A note on the phylogenetic signal in current morphomatrices that combined fossil and extant seed plants

Author: Guido W. Grimm

Affiliation: Swedish Museum of Natural History, Department of Palaeobotany, P.O.Box 50007, 10405 Stockholm

Correspondance to: guido.grimm@nrm.se

Abstract: The phylogenetic signal comprised in several morphological matrices compiled by different authors to put forward Hilton 2006 Friis 2007 Crepet 2010 or criticise evolutionary hypotheses Rothwell 2009 Crepet 2010 is explored using a multitude of approaches. The matrices in the original publications are used to infer phylogenetic trees and networks that fulfil one of three commonly applied optimality criteria: maximum likelihood (ML), maximum parsimony (MP), and the least-square distance criterion (LS; via neighbour-joining, NJ, and neighbour-network, NN, algorithms). Support of alternative, partly competing phylogenetic split patterns is compared based on bootstrap (ML, MP, LS/NJ) and jackknife analyses (ML, MP, LS/NJ) and Bayesian posterior probabilities (PP) in order to explore the consistency and robustness of the phylogenetic signal in the matrices. In contrast to what was claimed in the some of the original papers, all matrices favour similar relationships between the larger groups of seed plants. As a trend, the older matrices Hilton 2006, Friis 2007 with many binary characters or characters with only a few states appear to be more decisive, and indicate relationships that find ample support from different approaches. The newer matrices Rothwell 2009 Crepet 2010 one the other hand, being composed of a larger number of characters with many states, provide generally weaker and more ambiguous phylogenetic signals. This results in the phenomenon that relationships proposed by the strict-consensus trees of their few equally parsimonious trees are usually not supported by further analyses.

The use of comprehensive phylogenetic analyses is advocated.

Introduction

Despite the availability of a vast amount of molecular data certain aspects of plant evolution are still obscure. Different filtered sets of molecular data have provided highly supported, however, incongruent relationships regarding the position of the Gnetales, the root of the angiosperms, the root of all seed plants and, linked with this, the question whether the ‘gymnosperms’ (cycads, confers, *Gingko*, Gnetales) are paraphyletic or monophyletic in a Hennigian sense (see Mathews 2009 Mathews 2010 for a review and compilation of relevant literature). Because of the inconsistence, and partly insufficiency, of molecular-based reconstructions, some researchers kept alive the tradition of assembling morphological matrices as another source of phylogenetic information. In particular within the last years, relatively large morphological matrices have been compiled and made available to the scientific community that also include numerous fossil taxa and are thought to cover the whole bandwidth of seed plants (e.g., Hilton Bateman 2006 Doyle 2006 Rothwell 2009). Using the criterion of maximum parsimony to find the most optimal placement of individual taxa or fossils (e.g. Doyle 9999 Doyle Endress 2010 Friis 2009) or the most parsimonious tree inferred from the character matrix Hilton Bateman 2006 Rothwell 2009 Crepet Stevenson 2010, the hope is that the mysteries of seed plant evolution and, in particular, the origin of angiosperms can be solved, or at least, gaps and uncertainties left by the molecular-based reconstructions can be filled.

Unfortunately, the most parsimonious trees (MPT) based on these morphological matrices often contradict the alternative but consistent, typically well-supported molecular-inferred relationships (see Mathews 2009 Mathews 2010 for a review). For instance, although all recent molecular studies and maximum likelihood (ML)-based topological tests reject the so-called "anthophyte hypotheses", i.e. a sister relationship between angiosperms and Gnetales Mathews 2010, both groups are consistently placed in the same subtrees if morphological matrices are used (e.g. Doyle 2006 Hilton Bateman 2006 Friis 2007 Rothwell 2009, Crepet Stevenson 2010). However, the signal of the used morphological matrices is often too weak or internally too incompatible to establish high branch support despite the number of recorded characters Hilton Bateman 2006 Friis 2007 2009 Sareela 2007. This may explain why some researchers of seed plant evolution (e.g. Rothwell 2009 Crepet & Stevenson 2010) refrain from establishing any branch support at all. Instead, the aim seems to be to optimize the matrix so that the number of reconstructed, equally optimal (most parsimonious) trees (MPT) is minimized, hence, a more or less resolved strict consensus tree can be obtained. This is then used as the basis to draw evolutionary conclusions or as a "cladistic test" Crepet Stevenson 2010 Rothwell 2009, p. 296. Analyses leading to numerous MPT, and accordingly collapsed strict consensus trees (e.g. Friis 2007), are rejected as uninformative. Evidence from differential support values (e.g. using nonparametric bootstrapping, Felsenstein 1985) and significant support from filtered matrices that minimize the amount of missing data at the cost of included characters are considered not relevant (see Rothwell 2009's critique of Friis 2007).[1]

---

[1] Rothwell et al. (2009, p. 299) "intentionally omitted support values from the nodes of all trees" for "the reasons explained by Rothwell & Nixon (2006)" (i.e. GW Rothwell, KC Nixon, Int. J. Plant Sci., vol. 167, pp. 737–749). The reasons given there are (p. 739) that "… support values, whether low or high for particular groups, would only mislead the reader into believing we [Rothwell and Nixon] are presenting a proposed phylogeny for the groups in question. Differences among most-parsimonious trees are sufficient to illuminate the points we wish to make here, and support values only provide what we consider to be a false sense of accuracy in these assessments". They further state (p. 739) their phylogenetic inferences should "not [be] view[ed] as reasonable estimates of vascular plant phylogeny." How phylogenetic estimates that are considered to be inaccurate and unreasonable *per se* should provide a "cladistic test" for phylogenetic relationships inferred on the basis of direct evidence, differential similarity patterns, and traditionally used methods of phylogenetic inference (branch support; Friis et al., 2007) remains unclear.

To ignore the statistical significance of clades found in some or all of the MPT and to reduce phylogeny to the production of a strict consensus tree is, at best, naïve regarding the data structure, hence, information content, of the used matrices. Relationships found in the MPT based on matrices that combine fossil and extant taxa are characterized by a high instability and are naturally dependent on included characters and character coding (see e.g. Hilton Bateman 2006 pp. 999–999). This, coupled with the insufficiency of the strict consensus (SC) approach to catch alternative topologies in a collection of trees (Fig. 1; Felsenstein, 2004, p. 522ff), allows to collapse SC trees by slightest changes in the matrix. A nice example is Crepet and Stevenson's Crepet Stevenson 2010 'analysis 2': by recoding the *defined* state in *one* character in *one* taxon to *missing*, 14 additional branches collapsed in the SC tree, among them the newly found cycad-bennettitalean clade (including several seeds ferns), a potentially "*nouveau Cycadofilicales*" group Crepet and Stevenson, 2010, p. 233, of their 'analysis 1' (SC tree based on only four MPT). Moreover, the best-supported relationships, e.g. using bootstrap support, are not necessarily represented in the strict consensus or even majority rule consensus of the MPT (Friis 2007; this study), because according clades would necessitate additional changes in other parts of the tree making such trees slightly less optimal. Some taxa and subtrees (clades) can actually be moved freely around with the result that the tree becomes only one or a few steps longer than in the 'optimal' MPT Doyle 9999 Friis 2009. Such suboptimal trees are not necessarily "wrong" or "worse" phylogenetic hypotheses. Trees that fulfil uncontested or highly supported molecular constraints are typically longer than unconstrained trees Doyle Endress 2010 Sareela 2007 Friis 2007. In other words, potentially more correct topologies that show clades, which find support from the underlying and additional data, are simply *not* optimal under parsimony using available morphological matrices, and accordingly, will *not* be represented in the SC tree of MPT.

Before one combines molecular and morphological partitions to investigate deep-level relationships using mixed models as suggested by Mathews et al. Mathews 2010, one may want to assess the quality and quantity of primary phylogenetic signal provided in the available morphological matrices. Although rarely applied (but see e.g. Wiens 2005 Nylander 2004 Denk Grimm 2009 Schlee 2011), morphological matrices can be analysed in more than one fashion and under different optimality criteria (see Felsenstein 2004, for a comprehensive overview of methods of phylogenetic inference). A clade found in trees using different optimality criteria has more credibility than a clade found only in all MPTs, which may be all biased by a data-or method-inherent branching artefacts. Signals supporting alternative, (partly) incongruent relationships can be compared at hand of commonly used support values such as nonparametric bootstrap (BS) or jackknife (JK) percentages (e.g. Zander 2004 Grimm Renner 2006 Hedenäs 2007 Denk Grimm 2009 Friis 2007). Missing data, the limited number of characters, the unknown statistical properties of character changes, and the likely primitiveness of at least some fossil taxa make it impossible to receive levels of support as found in multigene or phylogenomic studies. With the respect to potential ancestors, or taxa resembling potential ancestors, that may be included in such matrices, low (<50) but differential branch support is the only source of information (Fig. 2; ?Wiens 9999? Zander 2004 Denk Grimm 2009 MORE REFS). Such information is lost if all branches are collapsed that received less than, e.g. 80% BS support, or if the collection of MPT is summarized simply using the SC approach. Since high support values are unrealistic for these kinds of matrices, one may want to elaborate on the ambiguous signals in the data and to investigate further indicated alternative phylogenetic hypotheses. Alternative topologies and support of competing phylogenetic splits can easily be visualized using consensus networks (Holland 2003; ‘bipartition networks’, Grimm 2006; see Denk Grimm 2009 for a case study on fossil and modern beech trees) and/or extracted from bipartition tables supplied by commonly used software such as PAUP* Swofford 2002 and MrBayes Huelsenbeck Ronquist 2001 Ronquist Huelsenbeck 2003. Distance matrices can provide further useful phylogenetic information. Delta values Holland 2002 allow estimating the 'treelikeliness' of matrices prior to any phylogenetic reconstruction. Matrices with high Delta values cannot be expected to result in a single or few optimal trees that are significantly better than alternative trees. The general levels of similarity ("*overall similarity*", Mayr Bock 2002 Hörandl 2006 2007) are often stable indicators of common origins. It can be expected that taxa that share a common origin are generally more similar to each other as they would be if they were of different origin ?Felsenstein 2004, p. 999?. If a distance matrix is the exact reflection of the true tree, i.e. the phylogeny (see e.g., Denk Grimm 2009, fig. 1A), the neighbour-joining algorithm, a cluster algorithm, will recover the *true* tree, i.e. the actual phylogeny Saitou Nei 1987 Felsenstein 2004, p. 166ff.

