

Problems with supertrees based on the subtree prune-and-regraft distance, with comments on majority rule supertrees

Pablo A. Goloboff* and Claudia A. Szumik

Unidad Ejecutora Lillo, Consejo Nacional de Investigaciones Científicas y Técnicas, Miguel Lillo 251, 4000 S.M. de Tucumán, Argentina

Accepted 12 December 2014

Abstract

This paper examines a recent proposal to calculate supertrees by minimizing the sum of subtree prune-and-regraft distances to the input trees. The supertrees thus calculated may display groups present in a minority of the input trees but contradicted by the majority, or groups that are not supported by any input tree or combination of input trees. The proponents of the method themselves stated that these are serious problems of “matrix representation with parsimony”, but they can in fact occur in their own method. The majority rule supertrees, being explicitly clade-based, cannot have these problems, and seem much more suited to retrieving common clades from a set of trees with different taxon sets. However, it is dubious that so-called majority rule supertrees can always be interpreted as displaying those clades present (or compatible with) with a majority of the trees. The majority rule consensus is always a median tree, in terms of the Robinson–Foulds distances (i.e. it minimizes the sum of Robinson–Foulds distances to the input trees). In contrast, majority rule supertrees may not be median—different, contradictory trees may minimize Robinson–Foulds distances, while their strict consensus does not. If being “majority” results from being median in Robinson–Foulds distances, this means that in the supertree setting a “majority” is ambiguously defined, sometimes achievable only by mutually contradictory trees.

© The Willi Hennig Society 2015.

Introduction

Supertrees are the equivalent of consensus trees when the input trees have different sets of taxa, and thus are useful to combine trees from data sets that cannot be combined in a supermatrix. If combined analysis is possible, neither supertrees nor consensus can truly be a replacement for it, which has been known since Barrett et al. (1991) showed that a group supported by each of the individual data sets may nonetheless be contradicted by the combined data set. However, even if supertrees cannot replace data set combination (as argued also by Gatesy et al., 2002; Goloboff and Pol, 2002; and Janies et al., 2013, among others), establishing the joint implications of sets of trees may be a valid goal in itself, and this obviously requires appropriate methods.

Methods for constructing supertrees are often studied by comparing the results when the input trees have identical taxon sets against the results of consensus methods. In one of the most important papers on supertrees, Cotton and Wilkinson (2007) proposed a method, majority rule supertrees, intended to behave as the majority rule consensus tree, and justified the intended behaviour on the grounds that.

The majority-rule consensus has proven particularly important because of its use in summarizing bootstrap or jackknife replicates (Felsenstein, 1985), quartet puzzling steps (Strimmer and von Haeseler, 1996), and Bayesian posterior probability distributions on trees. The majority-rule also seems quite natural when the input trees are inferred from independent data as is often the case in supertree construction (p. 445)

The justification invoked by Cotton and Wilkinson thus concerns a logical connection between those groups (clusters, monophyletic groups, higher taxa) in the result and those in the input trees. Other discus-

*Corresponding author:

E-mail address: pablogolo@yahoo.com.ar

sions of supertree methods based on their properties on groups and/or nestings can be found in Steel et al. (2000), Dong and Fernández-Baca (2009), and Dong et al. (2010a,b). In the case of majority rule consensus trees, being in more than half the input trees is a necessary and sufficient condition for the group to be in the result. Different connections between groups in the result and in the input trees can be established (e.g. as in Goloboff and Pol, 2002, which proposed a method intended to behave as the semi-strict supertrees), but a supertree analogue of the majority rule consensus is clearly the most appropriate in many contexts, as noted by Goloboff and Pol (2002), Goloboff (2005), and Cotton and Wilkinson (2007).

SPR supertrees

Other authors, instead, take a different approach to justifying a supertree method: the method is considered valid if it produces results that are closer to the “true” phylogeny. This justification can only be assessed in practice through simulations, and is always tied to the specific conditions of the simulations performed. A recent example of such an approach to justifying a supertree method is by Whidden et al. (2014), who proposed supertrees based on the subtree prune-and-regraft (SPR) distances (Hein et al., 1996). The SPR supertree is the binary tree with the minimum sum of SPR distances to the source trees (the source trees can have polytomies). SPR supertrees became possible with the work of Whidden and Zeh (2009) and Whidden et al. (2010, 2013), which made calculation of SPR distances orders of magnitude faster than previously possible (e.g. Nakhleh et al., 2005; Goloboff, 2008). The program SPRSupertrees (written by C. Whidden) implements the method and is available at <http://kiwi.cs.dal.ca/Software>.

