Building Supertrees: An Empirical Assessment Using the Grass Family (Poaceae)

Nicolás Salamin,¹,² Trevor R. Hodkinson,¹ and Vincent Savolainen²

¹Department of Botany, University of Dublin, Trinity College, Dublin 2, Ireland; E-mail: salaminv@tcd.ie, trevor.hodkinson@tcd.ie
²Molecular Systematics Section, Jodrell Laboratory, Royal Botanic Gardens, Kew, Richmond Surrey, TW9 3DS, United Kingdom; E-mail: v.savolainen@rbgkew.org.uk

Abstract.—Large and comprehensive phylogenetic trees are desirable for studying macroevolutionary processes and for classification purposes. Such trees can be obtained in two different ways. Either the widest possible range of taxa can be sampled and used in a phylogenetic analysis to produce a “big tree,” or preexisting topologies can be used to create a supertree. Although large multigene analyses are often favored, combinable data are not always available, and supertrees offer a suitable solution. The most commonly used method of supertree reconstruction, matrix representation with parsimony (MRP), is presented here. We used a combined data set for the Poaceae to (1) assess the differences between an approach that uses combined data and one that uses different MRP modifications based on the character partitions and (2) investigate the advantages and disadvantages of these modifications. Baum and Ragan and Purvis modifications gave similar results. Incorporating bootstrap support associated with pre-existing topologies improved Baum and Ragan modification and its similarity with a combined analysis. Finally, we used the supertree reconstruction approach on 55 published phylogenies to build one of most comprehensive phylogenetic trees published for the grass family including 403 taxa and discuss its strengths and weaknesses in relation to other published hypotheses.

[Matrix representation with parsimony; phylogeny; Poaceae; supertree.]

Comprehensive and well-resolved phylogenetic trees that contain, when possible, estimates of divergence dates underpin comparative biology and allow powerful tests of a wide range of hypotheses (Felsenstein, 1985; Harvey and Pagel, 1991; Pagel, 1999). Phylogenetic trees representing a large sample of taxa are also preferred to those based on limited numbers of taxa for classification purposes and for studying character evolution.

Two different approaches can be used to obtain comprehensive phylogenetic trees. The first uses characters gathered from the widest possible range of taxa directly in an analysis to produce a “big tree.” In this approach, phylogenetic analyses of true biological characters, either molecular or phenotypic (e.g., morphological), are performed, and the meaning of the evolutionary hypotheses underlining these characters can be interpreted. The second is the meta-analysis approach (Arnqvist and Wooster, 1995) used in supertree-building methods. The underlying idea of these methods is to combine the topologies (or source trees) resulting from multiple phylogenetic studies (Sanderson et al., 1998), rather than their respective biological data sets, to produce a supertree. Matrix elements derived directly from these published topologies, and for which no real phylogenetic interpretation can be obtained, represent the characters that are used to build the supertree or composite phylogenetic tree (Bininda-Emonds and Bryant, 1998). Although consensus techniques also work on topologies to produce a summarized phylogenetic tree, supertree reconstruction has the advantage of not requiring identical terminal taxa sets. Only overlapping sets are needed, which allows the method to produce more comprehensive phylogenies than the original ones.

A recent theoretical study (Hillis, 1996) suggested that large phylogenetic trees can be easier to analyze than previously thought, and empirical analyses have demonstrated that large, combined, multigene analyses can correctly infer large trees (Soltis et al., 1999, 2000; Savolainen et al., 2000). Because “big trees” may not be as easy to construct as suggested by these studies (Kim, 1996; Purvis and Quick, 1997a, b; Yang and Goldman, 1997) and because combinable data are not always available, the use of composite phylogenies to study evolutionary patterns is one of the few choices left to the investigators. However, choosing between the various supertree methods is not a straightforward task.

Two basic algorithms have been developed for supertree reconstruction. The strict
Supertree reconstruction method (Steel, 1992; Constantinescu and Sankoff, 1995) directly assembles the topologies of the source trees into a supertree; compatible trees (i.e., trees without conflicting nodes) are required as input. Incompatible trees cannot be incorporated in the analysis, a problem to which no solution has yet been found (Steel et al., 2000). This limitation precludes its use with real data sets, where incompatible nodes, reflecting either real evolutionary divergence between data sets (hard incongruence) or simply random error because of limited sample size (soft incongruence), are often the rule between different trees.