Distance matrices that are less exact will lead to trees that differ from the true tree. The most important case in which the distance between two or more taxa does not reflect the evolutionary (phylogenetic) distance is similarity because of (many) convergently evolved traits (in this case, also the MPT and the SC tree of all MPT will consistently show erroneous relationships; see Denk Grimm 2009, fig. 1B). Another problem is missing data that may lead to ill-defined distances between many pairs of taxa. In both cases, the signal from the distance matrix (and the character matrix to some degree, see e.g. Wiens 2003) will be highly incompatible because similarity reflecting a common origin competes with similarity due to convergent evolution or because potentially discriminating characters are not preserved in fossils. Planar phylogenetic networks based on distance matrices such as the neighbour-net splits graphs Bryant Moulton 2002 2004 can compensate for the deficiencies of phylogenetic trees in the case of data incompatibility. As a boon, distance-based phylogenetic networks seem to outperform all methods of phylogenetic tree-inference when it comes to ancestor-descendant relationships Spencer 2004 Denk Grimm 2009. This may be beneficial since taxa from different time slices are combined in a single matrix. It must be expected that at least some of the taxa represent direct ancestors, ancient members, or extinct close relatives of the modern, extant, taxa.

In this study I will investigate the actual phylogenetic signal comprised in several large-scale matrices. To do so, the extent of incompatible signal in the matrices is documented and the vulnerability of the matrices against taxon and character sampling and character coding is outlined. Results of phylogenetic inferences are discussed in the light of current competing molecular hypotheses about seed plant evolution. A set of analyses is introduced to provide a guide for any future studies that want to use these matrices, or modifications thereof, in order to establish robust as possible evolutionary hypotheses and to maintain a minimum of comparability and statistical reliability between and of studies.

Note on terminology

'Monophyla' in a strict, Hennigian sense, i.e. groups with an exclusive common origin Hennig 1950, often referred to as 'clades' in cladistic-phylogenetic literature (following the English translation of Hennig's original work, Hennig Schlee 1966), are often but not always represented as clades in phylogenetic trees (cf. Fig. 2). Paraphyla, termed 'grades', i.e. groups of inclusive common origin, may likewise be resolved as clades or grades in phylogenetic trees. In order to avoid confusion, the terms 'clade' and 'grade' will be exclusively used to address according topological features of phylogenetic trees (following Felsenstein's Felsenstein 2004, suggestion), and do not imply the Hennigian categories of 'monophyletic'/ 'monophyla' and 'paraphyletic' / 'paraphyla'. Although phylogenetic inferences may not necessarily allow for straightforward distinction between inclusive and exclusive common origins, they are informative regarding the distinction groups of taxa that share a, more or less direct, common origin, i.e. ‘monophyla’ in a pre-Hennigian, “evolutionary” sense (Haeckel 1866 Hörandl 2006; see e.g. Schwarz 1936 1936-1937 classification of oaks), from artificial groups of taxa that are not likely to be of common origin, i.e. polyphyletic in general-phylogenetic (pre-Hennigian and Hennigian) sense. ‘Common origin’ will hence be used exclusively in a general-phylogenetic sense.

Figure 1. Insufficiency of strict consensus trees when summarizing alternative topologies (e.g. MPT). **A**. Four trees, which only differ by the position of taxon T8. The four red (T1–T4) and three green taxa (T5–T7; OT, outtaxon) are consistently placed in different subtrees; the branches that refer to the common origins of both major groups are coloured accordingly. If T8 would be deleted from the trees, all four trees would be absolutely identical showing the same phylogenetic relationships (and clades). **B.** Strict consensus tree of all four trees. Because of the uncertain position of T8, most branches are collapsed, and the information about the uncontested common origin of the two groups, red and green, is lost. **C**. Rooted strict consensus network simultaneously showing all relationships indicated in the four trees. In contrast to the consensus tree (**B),** it is clear from the network that the green and red groups are phylogenetically distinct (no edges are shared by green and red taxa), and that all encountered topological ambiguities relate to T8. Parallel edge bundles refer to the same branch in individual trees, e.g. the red and green edge bundles refer to the red and green branches highlighted in all four trees (**A**).

Figure 2. Information provided by differential branch support. Some branches/edge bundles and associated supports are coloured for better traceability: correct (green) and wrong (red) alternative placements of T8; bluish: equally optimal placements of T7 in a phylogenetic

tree. **A.** A hypothetical evolutionary pathway (phylogeny), T7, T8 and the outgroup (OT) are fossil taxa, T7 being the direct ancestor of T5 and T6, T8 an ancient extinct sister lineage of T5–T7 with ambiguous signal because of convergent evolution, OT an ancient primitive taxon with only 20 uniquely derived characters out of 100. Both major lineages share 50 derived characters, 10 additionally derived characters in lineage T1-T2 and T3-T4, respectively, are evolved convergently to the T7 lineage. Of the 40 derived characters of T8, 10 represent the common origin with the T7 lineage, 10 are unique, and each 10 are either evolved in parallel in T5 or T6, the offspring of T7 (matrix included in the Electronic Supplement at www.palaeogrimm.org/data). **B.** Traditionally used majority-rule consensus tree based on 10,000 BS replicates. BS support annotated along branches. T6 is placed as sister to a clade comprising its predecessor T7 and its sibling T5. **C.** Alternative, slightly less supported topology, with BS support annotated based on the same 10,000 replicates. **D.** Bipartition network based on the same 10,000 BS replicates, showing all information at once that can be obtained from BS supports. Edge lengths are proportional to the number (or percentage) of replicates showing an according bipartition. Note that, pending on the amplitude (number of characters) and quality (proportion of convergent characters) of the underlying signal, differential support may either indicate ancestor-descendant relationships (T7 vs. T5, T6) or the lack of proper signal (as in the case of T8).

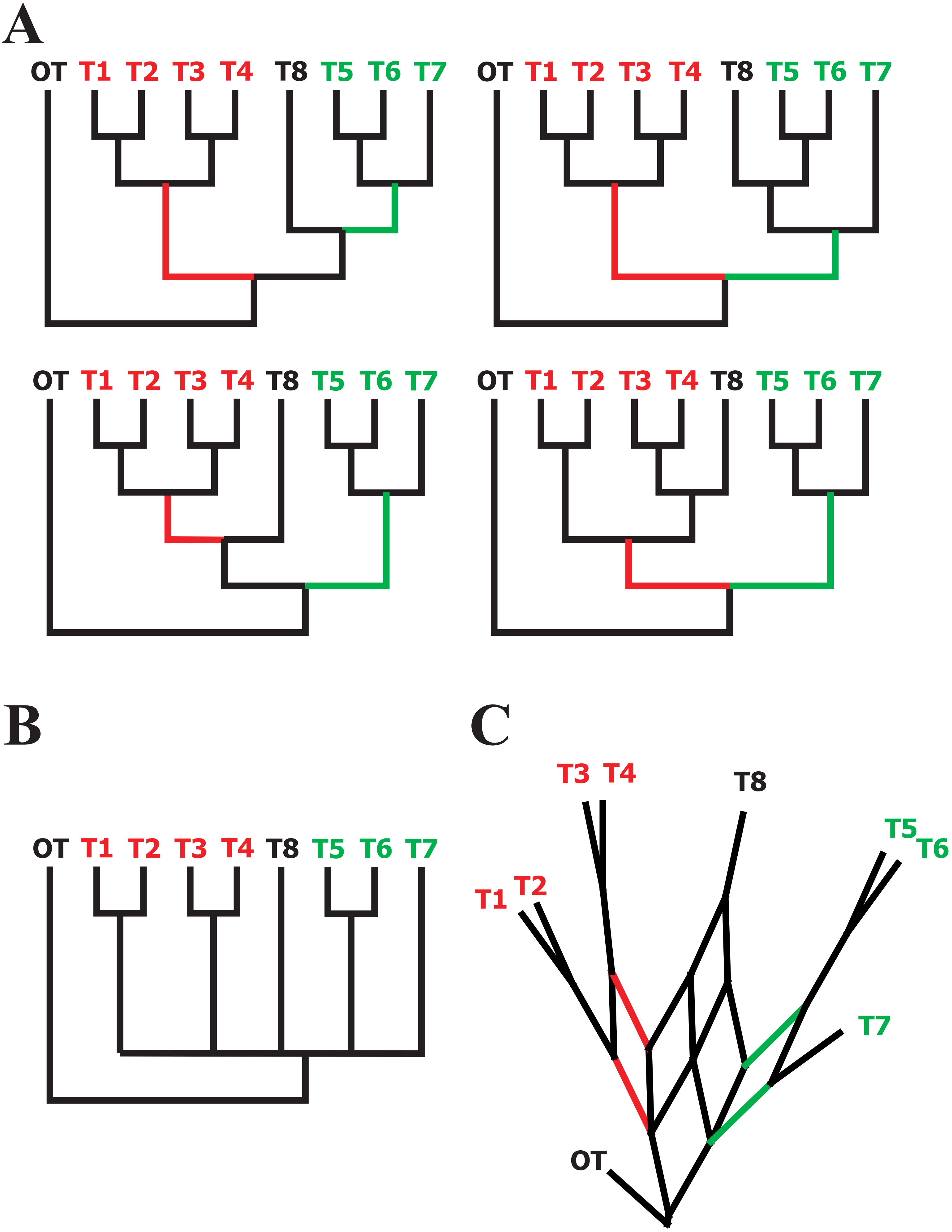

A
OT T1 T2 T3 T4 T8 T5 T6 T7
OT T1 T2 T3 T4 T8 T5 T6 T7
OT T1 T2 T3 T4 T8 T5 T6 T7
OT T1 T2 T3 T4 T8 T5 T6 T7
B
OT T1 T2 T3 T4 T8 T5 T6 T7
C
T3 T4
T8
T5
T6
T1 T2
T7
OT

