

Research Article

Exploring Diversification and Genome Size Evolution in Extant Gymnosperms through Phylogenetic Synthesis

J. Gordon Burleigh,¹ W. Brad Barbazuk,¹ John M. Davis,²
Alison M. Morse,² and Pamela S. Soltis³

¹Department of Biology, University of Florida, Gainesville, FL 32611, USA

²School of Forest Resources and Conservation, University of Florida, Gainesville, FL 32611, USA

³Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, USA

Correspondence should be addressed to J. Gordon Burleigh, gburleigh@ufl.edu

Received 2 June 2011; Accepted 20 September 2011

Academic Editor: Hiroyoshi Takano

Copyright © 2012 J. Gordon Burleigh et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Gymnosperms, comprising cycads, *Ginkgo*, Gnetales, and conifers, represent one of the major groups of extant seed plants. Yet compared to angiosperms, little is known about the patterns of diversification and genome evolution in gymnosperms. We assembled a phylogenetic supermatrix containing over 4.5 million nucleotides from 739 gymnosperm taxa. Although 93.6% of the cells in the supermatrix are empty, the data reveal many strongly supported nodes that are generally consistent with previous phylogenetic analyses, including weak support for Gnetales sister to Pinaceae. A lineage through time plot suggests elevated rates of diversification within the last 100 million years, and there is evidence of shifts in diversification rates in several clades within cycads and conifers. A likelihood-based analysis of the evolution of genome size in 165 gymnosperms finds evidence for heterogeneous rates of genome size evolution due to an elevated rate in *Pinus*.

1. Introduction

Recent advances in sequencing technology offer the possibility of identifying the genetic mechanisms that influence evolutionarily important characters and ultimately drive diversification. Within angiosperms, large-scale phylogenetic analyses have identified complex patterns of diversification (e.g., [1–3]), and numerous genomes are at least partially sequenced. Yet the other major clade of seed plants, the gymnosperms, have received far less attention, with few comprehensive studies of diversification and no sequenced genomes. Note that throughout this paper “gymnosperms” specifies only the approximately 1000 extant species within cycads, *Ginkgo*, Gnetales, and conifers. These comprise the *Acrogymnospermae* clade described by Cantino et al. [4].

Many gymnosperms have exceptionally large genomes (e.g., [5–7]), and this has hindered whole-genome sequencing projects, especially among economically important *Pinus* species. This large genome size is interesting because one suggested mechanism for rapid increases in genome size,

polyploidy, is rare among gymnosperms [8]. Recent sequencing efforts have elucidated some of genomic characteristics associated with the large genome size in *Pinus*. Morse et al. [9] identified a large retrotransposon family in *Pinus*, that, with other retrotransposon families, accounts for much of the genomic complexity. Similarly, recent sequencing of 10 BAC (bacterial artificial chromosome) clones from *Pinus taeda* identified many conifer-specific LTR (long terminal repeat) retroelements [10]. These studies suggest that the large genome size may be caused by rapid expansion of retrotransposons and may be limited to conifers, Pinaceae, or *Pinus*. Other studies have quantified patterns of genome size among gymnosperms, especially within *Pinus* and the other Pinaceae [6, 7, 11–14]. These studies have largely focused on finding morphological, biogeographic, or life history correlates of genome size, but the rates and patterns of genome size evolution in gymnosperms are largely unknown.

This study first synthesizes the available phylogenetically informative sequences to build a phylogenetic hypothesis of

gymnosperms that reflects the recent advances in sequencing and computational phylogenetics. The resulting tree provides a starting point for large-scale evolutionary and ecological analyses of gymnosperms and will hopefully be a resource to promote and guide future phylogenetic and comparative studies. We use the tree to examine large-scale patterns of diversification of the extant gymnosperm lineages and also to examine rates of genome size evolution.

2. Methods and Materials

2.1. Supermatrix Phylogenetic Inference. We constructed a phylogenetic hypothesis of gymnosperms from available, phylogenetically informative sequence data in GenBank that was available on June 30, 2009. We first downloaded from GenBank all core nucleotide sequence data from gymnosperms (Coniferophyta, Cycadophyta, Ginkgophyta, and Gnetophyta). Additionally, we downloaded sequences from the “basal angiosperm” lineages (e.g., *Amborella*, Nymphaeales, Chloranthaceae, and Austrobaileyales) to represent the angiosperms and a diverse sampling of Moniliformopses taxa (including species from *Equisetum*, *Psilotum*, *Ophioglossum*, *Botrychium*, *Angiopteris*, and *Adiantum*) to use as outgroups.