In contrast, matrix representation with parsimony (MRP; Baum, 1992; Ragan, 1992) can be used whether or not the source trees are compatible. MRP uses additive binary coding (Farris et al., 1970) to represent the hierarchical structure of trees as a series of matrix elements (Baum and Ragan, 1993). Every node on each source tree is represented by a binary matrix element, and parsimony analysis of the matrix is used to retrieve the tree or trees that represent the hierarchical information in the source trees. If multiple most-parsimonious composite trees are obtained, a strict or semistrict consensus can be used to generate a consensus composite tree (Baum, 1992; Ragan, 1992).

Several coding procedures have been proposed for the MRP method. Baum (1992) and Ragan (1992) independently first proposed that terminal taxa delimited by each node should be coded as 1 in the binary matrix, and all other taxa as 0. Missing taxa from individual source trees are then coded as missing values (typically, ?) for the matrix elements representing these trees. Purvis (1995a) argued that the elements derived from source trees lack independence and hence add redundant information to the matrix. He proposed removing this apparent redundancy by allocating the value 0 only to taxa within the immediate sister group to the particular clade under consideration and by assigning missing values to the other taxa of the source tree (i.e., coding them as ? instead of 0). Subsequently, Ronquist (1996) suggested that the bias would not be associated with redundant information, as stated by Purvis (1995a), but with the relative sizes of the source trees. He argued that the difference in the amount of information contributed by each source tree could be removed by inversely weighting each tree according to the number of its nodes. However, he favored other weighting schemes that were based on the support for nodes, which, he argued, would also compensate for any size bias.

Bininda-Emonds and Bryant (1998) discussed some of the properties of MRP and investigated modifications to the method. Little is known about the merits of such modifications or how they perform with real data. Bininda-Emonds and Bryant (1998) also discussed the issue that supertrees obtained from MRP are not always congruent with those based on an approach that uses combined data and suggested that different weighting schemes might help MRP better approximate the combined result. Furthermore, matrix elements represent membership (character state coded as 1) or lack of membership (character state coded as 0) of a particular taxon relative to a clade. Allowing reversals in the parsimony analyses can therefore produce cladets in the composite tree that are supported by a lack of membership in some components of the source trees. Bininda-Emonds and Bryant (1998) advocated using irreversible character states in a parsimony analysis to overcome this shortcoming.

In this paper, we report the results of an empirical study aimed at assessing the relative merits of the supertree approach, using the grass family as a case study. We used the combined molecular and morphological data set from the Grass Phylogeny Working Group (GPWG, 2000) to compare the various proposed modifications of the MRP methods, to investigate the effect of irreversible characters on supertree reconstructions, and to evaluate the differences between the supertrees we obtained and an approach using combined data. Accurate and meaningful comparisons between the combined analysis and the supertree approach were made by breaking down the GPWG combined data set into its eight character partitions and then rebuilding a phylogenetic tree by using the various MRP modifications. We also used the same MRP modifications to produce different large supertrees that include 403 genera from 55 published phylogenies. We discuss the strengths and weaknesses of each MRP modification obtained in
relation to other evolutionary hypotheses concerning the grass family.

**MATERIALS AND METHODS**

**Combined Phylogenetic Tree**

Using maximum parsimony, we reanalyzed the GPWG data set (GPWG, 2000), which combined for 61 species of grasses eight different data sets, representing molecular as well as morphological data, and defined well-supported clades according to bootstrap percentages (Felsenstein, 1985). One thousand bootstrap replicates were performed (with the tree bisection–reconnection swapping algorithm, random addition of taxa, and keeping 20 trees at each step) by using PAUP* 4.08b (Swofford, 1998). The GPWG matrix was also divided into its eight data partitions representing three plastid gene sequences (ndhF, rbcL, and rpoC2), three nuclear DNA regions (gbss1 and phyB genes and the 5.8S and internal transcribed spacer 2 of the nrDNA, hereafter termed ITS), plastid restriction site variations, and morphological data (Table 1). Parsimony analyses performed on each character set were followed by 1,000 bootstrap replicates with the same heuristic search options as previously described.

