to find; the calculation of the SRC profile can be demanding as it may contains an extremely large number of trees (Bryant, 2003; Wilkinson, 2003) and multiple derivative iterations may be required to obtain the primary RC (Wilkinson, 1995). This is not currently practical for large and complex cases because of implementation limits in RadCon (in terms of number of taxa and trees) and the lack of an algorithm for obtaining derivative trees from the SRC profile. Thus, although the primary RC may well serve to identify the most unstable taxa (i.e. those excluded from this tree), finding this tree is computationally challenging in many cases and the reduced consensus of the iterative PCR seems to be a more straightforward approach for our specific purposes.

## Evaluating unstable taxa

The outcome of the iterative PCR is the identification of unstable branches and a reduced (strict) consensus that display information shared by all MPTs. The latter is the primary objective of some consensus methods (e.g. Gordon, 1980; Wilkinson, 1994) and can provide useful summaries of the topological similarities of the MPTs, although this kind of consensus trees shows some groups whose monophyly was not unambiguously supported when all the taxa had been considered (Kearney, 2002; Grant and Kluge, 2003; Kearney and Clark, 2003). As mentioned above, our primary concern focused on the taxa (or clades) identified as unstable or problematic. The focus of this section is the evaluation of the causes that explain the multiple positions of unstable taxa (or clades) in the MPTs, i.e. the identification of the characters involved in (or potentially relevant to) such a problem. The aim of conducting this analysis is to guide and focus further research and data collection on these problematic taxa, which can be useful in empirical cladistic studies.

The unstable behaviour of taxa in phylogenetic analysis has usually been related to the presence of missing data, as several empirical studies have identified taxa with abundant missing entries (usually fossils) that take multiple alternative positions and produce poorly resolved strict consensus trees (e.g. Gauthier, 1986; Wilkinson and Benton, 1995). However, other empirical studies have noted that the completeness of a taxon is not necessarily correlated with its stability in the MPTs (Gauthier et al., 1988; Novacek, 1992; Gao and Norell, 1998). As discussed by Kearney (2002), multiple alternative positions of an unstable taxon can be caused by lack of information (missing entries) in critical characters, conflict in the character-state distribution, or (more commonly) a combination of these two situations. Kearney (2002) has correctly suggested that failing to distinguish these situations can be empirically misleading in guiding to the real solution of this problem, which lies in the character scorings of the analysed data.

Two types of characters are of interest for a close examination of the evidence that can be revised to help in defining the relationships of unstable taxa: (i) those that are not scored (missing data) in the unstable taxon and may help to determine its phylogenetic placement if the scoring is determined; and (ii) those that are scored and support the alternative positions of the unstable taxon. Here, we propose a protocol for identifying these two types of characters for each of the unstable taxa (or clades) that have been previously selected through the iterative use of PCR. The identification of both types of relevant characters shares a common characteristic: both are discovered by comparing differences in the character optimizations among the MPTs.

### *Potentially relevant missing entries*

Although missing data have been repeatedly related to undesirable effects in phylogenetic analysis, they represent absence of information and cannot bias the analysis in any particular way. Defining the scoring for empty matrix cells is a general goal in a phylogenetic research program, given that the new information increases the empirical content of the phylogenetic analysis. However, it is clear that defining the scorings of some characters may help more than defining others in resolving the phylogenetic position of a particular unstable taxon in a subsequent analysis.

In our approach, for each character scored as a missing entry in an unstable taxon, we evaluate its ancestral condition in all the MPTs. If the condition of character $j$ optimized in the ancestral node of taxon $x$ is identical in all the MPTs, it seems unlikely that defining the scoring of character $j$ for taxon $x$ will reduce the instability of this problematic taxon. By contrast, if the condition of character $j$ optimized in the ancestral node of taxon $x$ is optimized with different character states in different MPTs (Fig. 4), defining the condition of $x$ for character $j$ can provide evidence

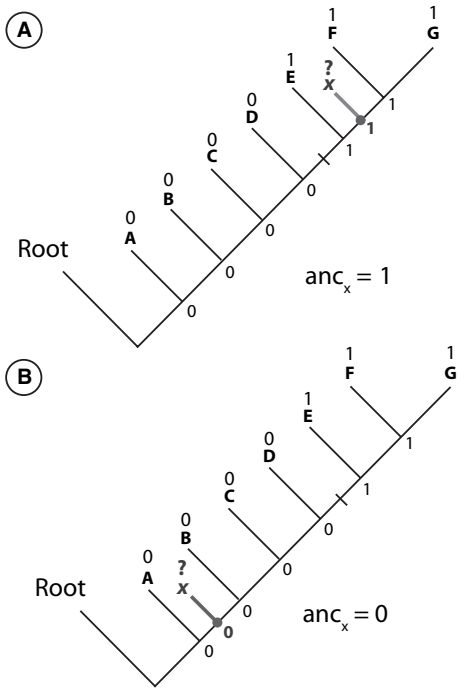


Fig. 4. Sample case showing two alternative positions of an unstable terminal taxon ($x$) and the optimization of a character scored as missing data for taxon $x$. Determining the character state for taxon $x$ could define one of the alternative positions of this taxon. Note that the ancestor node of $x$ is optimized with a different character state in tree A (anc$_x$ = 1) and in tree B (anc$_x$ = 0).

supporting only some of the alternative positions of *x*. Therefore, this type of character may help to reduce the instability of a given taxon in a subsequent phylogenetic analysis.