Figure 1

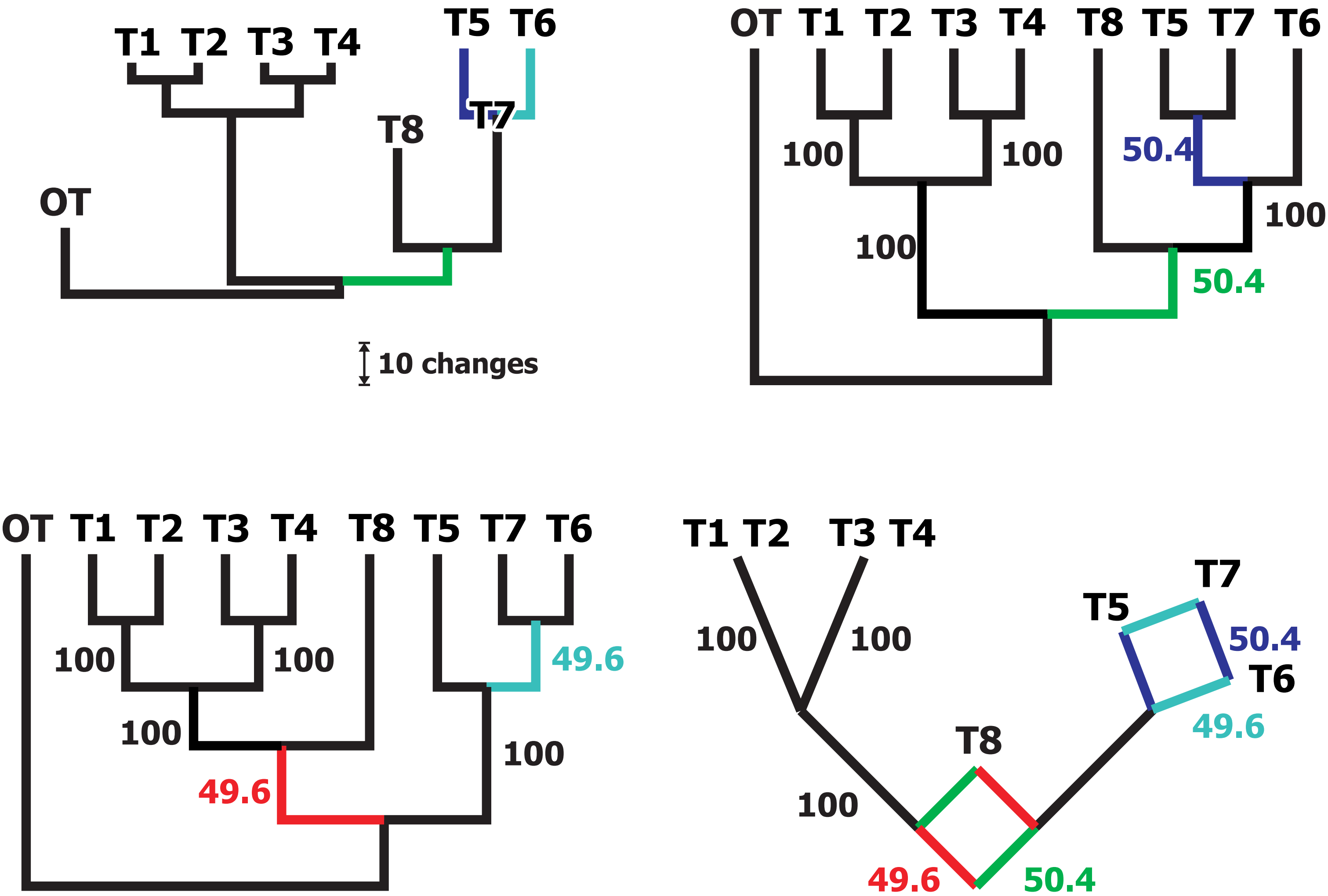

T1 T2 T3 T4
T5 T6
T7
T8
OT
10 changes
OT T1 T2 T3 T4 T8 T5 T7 T6
100
100
50.4
100
100
50.4
OT T1 T2 T3 T4 T8 T5 T7 T6
100
100
49.6
100
100
49.6
T1 T2 T3 T4
100
100
T7
T5
50.4
T6
49.6
T8
100
49.6
50.4

Figure 2

Material & Methods

Matrices

For my re-evaluation, I compared three matrices that have been used to infer, criticise and/or discuss hypotheses about spermatophyte phylogenetic relationships Friis 2007 Rothwell 2009 Crepet Stevenson 2010. The first matrix is essentially the same set of taxa and data than used by Hilton and Bateman Hilton 2006, for additions and recoding of one character see Friis et al. Friis 2007. The usefulness of this matrix was criticized by Rothwell et al. Rothwell 2009, who compared both matrices Hilton Bateman 2006 Friis 2007 to a third matrix with a set of taxa (and characters) that should be a "compromise" between sets used earlier Rothwell 2009, p. 309. In their compromise-matrix ('matrix 3', and based thereon, 'matrix 4', including two taxa joined by Friis et al., and 'matrix 5', including the joined taxon, Erdtmanithecales of Friis et al.), Rothwell and colleagues opt to drop, fuse, and/or recode several taxa and characters of the combined matrices. In particular, *Caytonia* was recoded to "remove embedded hypotheses in character codings" (p. 309) with the result that *Caytonia* became less similar to the angiosperms as it had been in the matrix of Hilton and Bateman Hilton 2006; Friis 2007, but also phylogenetically more ambiguous (see Results). In contrast to the matrices of Hilton & Bateman Hilton 2006 and Friis et al. Friis 2007, composed mostly of characters with two or three (rarely four) states, unordered characters are used with up to eight states (e.g., character 35 in Rothwell 2009 corresponds to three characters in Hilton Bateman 2006 Friis 2007). Thus, modified character sets represent essentially a re-weighting of original characters. The matrix of Crepet and Stevenson Crepet 2010 is a modification of the Rothwell et al. Rothwell 2009 'matrix 3'. The most important change according to the authors was replacing the extinct order Bennettitales by four member taxa and recoding characters in a “conservative” manner, making the Bennettitales less similar to angiosperms but more similar to Cycadales, but also primitive groups of seed plants (this study). Crepet and Stevenson also omitted Friis et al.’s charcoalified seed from Rothwell et al.’s ‘matrix 3’.

Phylogenetic reconstructions

Two general kinds of matrices were used to estimate phylogenetic relationships: the original character matrices and matrices of simple (Hamming) pairwise distances (as used in Friis 2007 Friis 2009 Denk Grimm 2009). All matrices and results of phylogenetic inferences can be downloaded at www.palaeogrimm.org/data.

Treelikeliness was estimated using Delta values Huber 2002. Delta values computed (with DIST_STATS; Auch 2006 for individual taxa allow predicting the amount of topological incongruence induced by these taxa, whereas the Delta value of a matrix reflects its treelikeliness (e.g. Auch 2006 Göker Grimm 2008). Tree inferences based on matrices with high Delta values are not likely to converge to a single optimal, well-supported solution, whereas matrices with low Delta values will result in a tree that is significantly better than all other alternatives (regarding the number of extra changes under parsimony if changed, its likelihood, and/or its support)

Phylogenetic trees were inferred based on the character matrices using maximum parsimony (MP) and maximum likelihood (ML), and based on distance matrices using the least-square fit (LS), as optimality criteria. MP inferences used the heuristic search algorithm of PAUP Swofford 2002 and default settings. In case analyses resulted in >10,000 MPT, “AddSeq” was set to “Random” using 100 addition replicates, and maximal 1,000 MPT were set per addition replicate (Option “NChuck = 1000”; “ChuckScore” set depending on matrix). The collection of MPT was summarized using the consensus network approach Holland 2003 as implemented in SplitsTree 4 (Huson Bryant 2006; threshold set to null, branch length option set to "None"). ML inference and bootstrapping relied on RAxML-HPC Stamatakis 2006 Stamatakis 2008 using one of the two implemented models for non-molecular, multistate data (MK; Lewis 2001) and allowing for site variation (+Γ). At the moment, polymorphic states cannot be handled in the case of standard categorical data under ML and were recoded as missing data. In addition to the inferred best-known tree, the collection of MPT and the LS tree were optimized under ML (option -f e) and subdued to a topological test (SH-test by Shimodaira Hasegewa 1999) in order to see if certain topologies can be rejected or not by this means. In addition to character-based analyses, the neighbour-joining (NJ) algorithm in its BioNJ implementation Gascuel 1997 was used to compute a tree based on a matrix of simple (Hamming) pairwise distances (with PAUP*) that fulfils the LS criterion (see Felsenstein 2004, p. 160ff). The Neighbour-net algorithm Bryant Moulton 2002 Bryant Moulton 2004 implemented in SplitsTree 4 was used to compute a planar phylogenetic network based on the same distance matrix. The NN algorithm basically adds a second dimension to the NJ algorithm, and was found to overcome a number of limitations inherent to tree-based phylogenetic reconstructions (see e.g., Lockhart 9999 Moulton 9999 Grimm Denk 9999).

Branch support was established using nonparametric bootstrapping (BS; Felsenstein 1985) and jackknifing (JK; REF) under MP and ML and, in addition, using Bayesian inference (BI), the latter providing posterior probabilities Rannala 1994. MP- and NJ-BS and -JK branch (edge) support was established using PAUP* with 10,000 (pseudo)replicates, and only one tree saved per replicate (Müller 2005); under ML the bootstop criterion (Pattengale 2009) implemented in RAxML (Stamatakis 2008) was used to determine the number of necessary replicates (option -# autoMRE), the maximum number were 1,000 replicates. A phyton script (programmed by xxx) was used to perform ML character-jackiknifing with RAxML. Since all matrices comprise numerous columns with undefined characters states for a significant number of taxa, it could be expected that BS supports are generally low. Therefore, JK supports were established with different deletion percentages: either 50 (^{50}JK; the default setting), 20 (^{20}JK), or only 5% (^{5}JK) of the characters were deleted per replicate. The high number of replicates (10,000) ensures to cover a huge amount of possible character combinations (in particular ^{5}JK), hence, allow testing the character-sampling dependence of inferred relationships. Posterior probabilities (PP) for bipartitions (which are resolved as branches in trees) rely on MrBayes 3.1 Huelsenbeck Ronquist 2003. In contrast to molecular analyses, PP based on morphological matrices are often as low as BS supports, because, in addition to incompatible signals, substitution patterns are not as decisive in a probabilistic environment as under MP, where each character change is treated as equally important. Competing phylogenetic splits (bipartitions) were visualized using modules implemented in SplitsTree 4 Huson Bryant 2008. The consensus network module can be used to compute a splits graph (network) that only shows bipartitions found in a certain percentage of the read-in replicate trees (BS or JK replicate trees, saved topologies under BI). Alternatively, bipartition tables produced by PAUP* and MrBayes can be directly used for coding a NEXUS-formatted SPLITS block for SplitsTree 4 (contact the first author for practical support).