Whidden et al. (2014) defended their method on the grounds that, when using it to combine source trees representing different genes with simulated lateral gene transfers (LGTs), it produces supertrees closer to the model tree than matrix representation with parsimony (MRP; Baum, 1992; Ragan, 1992; Baum and Ragan, 1993) and Robinson–Foulds (RF) supertrees (Bansal et al., 2010). We do not dispute the results of Whidden et al.’s simulations—it may well be true that under the specific conditions of LGT they simulated, SPR supertrees are more likely than MRP or RF supertrees to return the true tree. In addition to their simulations, based on the analysis of several real data sets, Whidden et al. also stated (p. 578) that their “results show that SPR is suited to a range of phenomena and not just LGT alone.” Given such optimistic statements, it can be anticipated that the method will be used under many situations regardless of whether LGT can be

assumed—indeed, C. Whidden (personal communication) states that the method is useful also for “hybridization, recombination, long branch attraction, or any other evolutionary scenario where specific bipartitions arise due to non-speciation events but ancestral and descendant bipartitions may still be attributed to speciation.”

Instead of considering the degree to which the method produces supertrees closer to the true tree (as Whidden et al. did), here we consider logical aspects of the method, and how it summarizes relationships from the source trees. As stated by Bansal et al. (2010, p. 10), any method that calculates a supertree so as to minimize some tree-distance measure (e.g. MRP, cliques, triplets, quartets, Robinson–Foulds, or SPR) is mathematically acceptable, and will find the supertree that is (for the distance measure used) as similar as possible to the original trees; furthermore, all those methods (or even the difference in results among them) can be useful in exploring the similarities and differences between the input trees. But this does not mean that any of those methods is equally useful for establishing systematic conclusions: relevant measures in this case must consider monophyletic groups (= clades). As Whidden et al. (2014, p. 566) themselves explicitly faulted MRP because it “can generate relationships that do not belong to any of the source trees or are contradicted by a majority of source trees,” the unwary reader might conclude that SPR supertrees lack those defects. That conclusion, however, would be incorrect. We show in the next section that SPR supertrees may display groups contradicted by each of the input trees and therefore cannot be interpreted in terms of clade monophyly or be expected to display groups present in the majority of the input trees.

SPR supertrees do not recover common clades

Whidden et al. (2014) compared their method with MRP and the RF supertrees of Bansal et al. (2010). The RF supertrees are a modification of Cotton and Wilkinson’s (2007) majority rule supertrees (which Whidden et al. do not cite). Whidden et al. conjectured that, because the RF distance (Robinson and Foulds, 1981) between the two trees resulting from moving a single taxon to a distant branch is very large (even when the two trees remain very similar), RF supertrees may be strongly affected by cases of LGT in some of the genes used to build the supertree.

When the aim of the supertree analysis is to identify groups supported by (or compatible with) the majority of the input trees, Whidden et al.’s method may produce unjustified results. Consider the case shown in Fig. 1, three trees that differ in the location of two taxa, X and Y. The first and last source trees differ in

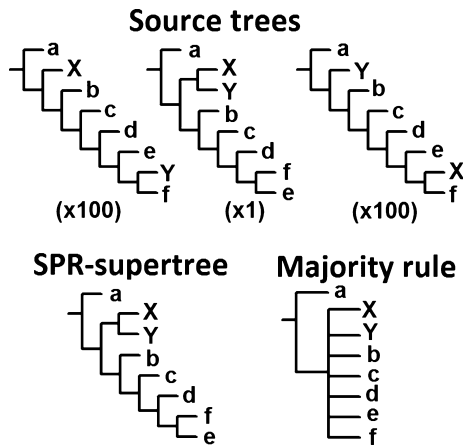


Fig. 1. A case with three types of topology in the input trees. For the SPR distance the middle tree is median between the other two. The SPR supertree is identical to the middle tree, even if there are 100 copies of the first and last input trees, and a single copy of the middle tree.

two SPR moves; the middle tree is median to the other two, differing in one move from each. Note that (with a single copy of each tree) every one of the groups is present in the same fraction of trees (just one tree, 0.333). However, a supertree identical to the middle tree will minimize the sum of SPR distances to all the source trees, and will be chosen. This will also be the case even when there are multiple copies (in equal numbers) of the first and last tree, with a single copy of the middle tree: the SPR supertree will continue being identical to the middle tree. When there are 100 copies of the first and third trees, most of the groups in the SPR supertree (again, identical to the middle tree) are present in only one of the 201 source trees. Furthermore, the group XY is present in the supertree, but 200 input trees display X and Y as far away from each other as is possible in a tree of this size. The middle tree is indeed median in terms of SPR distances, but very far from being so in terms of the groups or relationships displayed.