**MRP Reconstructions**

Two different types of MRP analyses were performed for the grass family. For the two sets of supertree reconstructions, we built five binary matrices, using the program SuperTree 0.85b (available at http://www.tcd.ie/Botany/NS/SuperTree.html). First, the eight different bootstrap trees obtained from each data partition of the GPWG matrix were used as source trees for supertree reconstructions [referred to hereafter as GPWG supertree(s)] with the same 61 terminal taxa as in the GPWG combined tree. Then, we used the information reported in 55 publications in supertree reconstructions (referred hereafter as large supertrees) to produce matrices of 403 genera (from a total of 635 for the whole family; Mabberley, 1993). The publications chosen did not represent an exhaustive sample of the published literature concerning the grass family, but that was not the goal of our study. Given the homogeneity of characters and methods used in the reviewed publications (indicated with an asterisk in the References list), we made no distinction between each of the source trees. Our decision followed Bininda-Emonds et al. (1999) and Purvis (1995b), who found that differential weighting of the source trees according to data or tree search method (or both) had little impact on the composite phylogenetic tree.

Supertree reconstruction requires an overlap of taxa sampling between source trees. Because of the large size of the family (about 10,000 species), however, few taxa were in common between published studies. To overcome this problem, we considered only generic names in the large analyses. Species were used when evidence against the monophyly of the genus was demonstrated in the published study. Three taxonomic groups—poooids, panicoids, and chloridooids—are consistently found as strong monophyletic groups, and no evidence has ever contradicted this view (Clark et al., 1995; Duvall and Morton, 1996; Mason-Gamer et al., 1998; Gaut et al., 1999; Hilu et al., 1999; Hsiao et al., 1999; Mathews et al., 2000). To ease the heuristic searches, we constrained our large analyses of the 55 published phylogenetic trees by forcing each of these three groups to form three monophyletic clades but allowed other taxa to be inserted within them (the “backbone” option in PAUP* 4.08b).

Because bootstrap percentages are missing in most studies before 1993, or because values <50% are not specified for most of the published trees, we used a transformed function of the bootstrap percentages to weight the characters into the coded matrices in both analyses. All percentages <50%, or nodes with missing values, were given the weight of 1. Bootstrap

### Table 1. Details of the eight character partitions of the GPWG data set.

<table>
<thead>
<tr>
<th>Data type</th>
<th>Genome</th>
<th>Number of characters</th>
<th>Number of taxa</th>
</tr>
</thead>
<tbody>
<tr>
<td>ndhF</td>
<td>plastid</td>
<td>2186</td>
<td>51</td>
</tr>
<tr>
<td>rbcL</td>
<td>plastid</td>
<td>1344</td>
<td>30</td>
</tr>
<tr>
<td>rpoC2</td>
<td>plastid</td>
<td>855</td>
<td>29</td>
</tr>
<tr>
<td>phyB</td>
<td>nuclear</td>
<td>1182</td>
<td>39</td>
</tr>
<tr>
<td>gbss1</td>
<td>nuclear</td>
<td>773</td>
<td>14</td>
</tr>
<tr>
<td>ITS</td>
<td>nuclear</td>
<td>424</td>
<td>42</td>
</tr>
<tr>
<td>Restriction sites</td>
<td>plastid</td>
<td>364</td>
<td>45</td>
</tr>
<tr>
<td>Morphology</td>
<td>N/A</td>
<td>52</td>
<td>61</td>
</tr>
</tbody>
</table>
percentages $\geq 50\%$ were weighted by using an exponential transformation (James S. Farris, pers. comm.):

$$e^{[\log(a) \times (b/100)]}$$

where $a$ represents the weight assigned to 100% of node support, and $b$ represents the bootstrap percentage. This transformation attempts to overcome the conservative bias found in bootstrap percentages. Indeed, Hillis and Bull (1993), using simulations, showed that bootstrap values $> 70\%$ in their study usually indicated, with $> 95\%$ probability, that the corresponding clade was real. Corrections to bootstrap estimates have been proposed (Rodrigo, 1993; Zharkikh and Li, 1995; Effron et al., 1996), but these corrections cannot compensate for lack of information in large analyses (Sanderson and Wojciechowski, 2000). As the number of taxa increases, the relative number of characters available declines, and phylogenetic accuracy suffers. Our transformation thus attempts to linearize the bootstrap percentages, allowing for a gradual increase in the character weights in subsequent parsimony analyses.