Results & Discussion

Phylogenetic inferences

For the reasons outlined in the Introduction, phylogenetic relationships, ambiguity in phylogenetic estimations, and differential support (incompatible signal) are investigated using splits graphs based on pairwise distances, collections of trees, and split frequencies in collections of trees. The Neighbour-net (NN) splits graph in Figure 3, a planar phylogenetic network, comprehensively visualizes the available phylogenetic signal in the matrix of Friis et al. Friis 2007. Members of the major groups of seed plants are distinctly grouped (clustered) in the graph. The more tree-like portions of the graph correspond to well-known groups that consistently form clades in (rooted) phylogenetic trees (e.g., angiosperms, Cycadales) that, in addition, receive relatively high support independent of the method ($BS_{ML/NJ/P} \geq 60$; $PP \geq 0.8$; Fig. 3). Clades comprising only the modern Gnetales and conifers, respectively, do not receive the same level of support because of the inclusion of fossil taxa more or less related to these but also further groups (Friis et al.'s "charcoalified seeds", Erdtmanithecales; *Emporia, Thucydia,* Cheirolepidaceae, see Table 1; box-like structures in Fig. 3). A clear distinction is found between the most primitive seed plants ("hydraspermalean" and "medullosan seed ferns", Hilton Bateman 2006; orders Lyginopteridales, Calamopityales, and Medullosales; Taylor 2009) and outtaxa (progymnosperms) and the more derived fossils, which are placed closer to the extant taxa (tentatively termed 'higher seed plants' in the following); the according bipartition receives significant support ($BS_{ML/NJ/P} \geq 70$; $PP = 1.00$). Within the higher seed plants, members of the putative 'anthophyte' and the Bennettitales-Erdtmanithecales-Gnetales (BEG) clades (cf. Friis 2007) are recognized as distinctly different from the remaining gymnosperms (cycads, *Ginkgo*, conifers). The Palaeozoic Cordaitales (cf. Taylor 2009) are recognized as closest relatives of the conifers, with *Ginkgo* being a possible sister lineage of a Cordaitales-conifer-lineage; a clade of the three taxa (GCC clade) would receive moderate support from NJ bootstrapping ($BS_{NJ} = 57$) and Bayesian inference ($PP = 0.72$), but negligible support under MP and ML ($BS_{P/ML} < 0.2$; better supported alternatives are, however, not found, Table 2). The various groups of seed ferns are placed according to their age and, linked with that, their derivedness. The only clearly placed seed fern is *Caytonia,* a "higher seed fern" Hilton Bateman 2006; Caytoniales, Taylor 2009, which appears to be the closest (but not a close) relative of the angiosperms that has been included in the matrix, i.e. representing a potential sister lineage (Figs 3, 4). An alternative sister group of the angiosperms would be the BEG clade, or only the Gnetales (including the Friis et al.'s charcoalified seed), but the signal for this is less pronounced (Fig. 3) and receives significantly less support than the alternative of a *Caytonia*-angiosperm clade (Fig. 4).

Box-like portions, relating to incompatible signals in the matrix, dominate in particular the central and some terminal parts of the NN splits graph. The matrix provides only weak signal to sort relationships between relatively close-related taxa, e.g. between the extant members of the angiosperms or the ancestral (Palaeozoic) group of 'hydraspermalean' seed ferns, which are only weakly differentiated in their defined characters (Table 2). Central parts of the graph are affected by the naturally ambiguous signal of potential ancestral members of well-defined clades (cf. Fig. 2; extinct conifers, Erdtmanithecales, and charcoalified seeds; see Fig. 4 for differential support) and fossil taxa lacking decisive, putatively apomorphic characters or showing combinations of putatively apomorphic characters typical of several modern groups ('mosaic' taxa; Table 3).

In contrast to trees, the network is able to further sort the ambiguous phylogenetic signal from the matrix. The phylogenetic position of 'higher' (except *Caytonia*) and "peltaspermalean" Hilton Bateman 2006 seed ferns, *Callistophyon*, and the Cycadales (but see note on *Gingko* below) in comparison to other derived seed plants relates only to weak and accordingly indecisive signal in the matrix: only short edge bundles connect these taxa to the inflated, spider web-like centre of the graph, which represents the root of all higher seed plants (support from phylogenetic tree-inference of any possible placement is accordingly low; Fig. 4). In conclusion, the phylogenetic position of these taxa within the 'higher' seed plants is essentially unresolved. In contrast, in the case of the GCC group, the graph structure (Fig. 3) is highly informative regarding competing phylogenetic signals (cf. Fig. 4), and translates directly into a phylogenetic sequence. The oldest and assumedly most primitive taxa of this lineage (Cordaitales) share edge bundles with the group of primitive seed plants and outtaxa, but the according bipartition is not found in the collections of BS replicate trees and BI saved topologies. The oldest potential conifers in the matrix, *Emporia* and *Thucydia* (Voltziales, Upper Pennsylvanian; Taylor 2009) share equally prominent edges both with the Cordaitales (possible ancestors of conifers), and the Pinales (Coniferales, putative descendants or sister lineages), which, in the matrix, are represented by the Mesozoic Cheirolepidaceae (Upper Triassic to Upper Cretaceous; Taylor 2009, p. 836f) and the extant conifer families (also known from the Mesozoic onwards; Taylor 2009). Thus, they provide a phylogenetic link between the two groups. The Cheirolepidaceae, considered a diverse extinct group of conifers with affinities to several of the extant groups, are placed accordingly. This phylogenetic scenario can be backed by the differential support patterns in this group (Table 2; cf. taxon T7 in Fig. 2). An analogue situation is found for the members of the BEG clade. Differentiation patterns appear more tree-like here (Fig. 3), but a missing data artefact must be taken into consideration ($BS_{ML/P}$ generally low; Figs 3, 4). *Ginkgo* (Gingkoales) on the other hand is phylogenetically closer to the Cycadales (and several extinct seed ferns: *Callistophyton*, *Autunia*, *Peltaspermum*) than the remainder of the GCC group. Excluding the unlikely possibility that *Ginkgo* was originally a conifer-cycad hybrid, there are only three scenarios left to explain this setting: (1) the signal that associates *Ginkgo* with conifers is due to convergent evolution, (2) the signal that associates *Ginkgo* with Cycadales is due to convergent evolution, (3) the association of *Ginkgo* and Cycadales relates to shared or retained primitive characters, (sym-)plesiomorphies, that were subsequently replaced (or lost) in the Cordaitales-conifer lineage (Fig. 5). The observed differential support patterns (Fig. 4; ES) clearly favour scenario three: an according branch (*Gingko*-Cycadales clade) would receive virtually no support (BS/PP ≤ 8/0.05; cf. taxon T8 in Fig. 2).

The recoding effort of Rothwell et al. Rothwell 2009 has relatively little effect on the inferred relationships. The NN splits graph based on Rothwell et al.'s Rothwell 2009 matrix 5 (Fig. 6A) recovers the same groups and relationships than shown in Figures 3 and 4. Members of the putative anthophyte, BEG and GCC clades are still clustered, as well as the members of the primitive groups of seed ferns. 'Anthophytes' are still distinct from the remaining gymnosperms (Cycadales, GCC clade), with several groups of seed ferns placed in between. The inclusion or exclusion of the Erdtmanithecales has little effect (compare Fig. 6A vs. 6B). Only the phylogenetic affinities of the 'higher' seed ferns Hilton Bateman 2006 change (Fig. 6), which is mainly due to the recoding of *Caytonia*. If the latter is excluded from matrix 3 (Fig. 6B shows the graph including *Caytonia*), *Glossopteris* swaps back to the same position than shown in Figure 6A and, in addition, the corystosperms then cluster again with the other two "peltaspermaelan" Hilton Bateman 2006 seed ferns (cf. Figs 3, 4). Excluding *Caytonia* from the Friis et al. (2009) matrix has, however, no effect on the remaining relationships (not shown; see ES). Deletion of Friis et al.'s charcoalified seed from Rothwell et al.'s matrix 3, further weakens the Bennettitales-Gnetales link, but a small edge bundle still links the Gnetales-genus *Welwitschia* to Bennettitales, so that the phylogenetic position of the latter is unaltered (according graphs supplied in the ES). The major effect of the rescoring by Rothwell et al. Rothwell 2009 is a general decrease of branch support for many relationships under MP (BS_P), less pronounced in BS_{NJ}, and an increase of support of several terminal relationships (compare Fig. 6A, B to Figs 3, 4); the latter is possibly due to the more restricted taxon sampling. This is most pronounced in the case of the Bennetittales, any possible placement of this taxon receives very low support from all methods ($BS_{NJ/P/ML} < 20$; PP < 999). Its position within the higher seed plants is basically unresolved. The best, but low supported alternative based on Rothwell et al.'s matrix 5 would be two place the Bennettitales as sister to the Erdtmanithecales ($BS_P = 21$; $BS_{NJ} < 5$) and based on their matrix 3 within a clade comprising all ‘anthophytes’ as circumscribed above (Figs 3, 4) except *Caytonia* ($BS_{NJ/P}$ = 24/30; also not included: *Pentoxylon,* cf. Rothwell et al., 2009, fig. 30). Thus, unlike stated by the original authors, the recoded matrices of Rothwell et al. Rothwell 2009 do not provide any phylogenetic signal that is at odds with the phylogenetic relationships inferred from the Friis et al. Friis 2007 matrix. However, many relationships indicated by their well-resolved SC tree (Rothwell et al., 2009, fig. 30) are at odds with the best supported relationships based on the very same matrix (highlighted by red colour in Fig. 6B). The only notable difference between the matrices used in the two studies is the amplitude of the signal for a given phylogenetic relationship and its potential alternatives (additional results for Rothwell et al.’s matrices 3–5 can be found in the ES).