The TNT script provides a list of this kind of character for each of the unstable taxa, which can help the researcher to reduce the phylogenetic uncertainty of some taxa in future phylogenetic studies. Obviously, defining some (or all) of these scorings does not guarantee that the unstable behaviour of a given taxon will be completely eliminated, but this set of characters is the one that most likely will help to achieve this goal.

*Scored problematic characters*

The second kind of character are those scored for an unstable terminal taxon that only supports some of its alternative positions among the MPTs. This type of character is necessarily in conflict with others and contains homoplastic instances in some of the MPTs. As such, these characters deserve special attention and a careful revision of them can be conducted to re-evaluate their homology statements. Therefore, the identification of these particular instances of homoplasies can also be useful for guiding revisions of the available data and may help to define the position of the unstable terminal taxon.

In our procedure, each of the characters scored for an unstable terminal taxon is evaluated, comparing the difference between the length of character *j* (*L*) for the original data and the length of character *j* when the scoring of the unstable taxon (*x*) is replaced by a missing entry (*L′*). This difference ($D = L-L'$) is calculated for all the MPTs. If some trees have larger differences (*D*) than others, then the information originally scored for *x* in character *j* introduces more steps in some topologies than in others. Or, conversely, character *j* is favouring some of the alternative positions of *x* over others. Therefore, character *j* participates in the character conflict that is (at least partially) responsible for the unstable behaviour of *x* (Fig. 5). If valid reasons are found when the homology statements scored for character *j* are re-evaluated, modifications in the scorings of *x* (or other putatively related taxa) for character *j* can change the amount of evidence supporting some of the alternative positions of *x* and may help reduce the instability of this taxon in a subsequent phylogenetic analysis.

As in the previous case, the outcome of the TNT script is the list of characters that only support some of the alternative positions of each unstable taxon (identified through the iterative PCR). These data can lead the researcher to re-evaluate specific characters to address the phylogenetic uncertainty of unstable taxa in future phylogenetic studies.

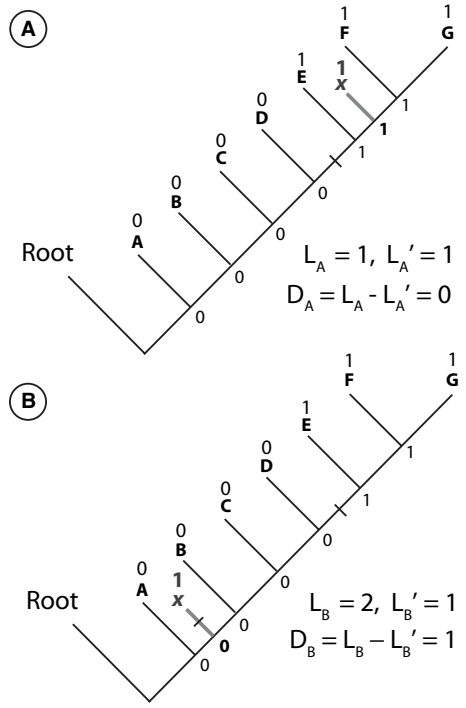


Fig. 5. Sample case showing two alternative positions of an unstable terminal taxon (*x*) and the optimization of a scored character that contributes to the instability of *x*. This character supports one of the alternative positions of taxon *x* (tree in A) but contradicts the other (tree in B) introducing a homoplastic step ($0 \rightarrow 1$ in the terminal branch leading to *x*). Therefore, after replacing the scoring of *x* with missing data, the length of this character ($D = L-L'$) decreases in tree B ($D_B = 1$) but remains constant in tree A ($D_A = 0$).

*Evaluating problematic clades*

So far we have discussed the causes that create an unstable behaviour in a terminal taxon. However, as noted above, in several cases an entire clade can be identified as problematic or unstable. Although the basic idea outlined above can be maintained when evaluating the causes of an unstable clade, these are more complex descendants and require a slightly different treatment.

Evaluating the influence of a given character (*j*) in the unstable behaviour of a given clade (node *X* in Fig. 6A) is conducted in two steps. First, the scorings of all the taxa included in the unstable clade are replaced with missing entries. Second, the character optimization at the base of clade (node *X*) is compared for all the MPTs (Fig. 6B,C). The optimization of node *X* (after the replacement with missing entries) will be determined solely by the optimization of its direct ancestral node. Therefore, if different trees depict different optimizations of node *X* ($opt_X$ in Fig. 6B,C) changes in the scorings of character *j* for the terminal taxa included in the unstable clade can lead to preference for some of its alternative positions over others.