Because minimizing the sum of SPR distances to the source trees has no special meaning in terms of the groupings, the SPR supertree can also display clades not present in any of the source trees, and contradict groups present in the majority of the source trees—both of which were explicitly considered by Whidden et al. (2014) as problematic in the case of MRP. Whidden et al. (2014) did not discuss whether SPR supertrees could have this behaviour, but Figs 2 and 3 show examples of these two problems.

Figure 2 shows the case of three source trees; the unique binary tree that minimizes the sum of SPR distances displays two groups (*degh* and *dge*, marked in Fig. 2) that are not present in any of the source trees—each of the individual source trees contradicts these two groups.

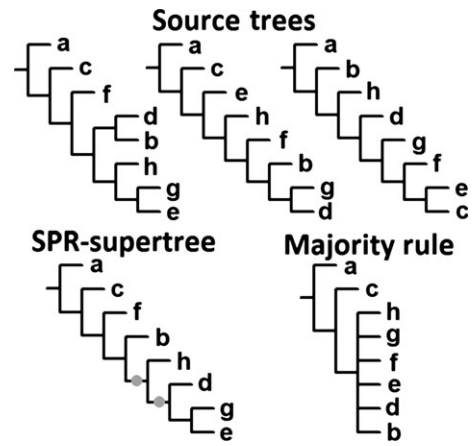


Fig. 2. A case where the SPR supertree displays two groups (marked) contradicted by each of the input trees.

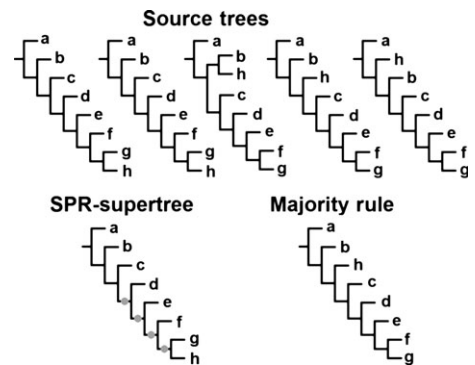


Fig. 3. A case where the SPR supertree displays four groups (marked) supported by two of the five input trees and contradicted by three. The SPR supertree places taxon *h* deeply nested at the tip of the tree, but *h* is located near the base in the majority of the trees.

Figure 3 shows a case with five source trees; the first two are identical, while all the others differ in the location of taxon *h*. The SPR distance for all tree-pairs is 1, except the first two trees (zero moves apart). Note that the groups *fg*, *efg*, *defg*, and *cdefg* are present in three of the five trees, while incompatible groups *gh*, *fgh*, *efgh*, and *defgh* are in only two of the five. Choosing any of the unique trees as supertree implies a sum of SPR distances equal to 4 (one move to each of the other trees), while choosing the tree in double copy implies a sum equal to 3 (zero moves to the other copy, and one move to each of the three remaining trees). Therefore, the supertree is identical to the first tree, which displays four groups present in the minority of the trees and contradicted by the majority. Note that with more taxa and more trees, if the non-duplicate trees correspond to more alternative positions of the floating taxon *h*, the frequency of those four groups (still present in the supertree) can be arbitrarily close to zero (and the frequency of the contradictory

groups *fg*, *efg*, *defg*, and *cdefg* arbitrarily close to unity).