The further matrix modifications applied by Crepet & Stevenson Crepet 2010, which used Rothwell’s ‘matrix 3’ as a basis but a more “conservative” (p. 999) character coding, the NN algorithm results in a somewhat distorted graph (Fig. 7), but the main relationships defined based on the Hilton and Bateman/Friis et al. matrix Hilton 2006/Friis 2007are still confirmed. Crepet & Stevenson Crepet 2010 excluded Friis et al.’s Friis 2007 charcoalified seeds and Erdtmanithecales and divided the Bennetittales into four different taxa. Based on their matrix, a Bennettitales clade receives high support; the according taxa are also distinctly grouped in the NN splits graph (Fig. 7). The major changes are that the Bennettitales now appear as a relatively primitive group, with little affinity to any of the major extant groups, except maybe for the Gnetales(sic!)-genera *Gnetum* and *Welwitschia* (confirming Friis 2007). *Caytonia*, on the other hand, shares (cf. Figs 3, 4) more pronounced edges with the angiosperms or part of this group, but with little support Fig. 7. The best supported alternative would however to place it as to the corystosperms (as in Crepet Stevenson 2010’s, SC tree based on ‘analysis 1’; oddly placed in the graph in Fig. 7) or as sister to *Glossopteris* (in agreement to the graph;

Fig. 7). The position of the remaining taxa agrees with the results outlined above, in particular regarding the relationships between the members of the putative GCC clade (see Figs 3, 4).

Figure 3. Neighbour-net splits graph illustrating the phylogenetic unfolding of seed plants (matrix of Friis 2007). Numbers at brackets refer to BS support (using MP, NJ, ML) and posterior probabilities (PP; using BI) of corresponding branches in phylogenetic trees ('clade' support).

Figure 4. Bipartition networks visualizing differential support of competing phylogenetic splits (box-like parts of the graph), and subsequently candidate clades (matrix of Friis 2007). Only bipartitions are shown that occurred in more than 15% of the bootstrap replicates. **A.** Splits graph based on the bipartition table of a NJ bootstrap analysis with 10,000 replicates. **B.** Splits graph based on the bipartition table of a parsimony bootstrap analysis with 10,000 replicates. Some competing split patterns and associated support values are highlighted by colours; e.g. a *Caytonia*-angiosperm clade (orange edges) would receive $BS_{NJ/P} > 50$; the alternative of a BEG-angiosperm clade only $BS_{NJ/P} < 20$ (light turquoise edges). Equally coloured edges in **A** and **B** refer to the same bipartition (phylogenetic split).

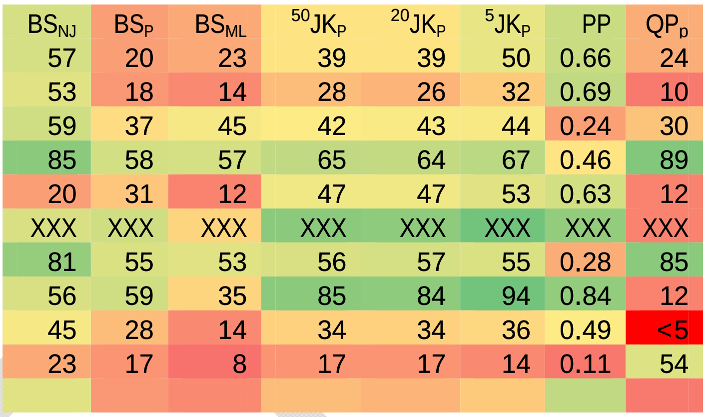
TODO Figure 5. Possible evolutionary scenarios that could explain the position of *Ginkgo* in the NN splits graph (Fig. 3), and the low support of alternative phylogenetic splits (bipartitions; Fig. 4). **A**. Scenario 1: *Ginkgo* sister to conifers; characters shared by *Ginkgo* and Cycadales are either convergences or plesiomorph. **B.** Scenario 2: *Ginkgo* sister to Cycadales; characters shared by *Ginkgo* and conifers (and Cordaitales) are either convergences or plesiomorph. **C**. Scenario 3: Cycadales and *Ginkgo* forming a grade to a Cordaitales-conifer clade. Characters shared by Cycadales and *Ginkgo* are symplesiomorphies, subsequently replaced by apomorphic characters in the Cordaitales-conifer lineage.

Figure 6. Neighbour-net splits graphs based on the “compromise” matrix of Rothwell et al. (Rothwell 2009). The de-coding of charcoalified seeds enforces their role as a phylogenetic link between Gnetales and Bennettitales. **A.** Graph based on morphological distances based on Rothwell et al.’s matrix 5 (including the criticized Erdtmanithecales). The ‘higher seed ferns’ including the recoded *Caytonia* are placed more distant to the angiosperms, but the BEG group of Friis et al. (Friis 2007) is still found despite the tasked weakening of the signal. **B.** Graph based on Rothwell’s et al.’s matrix 3 (excluding the Erdtmanithecales). The *Gingko*-Cordaitales-conifer (GCC) group/clade swaps position with *Caytonia, Glossopteris* and Corystosperms; the ‘higher seed ferns are dissolved (arrows). The Bennettitales are still placed between the Gnetales and *Pentoxylon*. Green coloured groups and splits (brackets and cross-lines) correspond with clades resolved in the original strict consensus (SC) trees (Rothwell 2009, figs 999; blue colour refers to herein inferred best-supported relationships (potential clades) not resolved in the according original SC trees; red colouration indicates best-supported relationships incongruent with the original SC trees. Numbers at brackets and cross-lines indicate branch support based on non-parametric bootstrapping (BS_{NJ}/BS_{P}) under different optimality criteria (see M&M).

Figure 7. Phylogenetic network based on the matrix of Crepet & Stevenson (Crepet 2010). Green coloured groups and splits (brackets and cross-lines) refer to clades resolved in the original strict consensus (SC) tree (analysis 1, Crepet & Stevenson, 2010, fig. 10.6); blue colour refers to herein inferred best-supported relationships (potential clades) not resolved in the according original SC tree; red colouration indicates best-supported relationships incongruent with the original SC trees. Cyan (replacing green) and purple (replacing red) denote branches collapsed in the SC tree based on Crepet & Stevenson’s analysis 2 (Crepet

2010, fig. 10.7). Numbers at brackets and cross-lines indicate branch support based on non-parametric bootstrapping (BS_{NJ}/BS_{P}) under different optimality criteria (see M&M).

NFSHD Table 1. Differential support of alternative systematic affinities of and between fossil and extant members of the conifers (Pinales; Earle 2010) based on the matrix used by Friis et al. Friis 2007

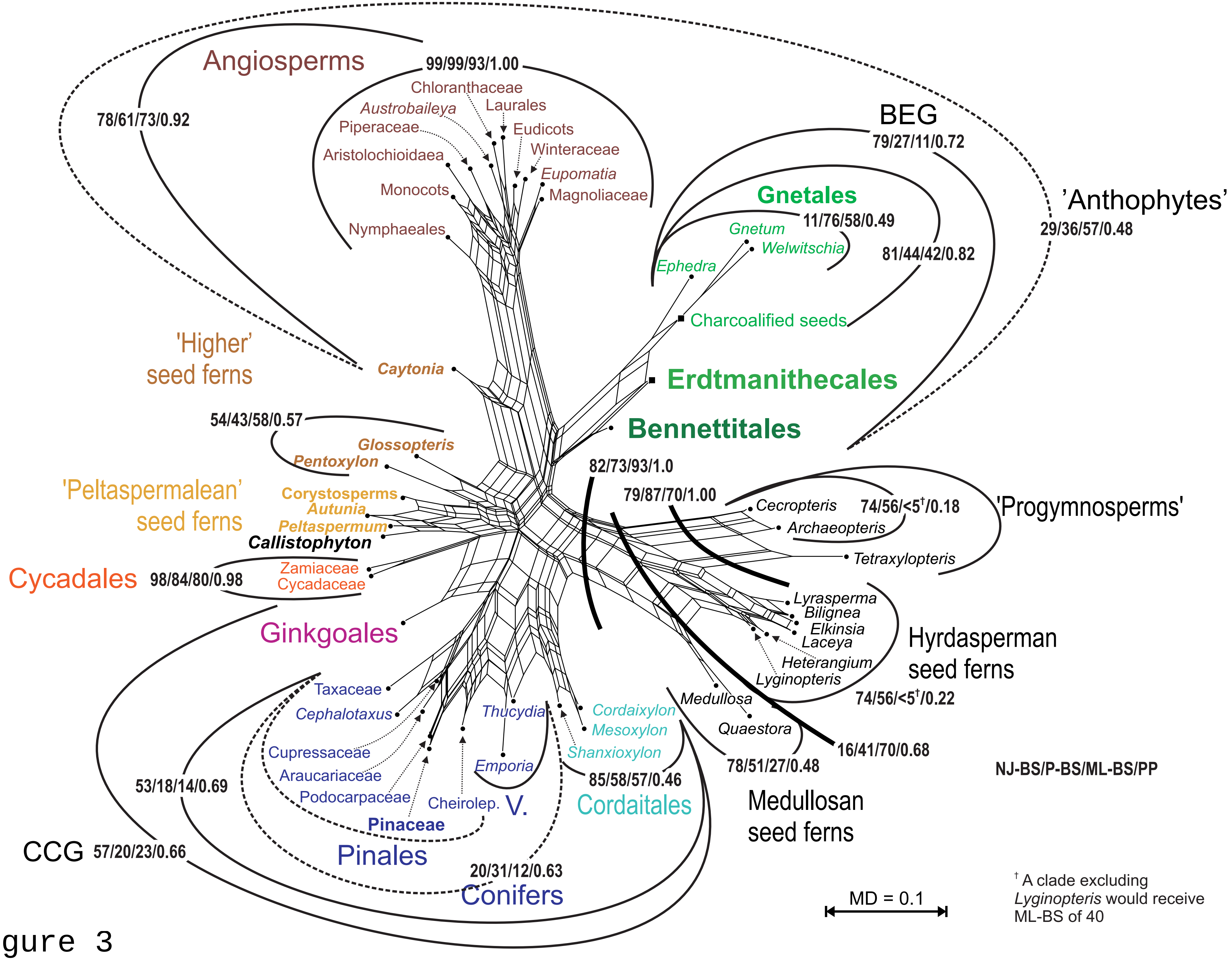


Putative clades	BS_{NJ}	BS_P	BS_{ML}	$^{50}JK_P$	$^{20}JK_P$	$^{5}JK_P$	PP	QP_p
[0] *Ginkgo*-Cordaitales-conifers (GCC clade)	57	20	23	39	39	50	0.66	24
[1] Cordaitales-conifers (extant + extinct)	53	18	14	28	26	32	0.69	10
[2A] Cordaitales-Voltziales	59	37	45	42	43	44	0.24	30
[3] Cordaitales	85	58	57	65	64	67	0.46	89
[2B] Extant and extinct conifers	20	31	12	47	47	53	0.63	12
[4A] Voltziales-Cheirolepidaceae	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX
[5] *Emporia-Thucydia* (Voltziales)	81	55	53	56	57	55	0.28	85
[4B] Extant and extinct Pinales*	56	59	35	85	84	94	0.84	12
[6A] Extant Pinales	45	28	14	34	34	36	0.49	<5
[6B] Cheirol.-Pinaceae-Podocarpaceae	23	17	8	17	17	14	0.11	54