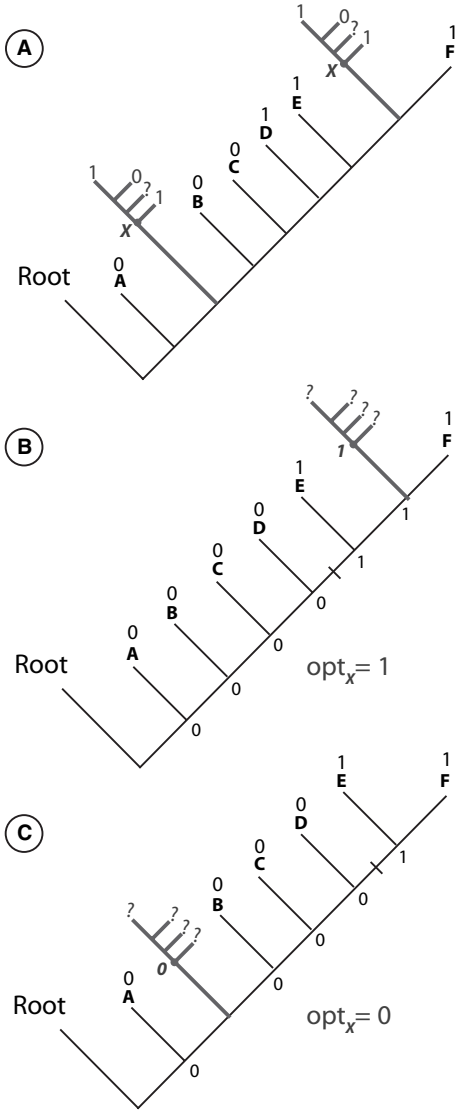


Fig. 6. Sample case showing the character evaluation for an unstable clade (node *X*). (A) Two alternative positions of this clade (marked in grey). The scorings of the analysed character are shown for all terminal taxa. (B) Optimization in one of the topologies (derived position of clade *X*) after replacing the scorings of members of *X* with missing entries. The basal node of clade *X* is optimized with character state 1. (C) Optimization in the alternative topology (basal position of clade *X*) after replacing the scorings of members of *X* with missing entries. The basal node of clade *X* is optimized with character state 0. Revising the scorings of terminal taxa of clade *X* (or defining their missing entries) could reduce the instability of this group.

This procedure is based on a replacement with missing entries as in the evaluation of problematic scored characters in terminal taxa, as internal nodes cannot be "scored with missing data". The uncertainty on the position of node *X*, however, is ultimately caused by the presence of missing data and/or character conflict in the scorings of the members of this clade (usually the most basal forms; see scorings in Fig. 6A). Each of the potentially relevant characters that explain the unstable behaviour of *X* is identified in this way, but the researcher should obviously examine the scorings of the most basal terminal taxa of *X* (either by defining missing entries or by re-evaluating conflictive scorings).

Two major classes of unstable clades can be detected by the iterative PCR procedure: unstable clades that appear in the (complete) strict consensus and unstable clades that only appear in a reduced strict consensus. The former class has a straightforward interpretation and represents clades of undisputed monophyly (in the context of the phylogenetic analysis) that have alternative phylogenetic positions among the MPTs. The latter class, by contrast, are groups of taxa of uncertain phylogenetic position that are non-monophyletic in some of the MPTs. A simple example of the second kind of group is shown in Fig. 7. In this example, the group DEF does not appear in the strict consensus as a monophyletic clade (Fig 7A) because an unstable terminal taxon (*u*) is depicted inside the DEF group in some of the MPTs (see alternative positions of *u* in the reduced consensus of Fig. 7B). Irrespective of the uncertain position of taxon *u* (and the identification of the characters relevant for this problem), the DEF group takes three alternative positions (see these positions in the reduced consensus of Fig. 7C). Therefore, it is necessary to identify which characters are responsible for the unstable behaviour of the DEF group in order to conduct further research to clarify the relationships of the analysed taxa.

The outcome of the TNT script, however, is similar for both kinds of unstable clades and includes a list of the potentially relevant characters as well as a list of the terminal taxa that comprise each unstable group. Thus, subsequent revisions can focus on revising the scorings and determining missing entries for the terminal taxa of the unstable clade (see Appendix S1).

## Empirical cases

In this section we describe the results of applying the procedure described above (implemented in a script for TNT) to 19 real datasets (see Appendix S2). These datasets show different degrees of complexity in terms of the number and degree of polytomies in the strict consensus of their MPTs. We first focus on the outcome of the iterative PCR procedure on a single dataset that illustrates the analysis of a large and complex polytomy with numerous taxa and clades of various degrees of instability. We then summarize the results of the character analysis conducted on unstable taxa, highlighting the frequency of the different causes of instability detected in the 19 analysed datasets and the use of this information in morphological and molecular phylogenetic studies. Further details on the results of each dataset are detailed in Appendix S2.