To identify sets of homologous sequences from the GenBank data, we clustered sequences less than 10,000 bp in length based on results from an all-by-all pairwise BLAST analysis. The all-by-all blastn search was done with blastall using the default parameters [15]. Significant BLAST hits had a maximum e-value of $1.0e^{-10}$ and at least 50% overlap of both the target and query sequences. A Perl script identified the largest clusters of sequences in which each sequence has a significant BLAST hit against at least one other sequence in the cluster. We only considered clusters containing loci that had been used previously for phylogenetic analyses. This included plastid and mitochondrial loci as well as nuclear 18S rDNA, 26S rDNA, and internal transcribed spacer (ITS) sequences. Among these clusters, those containing sequences from at least 15 taxa were aligned using Muscle [16], and the resulting alignments were manually checked and adjusted. The resulting alignments were edited for inclusion in the supermatrix by removing hybrid taxa and those that lacked a specific epithet and also keeping only a single sequence per species. The final cluster alignments were then concatenated to make a single phylogenetic supermatrix (e.g., [17]).

2.2. Phylogenetic and Dating Analysis. To estimate the optimal topology and molecular branch lengths for the gymnosperms, we performed maximum likelihood (ML) phylogenetic analysis on the full supermatrix alignment using RAXML-VI-HPC version 7.0.4 [18]. All ML analyses used the general time reversible (GTR) nucleotide substitution model with the default settings for the optimization of individual per-site substitution rates and classification of these rates into rate categories. To assess uncertainty in the topology and branch length estimates, we ran 100 nonparametric bootstrap replicates on the original data set [19].

We transformed the optimal and bootstrap trees to chronograms, ultrametric trees in which the branch lengths

represent time, using penalized likelihood [20] implemented in r8s version 1.71 [21]. We used a smoothing parameter of 10000, which was chosen based on cross-validation of the fossil constraints. For the r8s analysis, we used the same time constraints on seed plant clades used by Won and Renner [22]. The most recent common ancestor of seed plants was constrained to a maximum age of 385 million years ago (mya). The most recent common ancestor of the extant gymnosperms was fixed at 315 mya and *Gnetum* at 110 mya. The following clades were given minimum age constraints: angiosperms: 125 mya, cycads: 270 mya, Cupressaceae: 90 mya, Araucariaceae: 160 mya, Gnetales + Pinaceae: 225 mya, Pinaceae: 90 mya, and Gnetales: 125 mya.

2.3. Diversification Analysis. To examine the general patterns of diversification through time among the extant gymnosperm lineages, we first made lineage through time plots using the R package APE [23]. To account for uncertainty in the dating estimates, we plotted each bootstrap tree after it had been transformed into a chronogram and all nongymnosperm taxa were removed.

Since there appears to be much variance in the divergence time estimates among trees, and branch length estimates are often unreliable, especially when estimated from such a sparse, heterogeneous sequence matrix, we used a test for changes in diversification rate that relies on tree shape, not branch lengths. Specifically, we used the whole-tree, topology-based test described by Moore et al. [24] to detect nodes associated with significant shifts in diversification rate based on the Δ_1 statistic. The analyses were performed using the apTreeshape R package [25]. We used only the optimal tree estimate and again, pruned all non-gymnosperm taxa from the tree prior to analysis.

2.4. Rates of Genome Size Evolution. We first assembled a set of mean genome size data for all gymnosperms present in the phylogenetic tree (in pg DNA) from the Kew C-value database [26]. This includes data from the studies of Murray [6] and Grotkopp et al. [14]. When there were multiple estimates available from a single species, we used the mean of the estimates. We tested for shifts in rates of genome size evolution using Brownie v. 2.1.2 [27]. We used the censored rate test, which tests for differences in rates of evolution of a continuous character (genome size) in one clade versus another clade or paraphyletic group based on a Brownian motion model. We made the following comparisons: conifers + Gnetales versus cycads + *Ginkgo*, Pinaceae versus other conifers + Gnetales, non-Pinaceae conifers + Gnetales versus cycads + *Ginkgo*, *Pinus* versus other Pinaceae, the non-*Pinus* Pinaceae versus the other conifers + Gnetales, and *Pinus* subgenus *Strobus* versus *Pinus* subgenus *Pinus*. To account for topological and branch length uncertainty, we performed all hypothesis tests in Brownie on each bootstrap tree and weighted the results across replicates. The penalized likelihood analysis in r8s collapsed some branch lengths to 0, and Brownie does not work with 0 branch lengths in the tree. Thus, prior to the Brownie analysis, all 0 branch lengths were changed to 0.1.