* *Cephalotaxus,* Taxaceae, Cupressaceae, Araucariaceae, Podocarpaceae, Pinaceae, †Cheirolepidaceae

Table 2. Number of different (above diagonal) and identical (below diagonal) characters among members of the angiosperms. Note the difference in the number of coded characters (NDC). The number in the diagonal shows the number of defined characters for the respective taxon, the last column indicates the least-distinct (LDR) and most-similar relative (MSR) of the taxon within the angiosperm clade. NB: The placement a taxon is instable in trees if the least-distinct and most-similar relatives are different taxa. The same applies to more than once listed LDR and MSR (e.g. Monocots or Eudicots)

Taxon in matrix	NDC												LDR	MSR
Magnoliaceae	95%	**97**	4	9	15	11	13	14	16	17	18	13	*Eupomatia*	Same
Eupomatia	92%	90	**94**	10	15	7	13	13	16	17	21	14	Magnoliaceae	Same
Austrobaileya	95%	88	84	**97**	8	10	11	10	18	16	19	16	Chloranthaceae	Magnoliaceae
Chloranthaceae	87%	74	74	81	**89**	8	12	10	14	7	19	13	*Eupomatia,* Piperaceae	Piperaceae
Laurales	88%	79	83	80	81	**90**	12	11	17	17	19	17	*Austrobaileya,* Eudicots	*Eupomatia*
Winteraceae	95%	84	81	86	77	78	**97**	6	16	17	13	14	*Eudicots*	Same
Eudicots	94%	82	81	86	79	79	90	**96**	8	8	9	6	Winteraceae, Monocots	Same
Aristolochioidaea	93%	79	78	77	75	73	79	87	**95**	10	12	5	Monocots	Same
Piperaceae	92%	77	77	78	82	73	77	86	84	**94**	13	9	Chloranthaceae	Eudicots
Nymphaeales	91%	75	72	74	70	71	80	84	81	80	**93**	8	Monocots	Same
Monocots	95%	84	80	81	76	73	83	90	90	85	85	**97**	Aristolochioidaea	Aristolochioidae, Eudicots



Angiosperms
99/99/93/1.00
78/61/73/0.92
Chloranthaceae
Austrobaileya
Laurales
Piperaceae
Eudicots
Winteraceae
Aristolochioidaea
Eupomatia
Monocots
Magnoliaceae
Nymphaeales
BEG
79/27/11/0.72
'Anthophytes'
29/36/57/0.48
Gnetales
Gnetum
11/76/58/0.49
Welwitschia
81/44/42/0.82
Ephedra
Charcoalified seeds
'Higher' seed ferns
Caytonia
Erdtmanithecales
Bennettitales
54/43/58/0.57
Glossopteris
Pentoxylon
82/73/93/1.0
79/87/70/1.00
'Peltaspermalean' seed ferns
Corystosperms
Autunia
Peltaspermum
Callistophyton
Cecropteris
74/56/<5†/0.18
Archaeopteris
'Progymnosperms'
Tetraxylopteris
Cycadales
98/84/80/0.98
Zamiaceae
Cycadaceae
Lyrasperma
Bilignea
Elkinsia
Laceya
Ginkgoales
Hyrdasperman seed ferns
Heterangium
Lyginopteris
74/56/<5†/0.22
Taxaceae
Cephalotaxus
Thucydia
Cordaixylon
Medullosa
Mesoxylon
Quaestora
16/41/70/0.68
Cupressaceae
Emporia
Shanxioxylon
Araucariaceae
78/51/27/0.48
NJ-BS/P-BS/ML-BS/PP
53/18/14/0.69
85/58/57/0.46
Podocarpaceae
Cheirolep.
V.
Cordaitales
Medullosan seed ferns
Pinaceae
CCG
57/20/23/0.66
Pinales
20/31/12/0.63
Conifers
† A clade excluding Lyginopteris would receive ML-BS of 40
MD = 0.1

Figure 3

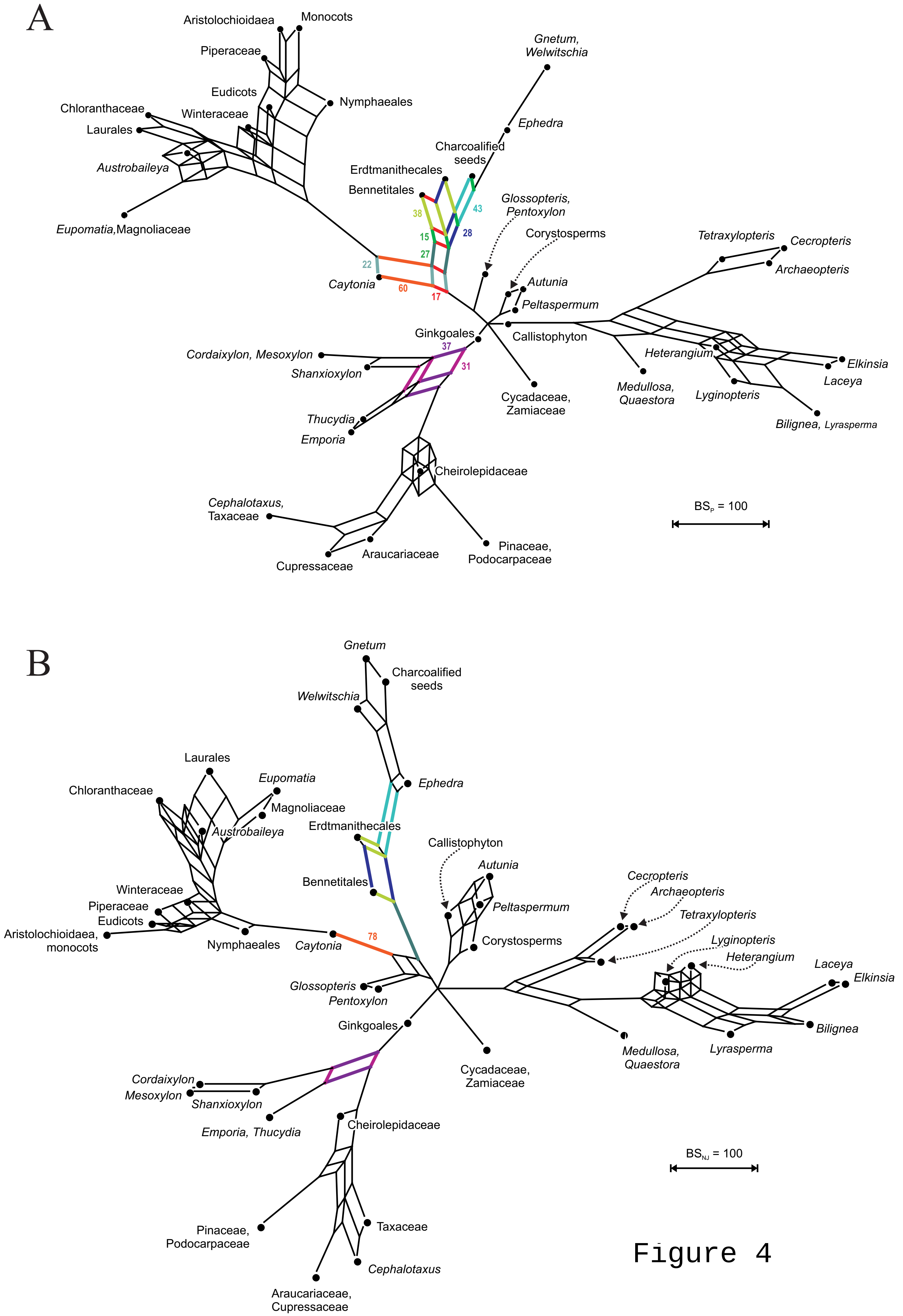
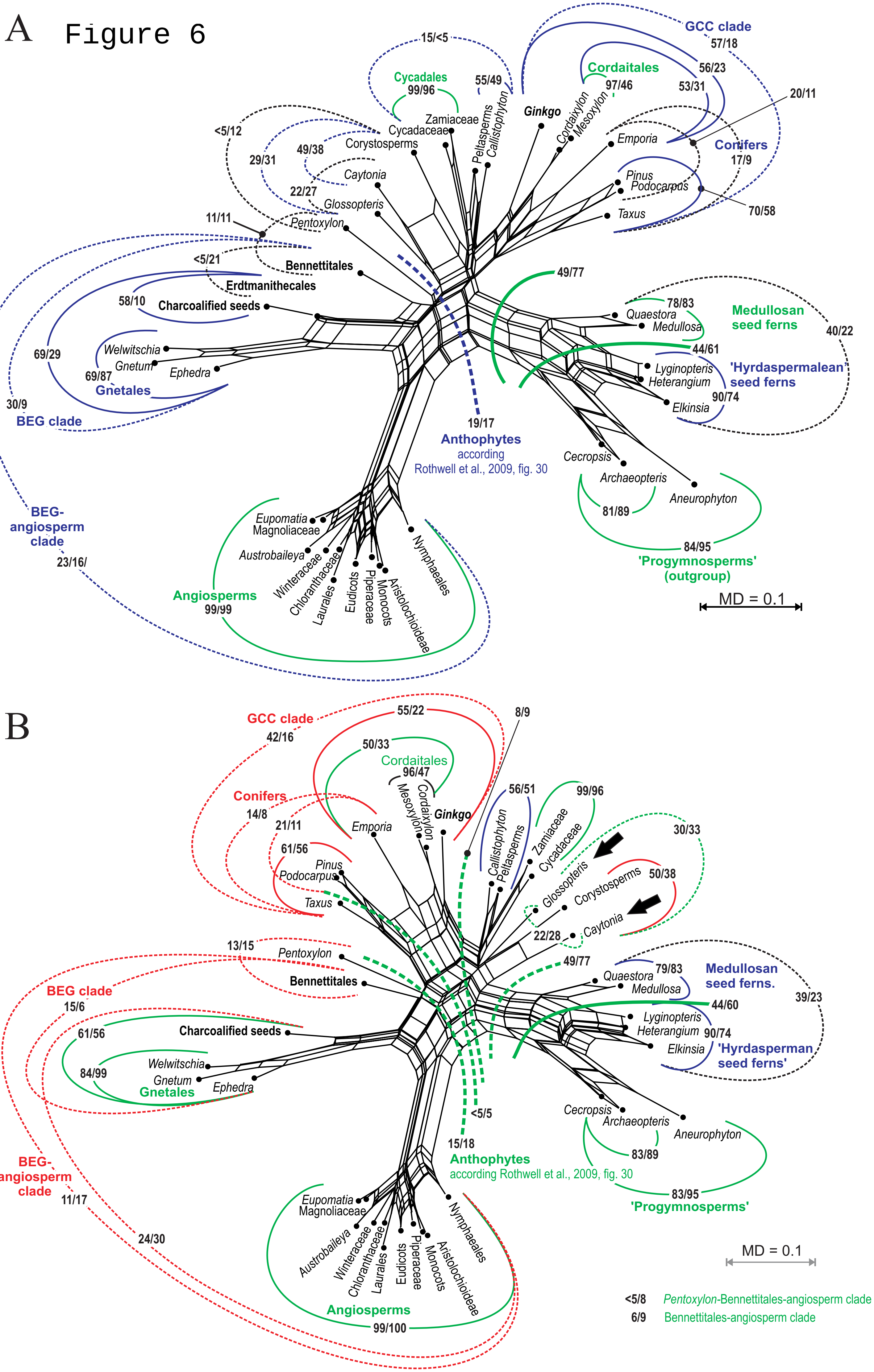