For examples similar to those in Fig. 3, one might try to justify the SPR supertree on the grounds that *h* is detected as a taxon of uncertain placement, and the most common placement is as sister to *g*, which is ignored by the majority rule consensus. C. Whidden (pers. comm.) used this argument to justify the results of SPR supertrees, and considered that this indicates that “the majority rule is problematic” because “naively combining compatible bipartitions containing incompatible subtrees may result in such problematic inferences.” That argument does not really justify the SPR supertree, however. First, the notion of “placement” is ambiguous—clades or groups of taxa are a well-defined concept, but “placement” of a taxon in certain “regions” of trees cannot be characterized in a general way, except in the particular circumstance of identical relationships for the *other* taxa (as in Fig. 3). Second, in the case Fig. 3 that justification is contradicted *even* if an intuitive characterization of the “regions” where taxon *h* can be located is accepted: taxon *h* is located near the base in three of the trees, and at the tip of the tree in only two. The majority rule consensus is then correct in placing *h* near the base of the tree, but the SPR supertree places it at the tip. With more taxa, and the floating taxon switching position to a more distant group, this difference can be made much more pronounced (given that every SPR move has the same cost, regardless of how far it moves the taxon), but Fig. 3 serves to illustrate the effect.

The results for Figs 2 and 3 were confirmed using Whidden et al.’s program (the most recent version, 1.2.1). The results shown are obtained only when the option *bipartition_cluster* is set to more than 0.6. With the default 0.5 for this option, the search for the SPR supertree will ignore “rearrangements that violate bipartitions supported by 0.5 of gene trees.” Whidden et al. (2014, p. 569) made clear that this is intended only as a heuristic to facilitate the search for optimal supertrees. The default value of 0.5 precludes cases like those in Figs 2 and 3, but at the expense of not allowing full minimization of the SPR distances to source trees. In other words, while the default 0.5 produces more reasonable results, it sets a limit on the degree to which the criterion proposed by Whidden et al. (2014) can be optimized. Thus, unrestricted optimization of the criterion—SPR distances to the source trees—leads to results that are hard to defend.

Note that the examples in Figs 1 and 3 would also produce the same results if a “TBR supertree” were applied. TBR distances are generally harder to calculate than SPR distances, but the trees in those examples differ only in the position of terminal taxa (so that their interconversion requires the same number of

TBR and SPR moves). Therefore, finding the tree that minimizes the sum of TBR distances to the source trees produces exactly the same results for those examples.

An additional problem with SPR supertrees concerns a point that, on first approach, seems merely technical, but in fact also concerns unsolved methodological issues. The SPR supertree is defined by Whidden et al. (2014) as the binary tree that minimizes the sum of SPR distances; computations are much easier on binary trees. Of course there may be multiple equally optimal binary trees, or a single optimal supertree but supported only weakly, for example in the case of conflicting input trees (as in the examples presented here), and then a less resolved supertree may be desirable. C. Whidden (personal communication) stated that the ideal situation is “that users of the software be aware of these concerns and apply tests such as bootstrapping before establishing systematic conclusions.” However, there are two problems with this. First, it is unclear how to perform the bootstrapping—should one resample the input trees themselves and re-infer the supertree for each pseudoreplicate? Or the original matrices and re-infer input trees? Second, once the resampling method is decided, there is still the problem of summarizing the results of the pseudoreplicates. In phylogenetics, bootstrapping assesses support of clades based on their frequencies¹—the more frequent the group, the higher its “support”—but C. Whidden himself considers that the majority rule tree is “problematic.” When bootstrapping SPR supertrees, one would thus be left with no method to summarize the results.

Majority rule supertrees

The only method that truly behaves like a majority rule consensus tree for identical taxon sets is Cotton and Wilkinson’s (2007) majority rule supertree, as it is based on optimizing the same metric, the symmetric difference or RF distance (Robinson and Foulds, 1981). For full taxon sets, given the definition of RF distances, it can be deduced (Barthélemy and McMorris, 1986) that the least resolved (i.e. most polytomous) tree which minimizes $\sum \text{RF}$ to source trees with identical taxon sets must have (i) any group present in the majority of the source trees as present, and (ii) any group not present in the majority of the source trees as absent, thus being necessarily identical to the majority rule consensus tree. Condition (ii) requires that polytomous trees be considered as well: $\sum \text{RF}$ may be minimized only by trees with

¹Goloboff et al.’s (2003) GC is also based on clade frequencies. Holmes (2003) presented some alternative views that, to our knowledge, have never been used or implemented.