For every supertree reconstruction, heuristic searches under maximum parsimony were performed with Baum/Ragan and Purvis coding schemes alone (hereafter BR-alone and PU-alone, respectively). We also used bootstrap percentages weighting schemes as described above on both coding schemes (hereafter BR + bootstrap and PU + bootstrap, respectively) and the Baum/Ragan coding scheme weighted by the inverse of the number of nodes present in each source tree (hereafter BR + nodes). We performed 1,000 replicates of random addition sequence, using the nearest-neighbor interchange swapping algorithm and keeping only 20 trees at each replicate. The stored trees were then swapped with the tree bisection–reconnection swapping algorithm, and a maximum of 1,000 trees were kept (the “maxtrees” option in PAUP*4.08b).

Two different parsimony analyses were performed in each of the cases described above: one considering characters as unordered (Fitch parsimony; Fitch, 1971), and one considering the same characters as irreversible (Camin–Sokal parsimony; Camin and Sokal, 1965); the result was a total of ten GPWG supertrees and ten large supertrees. Equally most-parsimonious solutions were summarized by using semistrict consensus.

**Topological Comparisons**

Two incongruence indices, expressed as distances, were used to compare the different topologies obtained. The partition metric (PM; symmetric difference in PAUP*4.08b; Robinson and Foulds, 1981) and the agreement subtree metric (D1; Gordon, 1980) were calculated by using PAUP*4.08b (Swofford, 1998). For the comparisons between the combined analysis and the GPWG supertree reconstructions, we performed Kishino–Hasegawa (KH) test (Kishino and Hasegawa, 1989) and the Shimodaira–Hasegawa (SH) test (Shimodaira and Hasegawa, 1999). The GPWG data set consisted of DNA sequences, restriction site data, and morphological characters (Table 1), and both tests were performed under the maximum parsimony criterion. Two sets of $P$-values were thus calculated—one according to the default options present in PAUP*4.08b (for the KH-test; Swofford, 1998), the other as described in Shimodaira and Hasegawa (1999) to ensure the validity of the test with topologies specified a posteriori. The latter test involved the creation of 500 bootstrapped replicates of the GPWG matrix, followed by the optimization of 1,000 random trees as well as the trees under consideration on these bootstrapped matrices (for details see Shimodaira and Hasegawa, 1999; Goldman et al., 2000).

**Results**

**GPWG Combined Tree**

The GPWG combined tree (Fig. 1) is one of the most comprehensive phylogenetic hypotheses concerning the grass family. Several features can be emphasized and will serve as a reference for the supertree comparisons. First, and as proposed by GPWG (2000), a large clade (PACCAD clade; Fig. 1) composed of six subfamilies—panicoids, arundinoids, chloridoinds, centothecoids, aristids, and danthonioids—forms a highly supported monophyletic group (100% bootstrap; Fig. 1). No bootstrap values $> 52\%$ support any group within the PACCAD clade. Outside the PACCAD...
clade, the bambusoids, ehrhartoids, and the pooids form another group called the BEP clade supported by a lower bootstrap value (87% bootstrap; Fig. 1). Finally, three clades—the anomochloids (Anomochloa, Streptochaeta), pharoids (Pharus), and puelioids (Guaduella, Puelia)—form the basal taxa of the Poaceae (early diverging lineages; Fig. 1).
TABLE 2. Partition metrics for the five MRP modifications, expressed as distances. A) Comparisons within modifications with unordered characters, B) within modifications with irreversible characters, and C) between modifications with unordered characters and modifications with irreversible characters.

<table>
<thead>
<tr>
<th>A) Unordered characters</th>
<th>BR-alone</th>
<th>BR + bootstrap</th>
<th>BR + nodes</th>
<th>PU-alone</th>
<th>PU + bootstrap</th>
</tr>
</thead>
<tbody>
<tr>
<td>BR-alone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BR + bootstrap</td>
<td>0.457</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BR + nodes</td>
<td>0.379</td>
<td>0.388</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PU-alone</td>
<td>0.440</td>
<td>0.362</td>
<td>0.422</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PU + bootstrap</td>
<td>0.474</td>
<td>0.379</td>
<td>0.457</td>
<td>0.362</td>
<td></td>
</tr>
</tbody>
</table>