A
Aristolochioidaea
Monocots
Piperaceae
Eudicots
Nymphaeales
Chloranthaceae
Winteraceae
Laurales
Austrobaileya
Eupomatia,Magnoliaceae
Gnetum,
Welwitschia
Ephedra
Charcoalified
seeds
Erdtmanithecales
Bennetitales
38
15
27
43
28
22
Caytonia
60
17
Glossopteris,
Pentoxylon
Corystosperms
Autunia
Peltaspermum
Callistophyton
Tetraxylopteris
Cecropteris
Archaeopteris
Heterangium
Elkinsia
Laceya
Medullosa,
Quaestora
Lyginopteris
Bilignea, Lyrasperma
Ginkgoales
37
31
Cordaixylon, Mesoxylon
Shanxioxylon
Thucydia
Emporia
Cycadaceae,
Zamiaceae
Cheirolepidaceae
Cephalotaxus,
Taxaceae
Araucariaceae
Cupressaceae
Pinaceae,
Podocarpaceae
BS_P = 100
B
Gnetum
Charcoalified
seeds
Welwitschia
Ephedra
Laurales
Chloranthaceae
Eupomatia
Magnoliaceae
Austrobaileya
Erdtmanithecales
Bennetitales
Callistophyton
Autunia
Peltaspermum
Corystosperms
Winteraceae
Piperaceae
Eudicots
Aristolochioidaea,
monocots
Nymphaeales
Caytonia
78
Glossopteris
Pentoxylon
Cecropteris
Archaeopteris
Tetraxylopteris
Lyginopteris
Heterangium
Laceya
Elkinsia
Bilignea
Lyrasperma
Medullosa,
Quaestora
Ginkgoales
Cycadaceae,
Zamiaceae
Cordaixylon
Mesoxylon
Shanxioxylon
Emporia, Thucydia
Cheirolepidaceae
BS_{NJ} = 100
Pinaceae,
Podocarpaceae
Taxaceae
Cephalotaxus
Araucariaceae,
Cupressaceae

Figure 4

Figure 6


A
15/<5
Cycadales
99/96
55/49
GCC clade
57/18
56/23
Cordaitales
97/46
53/31
20/11
Zamiaceae
Cycadaceae
Corystosperms
Peltasperms
Callistophyton
Ginkgo
Cordaixylon
Mesoxylon
Emporia
Conifers
17/9
<5/12
49/38
29/31
Caytonia
22/27
Glossopteris
Pinus
Podocarpus
70/58
Taxus
11/11
Pentoxylon
<5/21
Bennettitales
Erdtmanithecales
58/10
Charcoalified seeds
49/77
78/83
Quaestora
Medullosa
Medullosan seed ferns
40/22
44/61
69/29
Welwitschia
Gnetum
69/87
Ephedra
Gnetales
30/9
BEG clade
Lyginopteris
Heterangium
'Hyrdaspermalean' seed ferns
90/74
Elkinsia
19/17
Anthophytes
according
Rothwell et al., 2009, fig. 30
Cecropsis
Archaeopteris
Aneurophyton
81/89
84/95
'Progymnosperms' (outgroup)
MD = 0.1
BEG-angiosperm clade
23/16/
Eupomatia
Magnoliaceae
Austrobaileya
Winteraceae
Chloranthaceae
Laurales
Eudicots
Piperaceae
Monocots
Aristolochioideae
Nymphaeales
Angiosperms
99/99
B
GCC clade
42/16
55/22
8/9
50/33
Cordaitales
96/47
Mesoxylon
Cordaixylon
Ginkgo
Conifers
14/8
21/11
Emporia
56/51
99/96
61/56
Pinus
Podocarpus
Taxus
Callistophyton
Peltasperms
Zamiaceae
Cycadaceae
Glossopteris
30/33
Corystosperms
50/38
Caytonia
22/28
13/15
Pentoxylon
Bennettitales
49/77
Quaestora
79/83
Medullosa
Medullosan seed ferns.
39/23
44/60
BEG clade
15/6
61/56
Charcoalified seeds
84/99
Welwitschia
Gnetum
Gnetales
Ephedra
Lyginopteris
Heterangium
90/74
Elkinsia
'Hyrdasperman seed ferns'
<5/5
15/18
Anthophytes
according Rothwell et al., 2009, fig. 30
Cecropsis
Archaeopteris
Aneurophyton
83/89
83/95
'Progymnosperms'
BEG-angiosperm clade
11/17
24/30
Eupomatia
Magnoliaceae
Austrobaileya
Winteraceae
Chloranthaceae
Laurales
Eudicots
Piperaceae
Monocots
Aristolochioideae
Nymphaeales
Angiosperms
99/100
MD = 0.1
<5/8 Pentoxylon-Bennettitales-angiosperm clade
6/9 Bennettitales-angiosperm clade

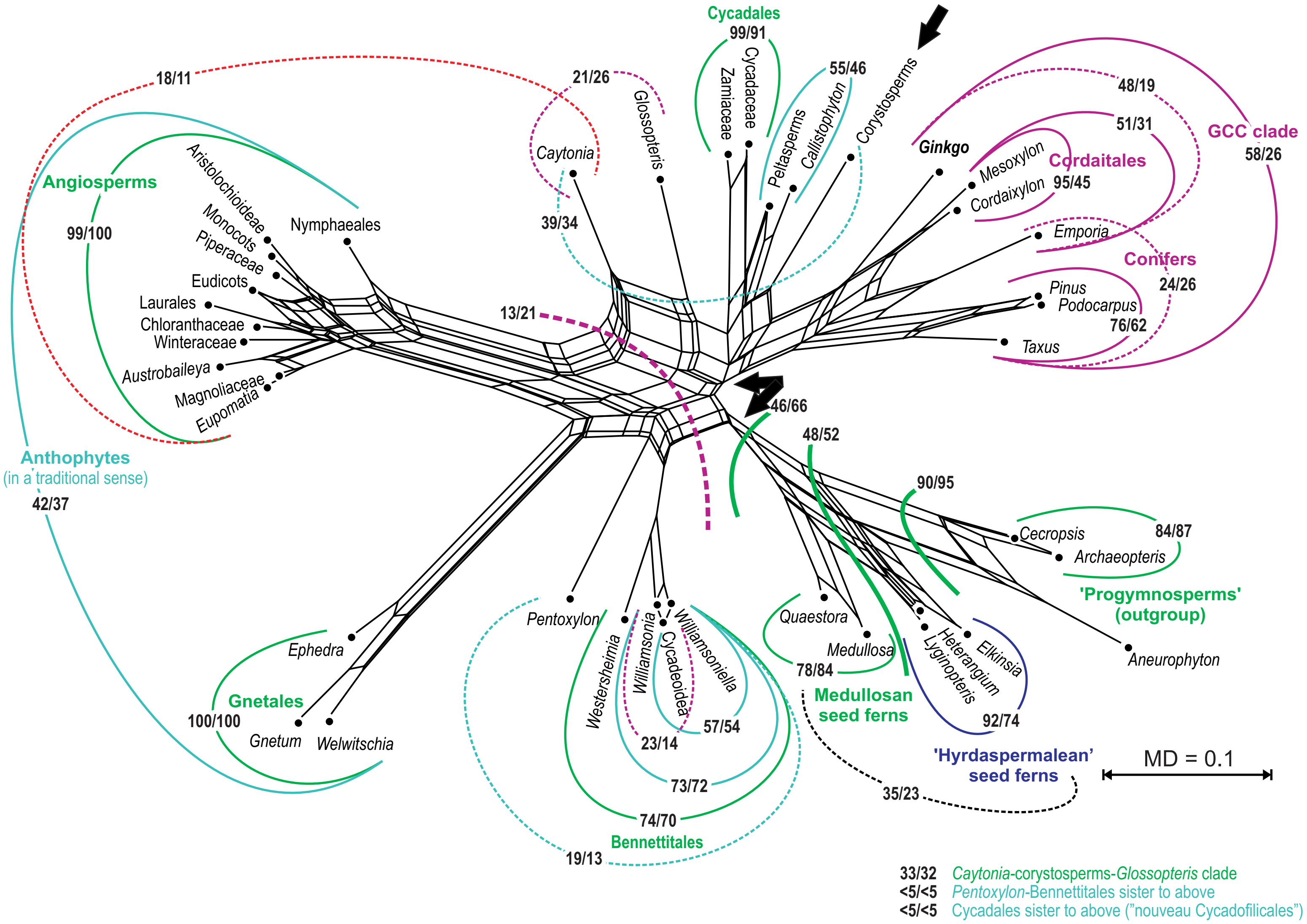

Cycadales
99/91
Zamiaceae
Cycadaceae
55/46
Peltasperms
Callistophyton
Corystosperms
18/11
21/26
Glossopteris
Caytonia
39/34
Angiosperms
99/100
Aristolochioideae
Monocots
Piperaceae
Nymphaeales
Eudicots
Laurales
Chloranthaceae
Winteraceae
Austrobaileya
Magnoliaceae
Eupomatia
13/21
Anthophytes
(in a traditional sense)
42/37
Ginkgo
48/19
51/31
GCC clade
58/26
Mesoxylon
Cordaitales
Cordaixylon
95/45
Emporia
Conifers
24/26
Pinus
Podocarpus
76/62
Taxus
46/66
48/52
90/95
Cecropsis
84/87
Archaeopteris
'Progymnosperms'
(outgroup)
Aneurophyton
Quaestora
Medullosa
78/84
Medullosan
seed ferns
Heterangium
Lyginopteris
Elkinsia
92/74
'Hyrdaspermalean'
seed ferns
35/23
MD = 0.1
Ephedra
Gnetales
100/100
Gnetum
Welwitschia
Pentoxylon
Westersheimia
Williamsonia
Cycadeoidea
Williamsoniella
23/14
57/54
73/72
74/70
Bennettitales
19/13
33/32 Caytonia-corystosperms-Glossopteris clade
<5/<5 Pentoxylon-Bennettitales sister to above
<5/<5 Cycadales sister to above ("nouveau Cycadofilicales")