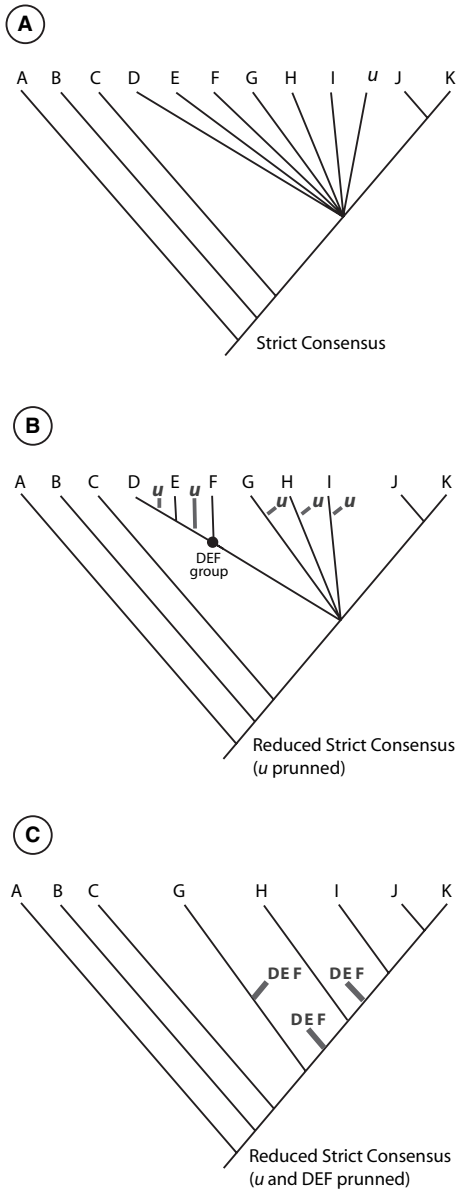


Fig. 7. Sample case of an unstable group (comprising terminal taxa D, E, and F) that does not appear in the strict consensus. (A) Strict consensus tree (including all taxa). (B) Reduced (strict) consensus obtained when unstable taxon *u* is excluded (alternative positions of *u* on this tree are shown in grey). (C) Reduced (strict) consensus when *u*, D, E, and F are excluded. The three alternative positions of the unstable group (DEF) are shown on this tree (in grey). Both the terminal taxon *u* and the group DEF will be selected for analysis of the characters involved in their instability.

### *Application of iterative PCR*

The phylogenetic analysis published by Smith and Stockey (2002) on a group of extant and fossil conifers provides an interesting test case for the iterative PCR (see Appendix S2 for further information). Although this data matrix has <4% of its cells scored as missing data, the strict consensus of the 79 MPTs is poorly resolved. This tree has a single major polytomy (degree = 24) in which 19 terminal taxa and four clades are direct descendants.

The results of the iterative PCR are summarized in Fig. 8. A first set of PCR values is calculated during Iteration 0 for all the descendant branches of the polytomy of degree 24 of the strict consensus (Fig. 8, Node B). The most unstable branch (i.e. lowest PCR value) of this polytomy is a clade composed by two taxa (Tax.21 and Tax.22). These taxa are pruned from the set of MPTs and the new set of reduced trees is the starting point of the next PCR iteration.

The reduced strict consensus obtained after pruning the unstable clade is still unresolved, showing a large polytomy (degree 23). After calculation of the new PCR values in Iteration 1, four branches are detected as the most unstable descendants of the large polytomy present in the (reduced) strict consensus. Three of them are terminal taxa (Tax.1, Tax.2, and Tax.3) and one of them is a clade of two taxa (Tax.25 and Tax.26). All these branches are pruned from the set of MPTs for the next iteration.

After pruning these four unstable branches, the reduced strict consensus is still unresolved, denoting the complex nature of the topological differences among the MPTs (i.e. more than five unstable branches create this large polytomy). The PCR values of Iteration 2 indicate a single terminal taxon (Tax.12) as the most unstable branch (PCR = 0.50).

After pruning Tax.12 the reduced strict consensus has gained three new nodes (see Iteration 3 in Fig. 8). Two of these new nodes are dichotomous but one has five descendant branches collapsed into a polytomy (Node C in Fig. 8). This new polytomous node is not present in the original strict consensus and will be analysed separately of Node B during the following iteration (Iteration 3). After calculating the new set of PCR values a single most unstable branch (Tax.18) is detected for Node B. The set of PCR values calculated for the descendants of Node C shows all the branches tied (PCR = 0), denoting that the five descendant branches are equally unstable. Therefore, these five branches are selected for the subsequent character analysis (see below). In the analysed datasets, cases like this are frequently found for polytomies of degree 4–6.

Pruning Tax.18 results in a single new node in the reduced strict consensus (see Iteration 4, Node D in Fig. 8). The PCR values of Iteration 4 for Node B indicate Tax.9 as the most unstable descendant of this polytomy (PCR = 0.10), whereas the PCR value of Node D denotes Tax.20 as the most unstable branch (PCR = 0). Note that a set of PCR values is not calculated again for Node C as its descendants were all identified as unstable in the previous iteration.