B) Irreversible characters

<table>
<thead>
<tr>
<th>B) Irreversible characters</th>
<th>BR-alone</th>
<th>BR + bootstrap</th>
<th>BR + nodes</th>
<th>PU-alone</th>
<th>PU + bootstrap</th>
</tr>
</thead>
<tbody>
<tr>
<td>BR-alone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BR + bootstrap</td>
<td>0.431</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BR + nodes</td>
<td>0.216</td>
<td>0.491</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PU-alone</td>
<td>0.293</td>
<td>0.431</td>
<td>0.216</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PU + bootstrap</td>
<td>0.414</td>
<td>0.379</td>
<td>0.440</td>
<td>0.379</td>
<td></td>
</tr>
</tbody>
</table>

C) Irreversible characters

<table>
<thead>
<tr>
<th>C) Unordered characters</th>
<th>BR-alone</th>
<th>BR + bootstrap</th>
<th>BR + nodes</th>
<th>PU-alone</th>
<th>PU + bootstrap</th>
</tr>
</thead>
<tbody>
<tr>
<td>BR-alone</td>
<td>0.000</td>
<td>0.431</td>
<td>0.216</td>
<td>0.293</td>
<td>0.414</td>
</tr>
<tr>
<td>BR + bootstrap</td>
<td>0.457</td>
<td>0.112</td>
<td>0.552</td>
<td>0.491</td>
<td>0.457</td>
</tr>
<tr>
<td>BR + nodes</td>
<td>0.379</td>
<td>0.483</td>
<td>0.319</td>
<td>0.328</td>
<td>0.483</td>
</tr>
<tr>
<td>PU-alone</td>
<td>0.440</td>
<td>0.457</td>
<td>0.500</td>
<td>0.422</td>
<td>0.474</td>
</tr>
<tr>
<td>PU + bootstrap</td>
<td>0.474</td>
<td>0.422</td>
<td>0.517</td>
<td>0.457</td>
<td>0.181</td>
</tr>
</tbody>
</table>

GPWG Supertrees

The proportion of nonidentical nodes between the ten MRP reconstructions taken by pairs was calculated by using the PM (Table 2). Analyses using unordered characters resulted in topologies with pairwise distances of between 0.362 and 0.474 (mean distance: 0.412; Table 2). Pairwise comparisons for each MRP modification using irreversible characters gave distances of between 0.216 and 0.491 (mean distance: 0.369; Table 2). Values for each unordered Baum/Ragan modification are closer to their irreversible counterparts than to any other method (mean distance unordered BR/irreversible BR: 0.328; mean distance unordered BR/irreversible PU: 0.411; Table 2). The same pattern is found for the unordered Purvis modifications (mean distance unordered PU/irreversible PU: 0.386; mean unordered PU/irreversible BR: 0.468; Table 2). Three modifications gave very similar topologies with both types of characters. BR-alone gave the exact same topology, whereas the two BR + bootstrap and the two PU + bootstrap values were very close to each other (0.112 and 0.181, respectively; Table 2).

For comparisons between the GPWG combined tree and the GPWG supertrees, D1 distances ranged from 0.32 for BR + bootstrap with irreversible characters to 0.59 for PU-alone with unordered characters; the other eight modifications had values from 0.42 to 0.52 (Fig. 2). When considering distances obtained with PM, BR + bootstrap with irreversible characters again gave the topology the most similar to the GPWG combined tree, yielding a distance of 0.25. The worst modification was the BR + nodes with irreversible characters, for which the distance was 0.53, whereas those in the eight other modifications ranged from 0.41 to 0.50.

We further examined the placement of the major subclades defined in the GPWG combined tree (Fig. 3). Two modifications, BR + bootstrap and PU + bootstrap, both with irreversible characters, are the only modifications to yield the same basal branching pattern as the combined GPWG tree (Fig. 3A). In all the other modifications involving irreversible characters, two basal grasses, *Guaduella* and *Puelia*, are inserted inside the BEP clade. BR-alone placed the pooids as sister group to the PACCAD clade, while BR + nodes and PU-alone placed a clade of the pooids and the ehrhartoids as sister group of the PACCAD clade (Fig. 3A). With unordered characters, and except for the BR-alone modification, which was identical to the BR-alone with irreversible characters modification, all modifications had an odd basal branching pattern, the outgroup being inserted between two clades of the early diverging lineages (Fig. 3B).
The two different implementations of the KH-test were used to estimate the validity of the ten GPWG supertrees as possible alternative hypothesis to the GPWG combined tree (Table 3). Based on the characters of the GPWG combined data set, the classical KH-test rejected all the GPWG supertrees \((P < 0.001)\) as suitable alternative phylogenetic trees. However, the modified KH-test found only the PU-alone with unordered characters significantly different \((P < 0.05)\) from the GPWG combined tree (Table 3).