Figure 7

TODO Table 3. Fossil taxa with incomplete data or ambiguous phylogenetic affinities. Listed for each taxon are the number of defined characters, its individual Delta value (iDV), and the bootstrap support for the clades, trees roots with the progymnosperms, including the taxon based on the Friis et al. Friis 2007 matrix and 'matrix 5' of Rothwell et al. Rothwell 2009.

Treelikeliness and phylogenetic signifance of optimal trees

The structure of the NN splits graphs and the complex differential support patterns (Figs 3–7) indicate that the phylogenetic signal from the matrices is often incompatible, and rarely unambiguous (in the sense that one placement in a tree is preferable over possible alternatives). This is corroborated by the matrix Delta values (mDV). All matrices comprise the same level of incompatibility (mDV ~ 0.31; Table 2), i.e. the matrices provide equally 'non-treelike' signals. This seems to contrast the number of MPT, and accordingly the resolution of the SC trees. The most non-treelike matrix (highest mDV) is the one of Crepet & Stevenson Crepet 2010, which, according to the authors, resulted in a well resolved strict consensus tree (based on four MPT, using NONA; but see Table 4). The mDV and individual Delta values (iDV) of Rothwell et al.'s matrix 3, which resulted in the best resolved SC tree, are in the same range than matrices resulting in much more MPT and accordingly less resolved SC trees (Table 4). The ranges of iDV show that each matrix is composed of taxa that, in general, induce a certain amount of ambiguous signal. As a trend, iDV of extinct taxa are higher than of extant taxa in all matrices, partly linked to the amount of missing data (Fig. 8). Extant taxa with high iDV (> median or average iDV of all taxa) in the matrix of Friis et al. (2007) are the members of the Cycadales and Gnetales, and the conifers *Taxus* fossil taxa with relatively low iDV (0.27–0.29; for comparison: angiosperms' iDVs range between 0.26–0.28) are the members comprising the 'hydraspermalean' seed ferns. In Rothwell et al.'s ‘matrices 3’, ‘4’, and ‘5’ [Rot3, Rot4, Rot5] Rothwell 2009, higher than median or average iDVs are found also in *Pinus* and *Podocarpus* ($iDV_{Rot3–5}$ c. 0.33). The taxon least compatible with tree-like evolution is the substantially recoded *Caytonia* ($iDV_{Rot5} = 0.38$) and *Pentoxylon* ($iDV_{Rot5} = 0.37$). The same hold for Crepet & Stevenson’s Crepet 2010 matrix (all iDV increased). Further taxa with high iDV are the taxa replacing the subdivided Bennettitales (Cycadeoidea, *Williamsonia*, *Williamsoniella*; iDV = 0.32–0.34), in particular the poorly defined (regarding number of characters) *Westerheimia* (iDV = 0.36). Thus, subdividing and recoding the Bennettitales, made the signal in the matrix even more ambiguous.

Taxa with high iDV account for the tree-inference problems (this includes procedures to establish branch support). In tree-inference, the indecisive signal from taxa like the Cycadales, *Glossopteris* and *Pentoxylon* result in many equally optimal or nearly optimal topologies (Fig. 9); each of them can be placed in different subtrees (clades) indicating various kinds of relationships. Typically all possible bipartitions separating a taxon (or group of taxa) with high iDV and one or several other taxa from the remainder will receive low support (Figs 4, 6, 7; Table 2). Eliminating taxa with high iDV from the matrix will increase directly the decisiveness of the phylogenetic inference. If the tree taxa with the highest iDV in the matrix of Friis et al. Friis 2007 are excluded (*Glossopteris, Pentoxylon,* corystosperms), support for the ‘anthophyte’ clade increases from low to moderate ($BS_{NJ} = 29 \rightarrow 53$), and a potential clade comprising *Autunia*, *Callistophyton,* and *Peltaspermum* would also receive moderate support ($BS_{NJ} = 63$). Deletion of the three extant (three angiosperms) or extinct taxa (three ‘hydraspermalean’ seed ferns) with lowest iDV has little effect on the inference, since all these taxa are anyhow well-embedded in according clades, i.e. they are consistently placed in the same subtrees, without affecting relationships outside the clade. Phylogenetic networks, trimmed to handle incompatible signal, are (should be) naturally less affected by inclusion or exclusion of taxa with high iDV (according graphs provided in ES).

Table 4 shows that the matrices analysed here only markedly differ in the number of obtained MPT (#MPT), and length of the MPT. The CI and RI are relatively low for all analyses, indicating a high amount of homoplasy in all of the underlying matrices. Given the low CI/RI values of the MPT and the level of mDVs and iDVs, one could expect that there should be a numerous possible, equally optimal phylogenetic trees. However, this is only found for the Friis et al. Friis 2007 matrix (Fig. 9), which, although being the largest matrix, results in MPT that are about 30 steps shorter. The low number of MPT (Table 4) found by Rothwell et al. Rothwell 2009 and Crepet & Stevenson Crepet 2010 clearly does not effectively represent the complexity of the phylogenetic signal in the matrices. This is most obvious in the case of *Caytonia*, which has the highest iDV in the according matrices, but is unambiguously placed in the MPT and subsequent SC trees as sister to *Glossopteris* (Rothwell et al., 2009, ‘matrix 3’) or corystosperms Crepet & Stevenson, 2010 in a *Caytonia*-corystosperms-*Glossopteris* clade. Given our results and the mDV and iDV values of the matrices (Table 4; ES), PAUP*’s heuristic search algorithms using the program’s default settings find a more realistic number of MPT for all matrices, but still too few in the case of ‘matrix 3’ of Rothwell et al. Rothwell 2009 and Crepet & Stevenson Crepet 2010. The number of slightly suboptimal topologies (Table 4) demonstrates that the found, limited numbers of MPT are not substantially better than alternative trees. Both the ML and NJ trees are (slightly) suboptimal under parsimony (regarding number of steps), which further demonstrates that the data provides little unambiguous signal. Vice versa, the SH-test did not reject any of the MPT as significantly worse than the RAxML-inferred ‘best-known’ tree. Based on these results it can be concluded that the mere topologies of optimal trees are of little relevance, if the intrinsic phylogenetic signal of the here analysed matrices should be comprehensively displayed.

Maybe add: possible reason for the few MPT is the many-state characters, which essentially increase the number of “autapomorphies” but also very restricted “synapomorphic” states and convergences.

TODO **Figure 8**. Plot of iDV vs. portion of defined characters (incl. polymorphic-coded).

TODO **Figure 9**. Strict consensus network of >20,000 MPT inferred using heuristic search on the matrix of Friis et al. Friis 2007

Table 4. Compilation of matrix- and reconstruction-related parameters.

Matrix	Friis et al., 2007	Rothwell et al., 2009, matrix 5	Rothwell et al., 2009, matrix 3	Crepet & Stevenson, 2010, matrix 1
Dimensions (Taxa × Characters)	50 × 102	40 × 110	39 × 110	41 × 111
mDV	0.308	0.309	0.303	0.312
#MPT (PAUP*)	21,974	684	60	180
#MPT (NONA)[a]	1,180	252	12	4
#MPT+1 (PAUP*)	9999	9999	1948	9999
Length MPT (MESQUITE/PAUP*)	329	365	360	364/365
Length MPT (NONA)[a]	314	334	328	327

Taken from Rothwell et al., 2009; Crepet & Stevenson, 2010

Conclusion

This little exercise shows the limitations and pitfalls of approaches that try to infer phylogenetic relationships based on morphological matrices by the mere inference of a somehow optimized tree, typically a strict consensus tree based on a sample of most-parsimonious solutions. In the tradition of plant phylogenetic studies, researchers have put a lot of stress on how to code characters for morphological traits (see e.g. Rothwell et al.'s extensive critique of character coding applied in Hilton 2006 and Friis 2007; see also works by Doyle 9999) and have grown somewhat reluctant regarding the actual nature of the phylogenetic signal that can be retrieved from such matrices. Despite their recoding efforts, neither Rothwell et al. Rothwell 2009 nor Crepet & Stevenson Crepet 2010 succeeded in producing matrices that would support fundamentally different relationships than the matrices of Hilton & Bateman Hilton 2006 and Friis et al. Friis 2007. The result of the recoding, here in particular the combination of simple binary, or few-state characters in the earlier matrices, into complex multi-state characters in the later matrices, was only that originally fairly congruent relationships among different methods and approaches, became less conclusive and less congruent. Plant morphological matrices are unlikely to produce unambiguous relationships on their own, why it is important to explore the many signals within the data and not to manipulate the data via coding in order to reduce the number of equally optimal tree topologies. Methods for such an exploration are available (e.g. consensus networks based on collections of trees, distance-based phylogenetic networks) and easy to use. Furthermore, this exercise shows that it is always risky to rely on a single approach to put forward a hypothesis. But it is naturally much less risky to accept relationships that find (some) support from different optimisations (here: BI, ML, MP, LS/NJ), which all have a more or less different angle on how evolution affects differentiation patterns. If they converge, the according relationship cannot be entirely circumstantial.