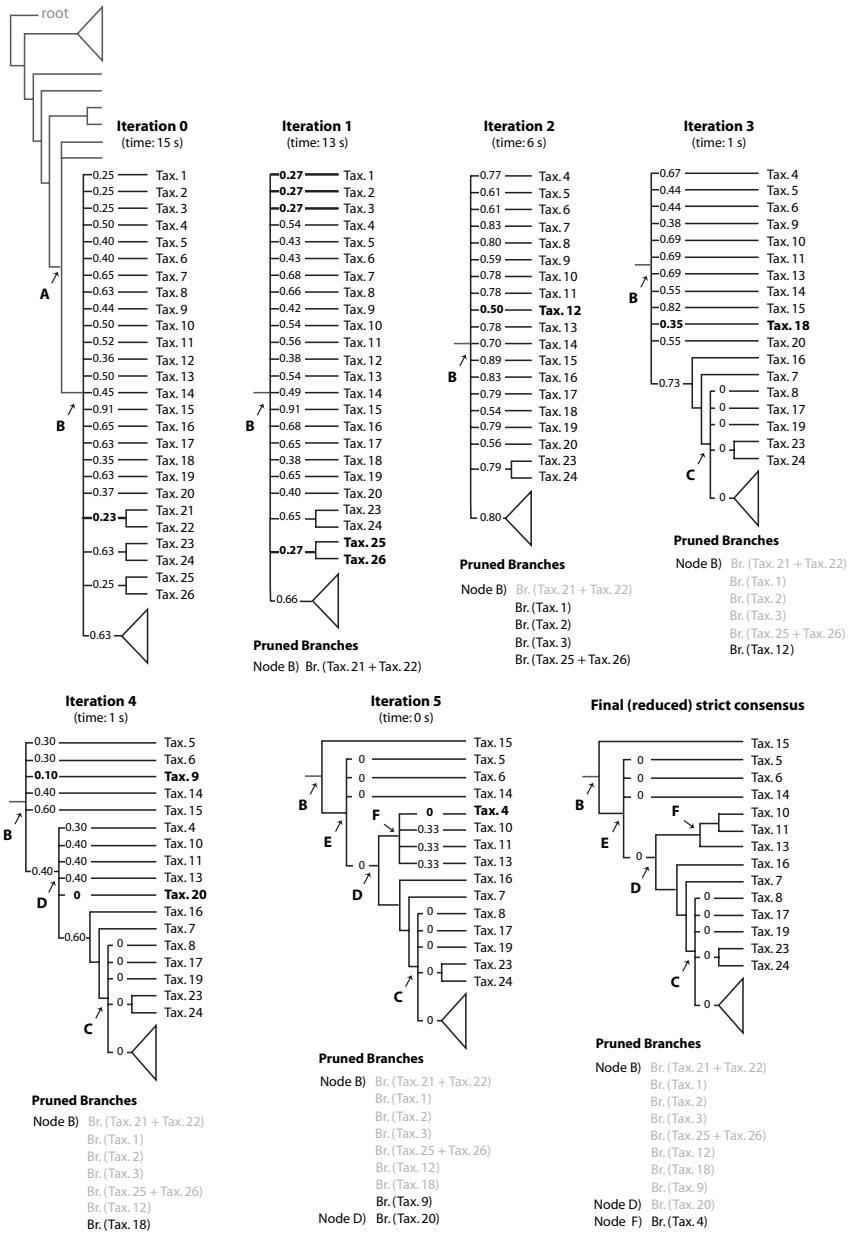


Fig. 8. Example of the iterative PCR procedure on the dataset of Smith and Stockey (2002). The outcome of each iteration of the iterative PCR is reproduced here from the output file of the script. The first tree (iteration 0) is the complete strict consensus showing the PCR values of the descendants of the large polytomy (Node B), whereas the following iterations show the reduced strict consensus obtained after pruning the most unstable branch identified in the previous iteration (with the new set of PCR values calculated for the descendants of each of the polytomies). The descendant branches with the lowest PCR value of each polytomy (in each iteration) are highlighted with a thick line and bold font. A list of the increasing set of pruned branches is listed at the bottom of each reduced consensus tree (the most recently pruned branches are marked in black). The computing time of each iteration is shown (in seconds).

Pruning Tax.9 results in the appearance of a new node (Node E) in the reduced strict consensus that is a direct descendant of Node B (see Iteration 5 in Fig. 8). Pruning Tax.20 also results in the appearance of a new node (Node F) in the reduced strict consensus. These two nodes (E and F) contain polytomies of degree 4, which are the only ones that need to be analysed in Iteration 5 (the only other polytomy is Node C). The PCR values of all the descendants of one of them (Node E) are equal to zero (terminal taxa 5, 6, 14, and Node D). The set of PCR values of the other node (F) indicates a single terminal branch (Tax.4) as the most unstable descendant and therefore it will be pruned for next iteration.

After pruning Tax.4, Node F is completely resolved and the iterative PCR procedure ends as there are no polytomies left to be analysed. The final reduced strict consensus is obtained after pruning all the unstable branches detected through the iterative PCR procedure (Fig. 8), except for those polytomies in which all the descendants have PCR equal to zero (e.g. descendants of Node C and Node E). These latter branches, although they are actually not pruned from the reduced consensus, will be subject of the subsequent character analysis as they have been identified as unstable (see below).