None of the other GPWG supertrees could be rejected, BR + bootstrap with both characters types being once again the modification giving the topology closest to that of the GPWG combined topology \((P = 0.588\) and \(P = 0.624\) for unordered and irreversible characters, respectively).

Finally, Figure 4 shows the supertree obtained with BR + bootstrap and irreversible characters. The discrepancies between this supertree and the GPWG combined tree were small and concerned mainly single taxa that were always positioned close to the clade to which they belonged in the GPWG tree (Fig. 4). The main differences concerned the placement of the ehrhartoids and bambusoids, which exchanged their position as sister group of the pooids in the supertree in comparison with the GPWG tree, and the placement of the three subfamilies arundinoids, aristidoids, and danthonioids.

**Large Supertree**

We analyzed the 55 publications the same way as the eight character partitions from the GPWG combined data set. For clarity, we do not show the results of BR + nodes, a MRP modification that resulted in topologies incompatible with the placement of subfamilies suggested in the GPWG combined tree. When irreversible characters were used (Fig. 5A), BR-alone, BR + bootstrap, and PU + bootstrap obtained the same basal branching pattern, but the placement of subfamilies inside the PACCAD clade was slightly different in each case (data not shown). PU-alone was the only modification to place the three subfamilies of the BEP clade together as a sister group to the PACCAD clade, agreeing with the GPWG supertrees and the GPWG combined tree. With unordered characters (Fig. 5B), BR-alone, BR + bootstrap, and PU-alone obtained the same basal branching pattern, with the BEP clade being paraphyletic. PU + bootstrap gave an odd combination, in which the early diverging lineages were embedded inside the bambusoids at the base of the tree (Fig. 5B).

**DISCUSSION**

The information given by the two incongruence indices helped pinpoint the differences present in the topologies obtained from the ten GPWG supertrees built using the
Figure 3. Summary of the subfamilial relationships in the GPWG supertrees obtained with the five MRP modifications with column A) characters considered as irreversible and column B) characters considered as unordered.

Eight character partitions from the combined data set. Analyses using unordered characters produced supertrees that were less similar, as measured by D1, to the combined tree than those obtained in analyses using irreversible characters. The only exception is for BR-alone, which obtained the same topology with both character types. However, unordered characters produced slightly better topologies than irreversible characters according to the PM incongruence index. D1 is defined as the number of taxa needed to be removed from both trees to obtain an identical subtree, whereas PM is the number of taxa bipartitions found between the two trees being compared (Johnson and Soltis, 1998). Having higher D1 values than PM values suggests that the topological differences do not involve single taxa (in which case, D1 would be smaller than PM), but rather a large set of taxa, placed differently in the supertree and the GPWG combined tree. Similar values of D1 and PM would suggest that a few discrepant taxa were involved, none being far from their proper locations. The D1 distances were much...
greater with unordered than irreversible characters and were much closer to the PM ones obtained with irreversible characters. Following this logic, we can emphasize that using unordered characters resulted in topologies with more differences from the GPWG combined tree as to the placement of larger subclades than did the use of irreversible characters. The type of characters (i.e., unordered or irreversible) used to reconstruct the supertrees did not greatly influence the placement of the terminal taxa in the GPWG supertrees. Therefore, topologies obtained by considering matrix elements as irreversible characters were closer to the combined phylogenetic tree represented by the GPWG combined analysis. This effect is clearly visible in the placement of the major subclades defined in the GPWG combined tree (Fig. 4). Unordered characters produced topologies in which the early diverging lineages were split into two groups—one including Puelia and Guaduella and one including Anomochloa, Streptochaeta, and Pharus. This extremely odd and unrealistic pattern does not correspond to any published phylogenetic trees concerning the grass phylogeny. Bininda-Emonds and Bryant (1998) found only a minor impact of irreversible characters compared to unordered characters in their analysis, but we found that topologies created by using irreversible characters are closer to an approach using combined data than are the topologies created by using unordered characters. The resolution in the large supertrees was extremely dependent on the type of characters used to perform the analysis; such was not the case with the GPWG supertrees, where both types of characters gave similar resolution (data not shown). Using irreversible characters produced only one or two large supertrees, depending on the modification used; in contrast, the “maxtrees” option in PAUP*4.08b was always reached when unordered characters were used. Of course, using irreversible characters should put more constraints on the parsimony analysis by preventing the reversion from state 1 to 0 (thus reducing the number of most-parsimonious trees), but it is not clear why such a drastic difference appeared only between the large supertrees.