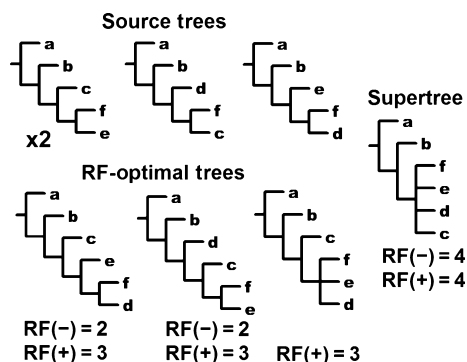


Fig. 4. A case where the majority rule supertree of Cotton and Wilkinson (2007) (i.e. the strict consensus of the trees minimizing the sum of RF distances to the input tree) does not itself minimize the sum of RF distances, in either the (–) or the (+) versions. This can happen only when the input trees have different taxon sets; for identical taxon sets, the strict consensus of all the trees which minimize the sum of RF distances is always also a median tree. Under the (–) method, two trees are RF-optimal (with $\sum\text{RF} = 2$); under the (+) method, the same two trees plus a partly polytomous one are RF-optimal (with $\sum\text{RF} = 3$). The majority rule supertree has a $\sum\text{RF} = 4$ in both the (–) and the (+) cases.

polytomies. This contrasts with the case of phylogenetic tree searches, where a polytomous tree can never have a better score (under either maximum-likelihood or maximum-parsimony) than any of its possible resolutions, and makes it difficult to calculate supertrees using tree-search algorithms on binary trees.

Cotton and Wilkinson (2007) proposed their method, in part, as a response to Goloboff and Pol's (2002) claim that an analogue of the majority rule consensus tree for the supertree case may not be possible. Dong et al. (2010b, p. 3–4) even suggested that “a strict generalization of majority-rule consensus might not be the ideal approach for supertree construction” because “different trees ‘specialize’ in different groups of taxa, leaving other groups largely unresolved or unrepresented”—which would be better summarized, in their view, by a supertree including those groups that are supported more frequently than they are contradicted (as in the frequency difference consensus of Goloboff et al., 2003), instead of just including groups in more than half the trees.

Goloboff and Pol's argument regarding the difficulty of a supertree analogue of the majority rule was that when trees have different taxon sets, the trees which support a given group without conflict (if combined with some trees) may be the same trees which contradict the group without conflict (if combined with other trees), and therefore an individual tree cannot be said either to support or to contradict a group, unless all the taxa included and excluded from the group are present in the tree. As a consequence, different contradictory supertrees may at the same time be compatible

with the majority of the input trees (see Goloboff and Pol, 2002; fig. 10). Cotton and Wilkinson (2007) cited that argument, but did not contradict it; they instead proposed to seek trees of minimum $\sum\text{RF}$, based on Barthélemy and McMorris' (1986) demonstration that (for identical taxon sets)

“The majority-rule consensus tree minimizes the sum of the symmetric difference metric ... between it and each of the input trees. The majority-rule consensus tree is thus shown to be a median of the input trees with respect to the (full-split) symmetric-difference metric.” (Cotton and Wilkinson, 2007, p. 446).

Cotton and Wilkinson (2007) then proposed the strict consensus of the tree(s) minimizing $\sum\text{RF}$ as a formal equivalent of the majority rule tree in the supertree setting. They proposed two types, majority rule (–) and (+), based on two different ways to calculate the RF distance between trees with different taxon sets. The (–) is based on removing from the candidate supertree the taxa absent from an input tree; the (+) is based on grafting onto the input tree each of the missing taxa and refining all polytomies into the possible binary trees (the distance is calculated as the shortest between the candidate supertree and one of the trees in the set derived from grafting and refining). Subsequently, Dong and Fernández-Baca (2009) proposed two variants of the (+), called (+)s and (+)g. The (+)s is the same as Cotton and Wilkinson's (+) but retaining the polytomous trees as well (instead of restricting the set to the binary trees), and the (+)g is based on grafting only (without refining polytomies). Given that the RF distance to input trees may well be minimized only by polytomous trees (see above), it seems more reasonable to retain polytomous trees derived from the grafting of missing taxa. Dong and Fernández-Baca (2009) also proved that only the variants (–) and (+)g produce the same results as the majority rule consensus for trees with equal taxon sets.

The majority rule supertrees of Cotton and Wilkinson (2007) have several desirable properties (as shown by Dong and Fernández-Baca, 2009). Being explicitly clade-based, the approach is among the most reasonable in the supertree setting, from a taxonomic perspective. However, there are two problems with the idea that such supertrees are indeed “majority” rule.