The computational cost of each iteration is directly related to the degree of the polytomies. In the example, the first three iterations took 95% of the total computing time, while the four last iterations took only 5% due to the successive reduction in the degree of the polytomies (i.e. the complexity of the problem). For instance, a total number of 220 triplets have to be evaluated for calculating the set of PCR values of Node B in Iteration 3, whereas only 40 triplets have to be evaluated in Iteration 4 (20 triplets for each of Nodes B and D). This drastic reduction is largely generated by the division of a node of degree 12 (Node B in Iteration 3) into two nodes of degree 6 (Nodes B and D in Iteration 4).

Detecting unstable taxa using different approaches could take equal or even less time for simple and small polytomies. However, in complex polytomies our script identifies unstable taxa several times faster than other approaches. The iterative PCR script takes 36 s to identify and prune 12 unstable taxa (eight terminal taxa and two clades of two taxa) in the example of Fig. 8. The command *prunnelsen* of TNT takes approximately 5000 s to obtain the reduced consensus that maximizes the number of nodes gained when up to 12 taxa are pruned. Note that the aim of these two methods is not the same and, therefore, these results are only compared in order to highlight why the use of the iterative PCR is preferred to the *prunnelsen* command for our particular purposes.

### *Evaluation of unstable branches*

As explained above, the unstable behaviour of a branch in a set of MPTs has traditionally been related to the amount of missing data. However, the instability of a branch can be caused by the lack of scorings in relevant characters (i.e. missing data) and/or the presence of character conflict (see Evaluating unstable taxa). Here we explore this general assumption by evaluating the causes underlying the unstable behaviour of branches in 19 datasets using the procedure described above and discuss the potential application of this information in empirical studies.

In contrast to this widespread assumption, 42 out of 125 unstable terminal taxa (34%) were identified as caused solely by the lack of scorings in relevant characters (i.e. exclusively due to their missing entries). A more common case (found in 42% of the unstable terminal branches) was the simultaneous presence of character conflict (i.e. some characters supporting some of their alternative positions) and the presence of missing data in relevant characters. A third case consists of terminal branches taking alternative positions that are supported by different sets of conflicting characters, but that lack critical cells with missing entries. Note that these taxa may have missing entries but these do not belong to characters with relevant information to decrease their instability. This case is poorly represented in our sample, being present in only 16 unstable terminal branches (13%). Note that most of the matrices used in this survey have a relatively high proportion of missing entries given that they usually contain fossil taxa. This is partly why the third type of case is not common among the analysed data matrices. It is likely that these cases would be more frequently found in datasets containing a much lower proportion of missing entries (e.g. containing well-preserved extant taxa).

Irrespective of the frequencies of the different cases, once the relevant characters are identified, the critical step is to dedicate efforts to revise the currently scored data or to score the missing entries that are likely to reduce the instability of the problematic taxa (or clades). The identification of characters related to the instability of a given branch is the final product of the TNT script and can be extremely useful for guiding further research in empirical phylogenetic analyses. In many of the datasets analysed here, the number characters identified as related to the instability of a given branch are $< 10\%$ of the total number of characters. In such cases, this information helps to markedly reduce the scope and focus of further research effort for solving the uncertain affinities of a particular taxon (or clade).

The identification of critical cells currently filled with missing entries can guide the collection of new data. This may involve either collecting more specimens or defining priorities for future research. In the case of morphological characters, researchers may conduct studies to score the unstable taxon for relevant characters that may require specific preparation (e.g. histological thin sections, SEM imaging, CT scanning). Some of these techniques may be destructive and/or expensive and therefore the researcher would be able to design a more efficient research effort for solving a particular phylogenetic problem. In the case of molecular datasets that include taxa represented by only some of the analysed genes, the proposed character analysis can also retrieve information for guiding further research. If the unstable behaviour of a branch is (at least partially) caused by the lack of information on a critical gene, the character analysis will probably retrieve groups of relevant characters with missing entries localized in such a gene. This can help to determine priorities on the

genes (or regions of DNA sequences) that need to be sequenced for refining the phylogenetic affinities of a given unstable taxon or clade.

The identification of characters that participate in the character conflict responsible for the alternative positions of a given unstable branch can lead to the revision of some of the hypotheses of homology represented in the data matrix. With morphological characters, this can lead to an examination of the scorings of the identified characters in order to re-evaluate the hypothesis of primary homology rejected in some of the optimal trees (Farris, 1983). Although homology statements and character definitions are recurrently revised during the development of morphological phylogenetic analyses, these types of revisions are time consuming as they involve examining the condition of multiple taxa as well as the definition of character states. Therefore, the identification of a reduced subset of characters to revise that may help to solve the uncertain affinities of a given taxon can result in a more efficient management of future research efforts. Fortunately, in the analysed datasets, usually fewer than ten characters of this type are detected for most of the unstable branches (out of the total number of characters, which are sometimes several hundreds). Finally, this type of revision can also lead to the discovery of influential but mistaken scorings, which can be more commonly found and more difficult to detect as the size of a morphological dataset increases (in datasets of moderate size, e.g. 200 characters of 50 taxa, a total of 10 000 cells have to be filled).

## Conclusions

The presence of unstable branches in phylogenetic analysis is problematic not only because they represent uncertainties on the affinities of some taxa, but also because it can produce collapsed strict consensus trees that obscure the otherwise stable relationships of the remaining taxa. Despite these problems there is no justification—either a *priori* or *a posteriori*—to definitively exclude unstable taxa from the data matrix as this involves the deletion of phylogenetic information that can be relevant (or even critical) for understanding the relationships of the entire group.