The Purvis coding scheme has been proposed as an improvement to the Baum/Ragan procedure as an attempt to reduce the dependency and redundancy between elements coming from the same topology (Purvis, 1995a; Ronquist, 1996; Bininda-Emond and Bryant, 1998). However, our results suggest that the Purvis coding scheme does not have a large impact on the MRP reconstructions and that the Baum/Ragan method works just as well or even better. Comparisons between the approach using combined data and BR-alone, or combined data and PU-alone, gave similar values for both PM and D1 (Fig. 2). Comparisons between MRP modifications also indicated a close relationship between the two modifications, especially with irreversible characters (Table 2). This conclusion is less well supported with unordered characters, because BR-alone is the only modification to
FIGURE 4. Grass supertree obtained with Baum and Ragan coding scheme with bootstrap support weighting using separate phylogenies built from the eight character partitions of the GPWG combined data set as source trees. Dashed lines indicate incongruent branches between the supertree and the GPWG combined tree.
Figure 5. Summary of the subfamilial relationships in the large supertrees based on 55 published source trees (including 403 genera) obtained using the five MRP modifications with column A) characters considered as irreversible and column B) characters considered as unordered.

give a basal topology very different from that obtained with any other modification (Table 2 and Fig. 2). Ronquist's (1996) proposition to weight each character in the binary matrix by the inverse of the number of nodes in the corresponding source tree does not perform as well as the other modifications: It yields the greatest partition metric distance and the second largest agreement subtree distance (Fig. 2). Although the two alternatives to the Baum/Ragan method are based on logical and plain arguments (i.e., nonindependence and redundancy of matrix elements, and impact of larger trees), their effects on MRP reconstruction are not obvious and do not result in topologies closer to our combined reference. Indeed, PU-alone with unordered characters was rejected as a suitable alternative hypothesis to the GPWG combined tree with the modified KH-test and gave the second smallest $P$-value when used with irreversible characters (Table 2). However, conclusions from the large supertree analysis differed, because PU-alone was the only modification to produce the same basal branching pattern as the GPWG combined tree (Fig. 5B), whereas BR-alone, with the pooid sister group to the PACCAD clade, corresponded to an alternative hypothesis for the grass family that is supported by some data from nuclear ($Adh$; Gaut et al., 1999) and plastid ($rbcL$; Duvall and Morton, 1996) genomes. Because no reference phylogenetic trees are available that contain a number of taxa similar to those of our large supertrees, it is difficult to assess their topologies and to determine whether PU-alone with irreversible characters really gave a more accurate phylogeny of large grasses than did the other methods. The placement of the pooids is possibly not the best criterion by which to judge the methods;
the evolutionary position of this subfamily within the grass family differs as inferred with different DNA sequence data sets.