3. Results

3.1. Phylogenetic Data. The alignment from the complete supermatrix contains sequences from 950 taxa (739 gymnosperms, 108 angiosperms, and 103 nonseed plant outgroups) and is 74,105 characters in length. The 739 gymnosperm taxa include at least one representative from every family as well as from 88 genera. In total, the matrix contains 4,511,144 nucleotides and 93.6% missing data. The number of nucleotides per taxon ranges from 252 to 33,138 (average = 4,749; median = 3,355).

3.2. Phylogenetic Inference. In the 950-taxon trees, 63.3% (601) of the nodes have at least 50% bootstrap support, 41.7% (396) have at least 70% support, 25.8% (245) have at least 90% support, and 9.7% (92) have 100% support. The seed plants have 100% support, and the angiosperms are sister to a clade of all gymnosperms (Figure 1). Within gymnosperms, a clade of *Ginkgo* + cycads (bootstrap support (BS) = 66%) is sister to a clade consisting of conifers + Gnetales (BS = 96%). Gnetales are sister to Pinaceae within the conifers, although the “Gne-Pine” clade has only 57% support. Within the major groups of gymnosperms (conifers, Gnetales, and cycads), family-level and generic relationships generally are congruent with those inferred in other analyses (Figure 1). Of the 54 gymnosperm genera represented by more than one species in the tree, 47 have at least 50% bootstrap support, 36 have at least 90% bootstrap support, and 26 have 100% bootstrap support. A full version of the bootstrap consensus tree is available as Supplemental Material.

3.3. Diversification. Although the lineage through time plots display much variation among bootstrap replicates (Figure 2), the general trend among the bootstrap trees is similar, with what appears to be high and possibly increasing diversification over the last 100 million years. Still, lineage through time plots are imprecise and difficult to interpret. If this trend of high recent diversification were true, we would expect to find evidence of increased rates of diversification in some relatively young clades.

The Δ_1 statistic indicated a significant shift in the rates of diversification at 10 nodes in the tree. Several are within the cycads. This includes the node dividing *Cycas* and *Epicycas* species from the other cycads ($P = 0.0474$) and its daughter node separating *Cycas*, *Epicycas*, and *Dioon* from the other cycads ($P = 0.157$). Also, two basal-most nodes of *Zamia* show significant shifts in diversification rates ($P = 0.014$ and 0.316). Within conifers, there is a significant shift in diversification at the most recent common ancestor of *Podocarpus* ($P = 0.017$). Also, there are significant shifts in diversification at the two basal nodes of Cupressaceae ($P = 0.0326$ and 0.0366) and within Cupressaceae, at the most recent common ancestor of *Callitris*, *Neocallitropsis*, *Actinostrobus*, *Widdringtonia*, *Fitzroya*, *Diselma*, and *Austrocedrus* ($P = 0.0387$). Finally, there is a significant shift in two of the basal nodes of *Picea* ($P = 0.0166$, $P = 0.0029$).

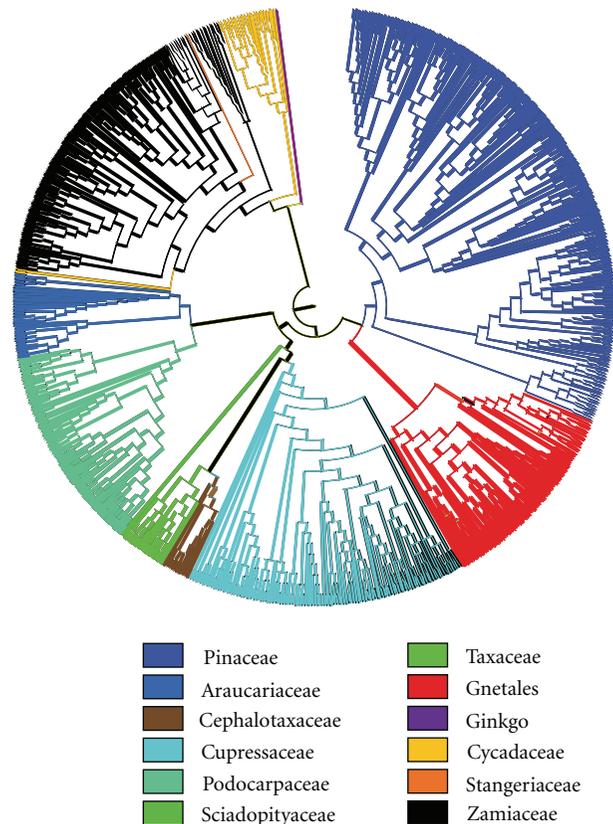


FIGURE 1: Overview of the ML tree of 739 gymnosperm taxa; angiosperms and outgroups have been removed. Colors represent the different families of conifers (Pinaceae, Araucariaceae, Cephalotaxaceae, Cupressaceae, Podocarpaceae, Sciadopityaceae, and Taxaceae), Gnetales, *Ginkgo*, and the families of cycads (Cycadaceae, Stangeriaceae, and Zamiaceae). A full bootstrap consensus tree is available as Supplementary Material available online at doi:10.1155/2011/292857.