The first problem is that Cotton and Wilkinson (2007) did not show that minimizing $\sum\text{RF}$ to trees of different taxon sets would *also* produce supertrees with groups present in the majority of the trees—the frequency of such groups apparently cannot be counted, based on Goloboff and Pol's (2002) examples. Majority rule consensus trees are not used by taxonomists because they minimize $\sum\text{RF}$; instead, they are useful *because they display the groups present in the majority of the source trees.*

The second problem is that, even if what makes a tree “majority” is the fact that it minimizes $\sum\text{RF}$, input trees with different taxon sets do not behave in the same way as with identical sets of taxa. For identical taxon sets, the strict consensus of all the trees minimizing $\sum\text{RF}$ is itself of minimum $\sum\text{RF}$; in other words, the majority rule consensus is the *least* resolved among all trees of minimum $\sum\text{RF}$ (Dong and Fernández-Baca, 2009; p. 361). For different taxon sets, instead, several trees may imply the minimum possible $\sum\text{RF}$ while their strict consensus has a higher $\sum\text{RF}$. When analysed under the (–) option, this is the case for Fig. 1 of Cotton and Wilkinson (2007), which reproduces the example of Goloboff and Pol (2002, fig. 10) (a set of four trees, (a (ed)), (e(cd)), (d(bc)), and (c(ab))). In that example, each of the four combinations of three trees is RF-optimal (with $\sum\text{RF} = 2$),² as well as the additional partly resolved tree (ed(abc)), but the strict consensus of those five optimal trees has $\sum\text{RF} = 4$. Although the majority rule (+) supertree for that example is itself of minimum $\sum\text{RF}$, there are other cases for which neither the majority rule (–) nor (+) is RF-optimal. Consider the example of Fig. 4, with four input trees which produce the same supertree under either (–) or (+). The $\sum\text{RF}$ of the optimal trees is 2 under (–) and 3 under (+); the majority rule supertree has a $\sum\text{RF} = 4$ under either (–) or (+). Note that because the source trees contain no polytomies, the (+)g modification (Dong and Fernández-Baca, 2009; based on not refining the polytomies of the input trees) is irrelevant in this case and produces the same result as the original method of Cotton and Wilkinson (2007). If the condition of “majority” results from a tree being of minimum $\sum\text{RF}$, then the strict consensus of the trees with minimum $\sum\text{RF}$ does not represent a “majority.” Only the individual trees of minimum $\sum\text{RF}$ can be said to be “majority,” and they contradict each other. In other words, for unequal taxon sets, the minimization of $\sum\text{RF}$ is ambiguous and not uniquely defined—as Goloboff and Pol (2002) had claimed.

Discussion

The examples in Figs 1–3 show that SPR supertrees may produce results that are hard to justify, from the point of view of the clades obtained, even in cases where the source trees have identical taxon sets. Whidden et al. (2014) speculated that cases of taxa placed in distant parts of the trees (attributable, in their context, to LGT) might have an adverse effect on supertrees based on RF distances, given that the RF

distance between trees with a single taxon placed in two distant parts of the tree is very large (despite the trees being very similar). However, the way in which SPR supertrees summarize otherwise similar trees but with a few taxa in different positions (such as Figs 1 and 3) seems even more inappropriate, at least for the purpose of identifying monophyletic groups in common. When the input trees are very different (as in Fig. 2), the SPR supertree may display groups that are neither present in nor implied by any input tree or combination of input trees—these spurious groups are widely recognized by most authors (Whidden et al., 2014 included!) as the most serious problem of MRP.

It remains to be seen whether, in the presence of extensive LGT, the SPR supertree is indeed most likely to be the “true” phylogeny. Of course, a problem with extensive LGT or hybridization is that they challenge the very notion of choosing a single tree—when a species is thought to have exchanged numerous genes, or hybridized, with several other taxa, choosing a unique location for that species may not be the best way to represent that situation (as Whidden et al., 2014 themselves admitted, p. 578).