We agree with recent suggestions that the problem of unstable taxa should be focused on both their detection and the identification of the underlying causes of their instability (Kearney, 2002). Within this context, we propose here an iterative procedure based on the comparison of triplets to detect unstable branches (either terminal taxa or clades). These branches are then subjected to an automated examination of the alternative character optimizations in the MPTs, aiming to identify the relevant characters that cause the instability (by character conflict) or that can ameliorate the instability (by defining missing entries) of problematic branches. The empirical cases analysed here show that a combination of both cases (lack of information and conflict of characters) is most frequently the explanation for the behaviour of unstable taxa. This approach, focused on the causes of instability, can also lead to a more efficient use of time and resources while conducting further research efforts on the affinities of unstable branches.

Finally, although we have focused on the analysis of MPTs, a similar approach could be conducted on a set of suboptimal topologies for identifying unstable (low-support) branches and assessing the relevant characters that could diminish their instability and therefore guide future research toward a more robust understanding of their relationships.

## Acknowledgments

This contribution is based on the development of a script for TNT, a program made freely available thanks to a subsidy by the Willi Hennig Society. M. Kearney and M. Ramírez are thanked for providing some of the analysed datasets. P. Goloboff is thanked for comments on the manuscript and useful suggestions on the *IterPCR* script. Julian Faivovich and an anonymous reviewer are also thanked for their comments.

## References

Arnold, E.N., 1981. Estimating phylogenies at low taxonomic levels. Zeitschrift für Zoologische und Systematik Evolutionsforschung 19, 1–35.

Benton, M.J., 1990. The species of *Rhynchosaurus*, a rhynchosaur (Reptilia, Diapsida) from the Middle Triassic of England. Phil. Trans. Roy. Soc. Lond., Ser. B 328, 213–306.

Bryant, D.. 2003. A classification of consensus methods for phylogenetics. In: Janowitz, M., Lapointe, F.-J., McMorris, F.R., Mirkin, B., Roberts, F.S. (Eds.) BioConsensus, DIMACS. American Mathematical Society, Providence, RI, pp. 163–184.

Donoghue, M.J., Doyle, J.A., Gauthier, J., Kluge, A.G., Rowe, T., 1989. The importance of fossils in phylogeny reconstruction. Ann. Rev. Ecol. Syst. 20, 431–460.

Estabrook, G.F., 1992. Evaluating undirected positional congruence of individual taxa between two estimates of the phylogenetic tree for a group of taxa. Syst. Biol. 41, 172–177.

Estabrook, G.F., McMorris, F.R., Meacham, C.A., 1985. Comparison of undirected phylogenetic trees based on subtrees of four evolutionary units. Syst. Zool. 34, 193–200.

Farris, J.S.. 1983. The logical basis of phylogenetic analysis. In: Platnick, N., Funk, V. (Eds.), Advances in Cladistics 2. Columbia University Press., New York, pp. 7–36.

Gao, K., Norell, M.A., 1998. Taxonomic review of *Carusia* (Reptilia : Squamata) from the late Cretaceous of the Gobi Desert and phylogenetic relationships of anguimorphan lizards. Am. Mus. Novit. 3230, 1–51.

Gauthier, J.A., 1986. Saurischian monophyly and the origin of birds. Mem. Cal. Acad. Sci. 8, 1–47.

Gauthier, J.A., Kluge, A.G., Rowe, T., 1988. Amniote phylogeny and the importance of fossils. Cladistics 4, 105–209.