Weighting by node support, as suggested by Ronquist (1996), improved the fit between the additive binary matrices and the GPWG combined tree. The beneficial impact of the bootstrap weighting scheme is evident from the increased $P$-values obtained with the modified KH-test (Table 2). This trend is visible for the Baum/Ragan coding scheme, but is less obvious with the Purvis coding scheme (Fig. 2). Bootstrap weighting was proposed (Ronquist, 1996) as an alternative to the Purvis coding scheme both to reduce the greater impact of large source trees over smaller ones and to improve the effect of well-supported nodes in the MRP analysis. It is not clear how bootstrap weighting could reduce the impact of larger trees if the smaller trees do not have much greater support values than those of the larger trees. However, weighting of characters by bootstrap support within the Purvis coding scheme appears to be a redundant procedure, unable to greatly improve the MRP reconstruction from the important information present in bootstrap support values (Fig. 2). In the large supertree analysis, $PU + bootstrap$ with unordered characters produced unrealistic topologies with the early diverging lineages embedded in the bambooids (Fig. 5B). The Purvis coding schemes removes important restrictive information from the matrix (Bininda-Emonds and Bryant, 1998); possibly, therefore, weighting by bootstrap support would randomly assign high values to characters in the matrix where this restrictive information has been removed and would prevent the weighting schemes from being as effective as with the Baum/Ragan coding schemes. Weighting the matrix elements by node support also poses problems. Not all bootstrap analyses can be considered identical, and the number of replicates, the type of searches done, or both, will influence the support found. Moreover, when node support is not provided for a source tree, weights cannot be assigned, and information from this tree will be down-weighted in the MRP analysis. The effect of this is great in supertree reconstructions done with older publications, which in general do not report node support and are hence down-weighted in the MRP analyses.

An important aspect affecting the large supertrees that is not resolved in our comparisons is the placement of some rarely sampled taxa. taxa that are present in only a few publications have been allocated a high proportion of missing character values, which makes parsimony analyses much more difficult. How many of these taxa are misplaced is difficult to determine because no reference phylogeny for the 403 genera is available. MRP reconstruction done without constraining three clades (panicoids, pooids, and chloridoids) to be monophyletic ended with rarely sampled taxa being scattered all around the supertrees (data not shown). That can be an important problem when MRP analyses are performed on a wide taxonomic group comprising large numbers of rarely sampled taxa.

MRP coding, especially when weighted by node support, can be considered an indication of the signal in the primary data, with each node represented by one (weighted) synapomorphy (see Bininda-Emonds et al., 1999). However, we stress that supertrees cannot be viewed and interpreted in exactly the same way as phylogenetic analyses based on biological characters. One major problem is the difficulty of assigning node support for supertrees. Bootstrapping and other resampling procedures cannot be applied, given the clear nonindependence of the characters present in the binary matrix, and Bremer support (Bremer, 1988) is certainly equally affected by such nonindependence (as is the parsimony itself; hence, the different MRP modifications proposed). Moreover, branch lengths associated with supertrees are treated in the same way as morphological branch lengths, and unlike molecular branch length, they cannot be used to establish divergence times. This weakness is even more significant in taxonomic groups, such as the grasses, for which dates are difficult to gather from other sources (e.g., fossil records).

**CONCLUSIONS**

We have demonstrated that supertrees offer an easy way of producing phylogenies with a large number of taxa and that they can give good estimates of relationships within these groups (in this case, the grass family). Supertrees based on the eight character partitions of the GPWG data sets were found to roughly match the combined analysis, the discrepancies of the best MRP
modifications being found mainly in weakly supported branches of the combined tree. The Baum/Ragan and the Purvis modifications gave similar results, but incorporating bootstrap support associated with preexisting topologies improved the Baum/Ragan modification. Moreover, supertrees can be useful for comparative studies (Purvis, 1995b; Bininda-Emonds et al., 1999), whether of adaptation, coevolution, rates of evolution, cospeciation, or rates of effective cladogenesis, where accuracy in the branch length of phylogenetic reconstructions is not the primary problem (Purvis et al., 1994). Supertree reconstructions are also a useful way to help highlight poor taxonomic sampling and identify where previous studies are inconsistent. They can therefore be used as an exploratory tool capable of developing new hypotheses and indicating where future research should be focused.

ACKNOWLEDGMENTS

We are grateful to Elizabeth Kellogg and Lynn Clark for providing the GPWG matrix and for their comments on the manuscript, James Farris for helpful discussion regarding bootstrap weighting schemes during the XIX Hennig Society meeting, Stephen Renvoize and Mark Chase for helpful discussion on the project, Olaf Bininda-Emonds for his comments on the manuscript, and Jonathan Davis for the testing of the SuperTree 0.85b program. This study was supported by the Irish Higher Education Authority, the Roche Research Foundation (99-186), Enterprise Ireland (IC/2001/29) the Agassiz Foundation, and the Swiss NSF (31-52378).

REFERENCES

(References marked * were used as sources of phylogenies for the supertree reconstructions.)


Received 2 March 2001; accepted 26 September 2001

Associate Editor: Mike Steel