Taxa that can switch positions in multiple trees may also decrease resolution of consensus trees, and a number of heuristic methods have been proposed to identify taxa (or clades) responsible for the loss of resolution in that case (e.g. Wilkinson, 1994, 1995; Aberer et al., 2013; and a number of methods implemented in TNT, see Goloboff et al., 2008; p. 782; the method of Pol and Escapa, 2009, iterPCR, was also implemented in TNT subsequent to 2008). Ignoring floating taxa when calculating consensus or supertrees solves the problem of long LGTs mentioned by Whidden et al. (2014), diminishes the need for alternative supertree methods, and is perhaps a more appropriate way to represent cases of LGT and hybridization (as the jumping taxa are not placed in any definite position). However, in the case of supertrees, the caveat still remains that when some taxa are pruned from the source trees the way in which the *other* taxa are related may change (as exemplified by Goloboff and Pol, 2002, fig. 12; this can never happen in the consensus setting).

The problems with SPR supertrees, and the fact that the problems with wildcard taxa they attempt to correct can be easily solved by pruned (super) trees, suggest that majority rule supertrees are much better justified for summarizing results. Unfortunately, no computer program allows direct calculation of majority rule supertrees. For this paper, we used TNT scripts (available from the first author) for the (–) and (+)g options, based on enumerating each possible (binary and polytomous) candidate supertree and calculating $\sum\text{RF}$ for each; this enumerative approach is extremely slow and practical only for very low

²As in Cotton and Wilkinson (2007, p. 449), these RF values are not normalized, because that would attribute different weights to splits from trees of different size.

numbers of taxa (eight or nine at the most). Dong et al. (2010b) described an integer linear programming implementation for majority rule supertrees, but their implementation is not available and—due to the nature of the formulation—it can handle only modestly sized problems. Bansal et al. (2010) proposed heuristics for RF supertrees (i.e. the *binary* trees with minimum Σ RF), which may be useful in approximating the majority rule supertree, although their suggestion to use greedy contractions of those binary trees to obtain the majority rule supertree is problematic (see below).

The only conceptual caveat regarding majority rule supertrees is that the degree to which they can be considered to truly display clades in the majority of the input trees is uncertain. In this regard, it is surprising that despite the meticulous studies previously published on properties and calculation of majority rule supertrees (e.g. Cotton and Wilkinson, 2007; Dong and Fernández-Baca, 2009; Dong et al. 2010a), none of these papers has discussed the fact that the primary goal of so-called “majority” rule supertrees, the minimization of Σ RF, is not necessarily achieved by the strict consensus of the individual trees of minimum Σ RF, in either the (–) or the (+) versions. This implies that, if “the majority-rule consensus tree minimizes the sum of the symmetric difference metric” (Cotton and Wilkinson, 2007), then majority rule supertrees are not an equivalent, because they do not minimize that sum. Another implication is that, although Bansal et al.’s (2010) heuristic algorithms for searching binary RF supertrees may be important in helping to find the majority rule supertree, their suggestion (Bansal et al., 2010, p. 11) that the majority rule supertree be obtained by the greedy contraction of groups from one of the binary RF supertrees will fail in many cases. This greedy contraction (accepting all contractions that improve or match the Σ RF) always produces the correct results for identical taxon sets, but when taxon sets differ, calculating the majority rule supertree requires finding all individual trees of minimum Σ RF and then calculating their strict consensus.

Acknowledgements

We thank S. Arias, R. Beiko, S. Catalano, D. Fernández-Baca, and M. Wilkinson for comments and discussion. Mark Holder, and reviewers C. Whidden, K. Cranston and Anonymous provided useful comments when processing an earlier version of the manuscript, which led to significant improvement. We also acknowledge funding from CONICET (PIP 112 201101 00687 to P.A.G.), and a grant (to J. Cracraft and L. Lohman, entitled “Assembly and evolution of the Amazonian biota and its environment: an inte-

grated approach”), from NSF, NASA, and Fundação de Amparo à Pesquisa do Estado de São Paulo.