Goloboff, P.A., 1999. Analyzing large data sets in reasonable times: Solutions for composite optima. Cladistics 15, 415–428.
Goloboff, P., Farris, J., Nixon, K. 2008a. TNT: Tree Analysis Using New Technology, vers. 1.1 (Willi Hennig Society Edition). Program and documentation available at http://www.zmuc.dk/public/phylogeny/tnt.
Goloboff, P., Farris, J., Nixon, K. 2008b. TNT, a free program for phylogenetic analysis. Cladistics 24, 774–786.
Gordon, A.D., 1979. A measure of the agreement between rankings. Biometrika 66, 7–15.
Gordon, A.D. 1980. On the assessment and comparison of classifications. In: Tomassone, R. (Ed.), Analyse de données et informatique. INRIA, LeChesnay, pp. 149–160.
Grande, L., Bemis, W.E., 1998. A comprehensive phylogenetic study of amiid fishes (Amiidae) based on comparative skeletal anatomy. An empirical search for interconnected patterns of natural history. J. Vert. Paleont. 18, 1–690.
Grant, T., Kluge, A., 2003. Data exploration in phylogenetic inference: scientific, heuristic, or neither. Cladistics 19, 379–418.
Hillis, D.M., 1996. Inferring complex phylogenies. Nature 383, 130–131.
Hillis, D.M., 1998. Taxonomic sampling, phylogenetic accuracy, and investigator bias. Syst. Biol. 47, 3–8.
Hillis, D.M., Pollock, D.D., McGuire, J.A., Zwickl, D.J., 2003. Is sparse taxon sampling a problem for phylogenetic inference? Syst. Biol. 52, 124–126.
Kearney, M., 2002. Fragmentary taxa, missing data, and ambiguity: Mistaken assumptions and conclusions. Syst. Biol. 51, 369–381.
Kearney, M., Clark, J.M., 2003. Problems due to missing data in phylogenetic analyses including fossils: a critical review. J. Vertebr. Paleontol. 23, 263–274.
Nixon, K.C., 1999. The parsimony ratchet a new method for rapid parsimony analysis. Cladistics 15, 407–414.
Nixon, K.C., Wheeler, Q.D. 1993. Extinction and the origin of species. In: Novacek, M.J., Wheeler, Q.D. (Eds.), Extinction and Phylogeny. Columbia University Press, New York, pp. 119–143.
Novacek, M.J., 1992. Fossils, topologies, missing data, and the higher level phylogeny of eutherian mammals. Syst. Biol. 41, 58–73.
Poe, S., 1998. Sensitivity of phylogeny estimation to taxonomic sampling. Syst. Biol. 47, 18–31.
Pollock, D.D., Zwickl, D.J., McGuire, J.A., Hillis, D.M., 2002. Increased taxon sampling is advantageous for phylogenetic inference. Syst. Biol. 51, 664–671.
Rannala, B., Huelsenbeck, J.P., Yang, Z., Nielsen, R., 1998. Taxon sampling and the accuracy of large phylogenies. Syst. Biol. 47, 702–710.
Rowe, T., 1988. Definition, diagnosis, and origin of Mammalia. J. Vertebr. Paleontol. 8, 241–264.
Santini, F., Tyler, J.C., 2004. The importance of even highly incomplete fossil taxa in reconstructing the phylogenetic relationships of the Tetraodontiformes (Acanthomorpha: Pisces). Integr. Comp. Biol. 44, 349–357.
Smith, A.B. 1994. Systematics and the Fossil Record. Blackwell Science, Oxford, UK.
Smith, S.Y., Stockey, R.A., 2002. Permineralized pine cones from the Cretaceous of Vancouver Island, British Columbia. Int. J. Plant Sci. 163, 185–196.
Swofford, D.L. 2002. PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4. Sinauer Associates, Sunderland, Massachusetts.
Thorley, J.L. 2000. Cladistic Information, Leaf Stability and Supertree Construction. Ph.D. thesis. University of Bristol, U.K.
Thorley, J.L., Page, R.D.M., 2000. RadCon: phylogenetic tree comparison and consensus. Bioinformatics 16, 486–487.
Thorley, J.L., Wilkinson, M., 1999. Testing the phylogenetic stability of early tetrapods. J. Theor. Biol. 200, 343–344.
Thorley, J.L., Wilkinson, M. 2000. *The RadCon Manual 1.1.2*. Program and Documentation available at http://www.poisson consulting.ca/radcon/radcon.html.
Thorley, J.L., Wilkinson, M., Charleston, M.A. 1998. The information content of consensus trees. In: Rizzi, A., Vichi, M., Bock, H.-H. (Eds.), Advances in Data Science and Classification. Springer, Berlin, pp. 91–98.
Wheeler, W.C., Giribet, G., Edgecombe, G.D. 2004. Arthropod systematics. The comparative study of genomic, anatomical, and paleontological information. In: Cracraft, J., Donoghue, M.J. (Eds.), Assembling the Tree of Life. Oxford University Press, New York, pp. 281–295.
Wilkinson, M., 1994. Common cladistic information and its consensus representation: reduced Adams and reduced cladistic consensus trees and profiles. Syst. Biol. 43, 343–368.
Wilkinson, M., 1995. More on reduced consensus methods. Syst. Biol. 44, 435–439.
Wilkinson, M., 2003. Missing entries and multiple trees: instability, relationships, and support in parsimony analysis. J. Vertebr. Paleontol. 23, 311–323.
Wilkinson, M., Benton, M.J., 1995. Missing data and rhynchosaur phylogeny. Hist. Biol. 10, 137–150.
Zwickl, D.J., Hillis, D.M., 2002. Increased taxon sampling greatly reduces phylogenetic error. Syst. Biol. 51, 588–598.

## Supporting Information

Additional Supporting Information may be found in the online version of this article:

IterPCR.run
TNT script implementing the entire procedure described in this paper.

matrices-trees.zip
A ZIP file containing the 19 matrices and tree files used for the empirical analyses presented in this paper.

output_files.zip
A ZIP file containing all the output files generated by the IterPCR script during the analysis of the 19 empirical datasets.

**Appendix S1.** Implementation of TNT script
Electronic_Appendix_1_Implementation.doc. The complete procedure described in this paper has been implemented in a script for TNT. The structure of the IterPCR script is described in this appendix, providing detailed instructions for its use.

**Appendix S2.** Datasets
Electronic_Appendix_2_Datasets.doc. This appendix includes a short description of the 19 datasets used in the empirical part of this paper, and a short summary of the results of the IterPCR script of these files is provided.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.