References

- Aberer, A., Krompas, D., Stamatakis, A., 2013. Pruning rogue taxa improves phylogenetic accuracy: an efficient algorithm and webservice. *Syst. Biol.* 62, 162–166.
- Bansal, M., Burleigh, J.G., Eulenstein, O., Fernández-Baca, D., 2010. Robinson–Foulds supertrees. *Algorithms Mol. Biol.* 5 (18), 1–12.
- Barrett, M., Donoghue, M.J., Sober, E., 1991. Against consensus. *Syst. Zool.* 40, 486–493.
- Barthélemy, J.P., McMorris, F.R., 1986. The median procedure for n -trees. *J. Classif.* 3, 329–334.
- Baum, B., 1992. Combining trees as a way of combining data sets for phylogenetic inference, and the desirability of combining gene trees. *Taxon* 41, 3–10.
- Baum, B., Ragan, M., 1993. A reply to A.G. Rodrigues. A comment on Baum’s method for combining phylogenetic trees. *Taxon* 42, 637–640.
- Cotton, J.A., Wilkinson, M., 2007. Majority-rule supertrees. *Syst. Biol.* 56, 445–452.
- Dong, J., Fernández-Baca, D., 2009. Properties of majority rule supertrees. *Syst. Biol.* 58, 360–367.
- Dong, J., Fernández-Baca, D., McMorris, F.R., Powers, R., 2010a. Majority-rule (+) consensus trees. *Math. Biosci.* 228, 10–15.
- Dong, J., Fernández-Baca, D., McMorris, F.R., 2010b. Constructing majority-rule supertrees. *Algorithms Mol. Biol.* 5 (2), 1–16.
- Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783–791.
- Gatesy, J., Matthee, C., DeSalle, R., Hayashi, C., 2002. Resolution of a supertree/supermatrix paradox. *Syst. Biol.* 51, 652–664.
- Goloboff, P., 2005. Minority rule supertrees? MRP, compatibility, and minimum flip may display the least frequent groups. *Cladistics* 21, 282–294.
- Goloboff, P., 2008. Calculating SPR distances between trees. *Cladistics* 24, 591–597.
- Goloboff, P., Pol, D., 2002. Semi-strict supertrees. *Cladistics*, 18, 514–525.
- Goloboff, P., Farris, J., Källersjö, M., Oxelmann, B., Ramírez, M., Szumik, C., 2003. Improvements to resampling measures of group support. *Cladistics* 19, 324–332.
- Goloboff, P., Farris, J.S., Nixon, K.C., 2008. TNT, a free program for phylogenetic analysis. *Cladistics* 24, 774–786.
- Hein, J., Jiang, T., Wang, L., Zhang, K., 1996. On the complexity of comparing evolutionary trees. *Discrete Appl. Math.* 71, 153–169.
- Holmes, S.P., 2003. Bootstrapping phylogenetic trees: theory and methods. *Stat. Sci.* 18, 241–255.
- Janies, D., Studer, J., Kandelman, S., Linchangco, G., 2013. A comparison of supermatrix and supertree methods for multilocus phylogenetics using organismal datasets. *Cladistics* 29, 560–566.
- Nakhleh, L., Ruths, D., Wang, L.-S., 2005. RIATA-HGT: a fast and accurate heuristic for reconstructing horizontal gene transfer. In: Wang, L. (Ed.), *Proceedings of the Eleventh International Computing and Combinatorics Conference (COCOON 05)*. Lecture Notes in Computer Science, LNCS no. 3595, Springer, Berlin, pp. 84–93.
- Pol, D., Escapa, I., 2009. Unstable taxa in cladistic analysis: identification and the assessment of relevant characters. *Cladistics* 25, 515–527.
- Ragan, M., 1992. Phylogenetic inference based on matrix representation of trees. *Mol. Phylogenet. Evol.* 1, 51–58.
- Robinson, D.F., Foulds, L.R., 1981. Comparison of phylogenetic trees. *Math. Biosci.* 53, 131–147.
- Steel, M., Dress, A., Bocker, S., 2000. Simple but fundamental limitations on supertree and consensus methods. *Syst. Biol.* 49, 363–368.

- Strimmer, K., von Haeseler, A., 1996. Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* 13, 964–969.
- Whidden, C., Zeh, N., 2009. A unifying view on approximation and FPT of agreement forests. In: WABI 2009. LNCS, Springer-Verlag, Berlin, Vol. 5724, pp. 390,401.
- Whidden, C., Beiko, R.G., Zeh, N., 2010. Fast FPT algorithms for computing rooted agreement forests: theory and experiments. In: Festa, P. (Ed.), *Experimental Algorithms. Lecture Notes in Computer Science*. Springer, Berlin Heidelberg, Vol. 6049, pp. 141–153.
- Whidden, C., Beiko, R.G., Zeh, N., 2013. Fixed-parameter algorithms for maximum agreement forests. *SIAM J. Comput.*, 42.4, 1431–1466.
- Whidden, C., Zeh, N., Beiko, R.G., 2014. Supertrees based on the subtree prune-and-regraft distance. *Syst. Biol.* 63, 566–581.
- 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.