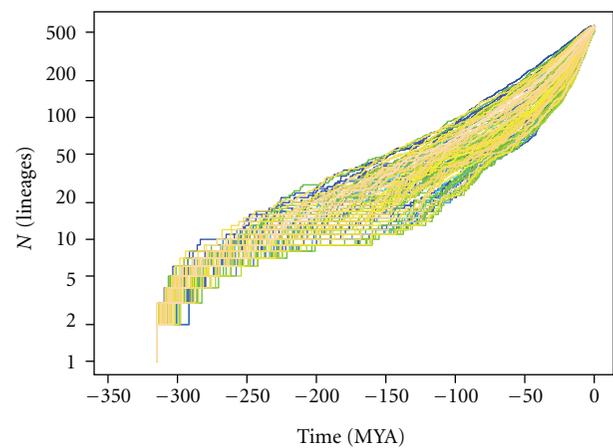


FIGURE 2: Lineage through time plot for the gymnosperm species. All bootstrap trees, with ultrametric branch lengths from r8s, were pruned to include only the gymnosperm taxa. Each line represents a single ML bootstrap tree. The graph shows the pattern of diversification of the gymnosperm taxa in the tree through time, as the tree grew from a single lineage at the root to the current sampling of 739 species.

3.4. Genome Size Evolution. Based on the large size of genomes of *Pinus* species, we hypothesized that there may be an increase in the rate of genome size evolution (Figure 3). We performed a series of likelihood ratio tests to examine the patterns of rate variation throughout gymnosperms, with a focus on testing for rate variation associated with conifers, Pinaceae, and *Pinus* (Table 1). In all comparisons in which *Pinus* (or a group containing *Pinus*) was compared to another group, the group with *Pinus* showed significantly elevated rates of genome size evolution (Table 1). We detected no significant shifts in rates of evolution between any groups that did not contain *Pinus*, and there was no significant difference in rates of evolution between the two *Pinus* subgenera (*Pinus* and *Strobus*; Table 1).

4. Discussion

The analyses of gymnosperm diversification and genome size evolution demonstrate the dynamic evolutionary processes of the extant gymnosperms, which sharply contrasts with their reputation as ancient, relictual species. The lineage through time plots are consistent with high, and possible growing, rates of diversification within the last 100 million years, concurrent with major radiations of angiosperms (e.g., [1, 2, 28]) and extant ferns [29]. There is evidence of numerous significant shifts in diversification within both cycads and conifers, and there is strong evidence for a recent, large increase in the rate of genome size evolution in *Pinus*. Although *Pinus* is a species-rich genus, we find no links between increased rates of diversification and shifts in rates of genome size evolution.

Advances in sequencing technology and computational biology over the past decade enable phylogenetic estimates comprising large sections the plant diversity. This study demonstrates that it is possible to construct credible phylogenetic hypotheses including nearly three quarters of the extant gymnosperm species. Unlike supertree approaches (e.g., [14]), the supermatrix methods easily incorporate branch length estimates and estimates of topological and branch length uncertainty. Still, until there is far more data per taxon, estimates of the gymnosperm phylogeny will continue to improve, and thus, it is important to consider error and uncertainty in phylogenetic estimates when using these trees to infer evolutionary processes. There are other reasons to interpret this gymnosperm tree with caution. For example, both heterogeneity in the patterns of molecular evolution and missing data can lead to erroneous estimates of trees and branch lengths in ML phylogenetic analyses (e.g., [30, 31]). Furthermore, our analysis does not attempt to incorporate evolutionary processes, such as incomplete lineage sorting or gene duplication and loss or reticulation, that may cause incongruence between the gene trees and the species phylogeny (e.g., [32]). Although this study used thousands of sequences, it does not incorporate the evolutionary perspective of low-copy nuclear genes.

Still, in many cases, evolutionary or ecological analyses that use phylogenetic trees may be robust to topological and branch length error (e.g., [33]), and the large tree of

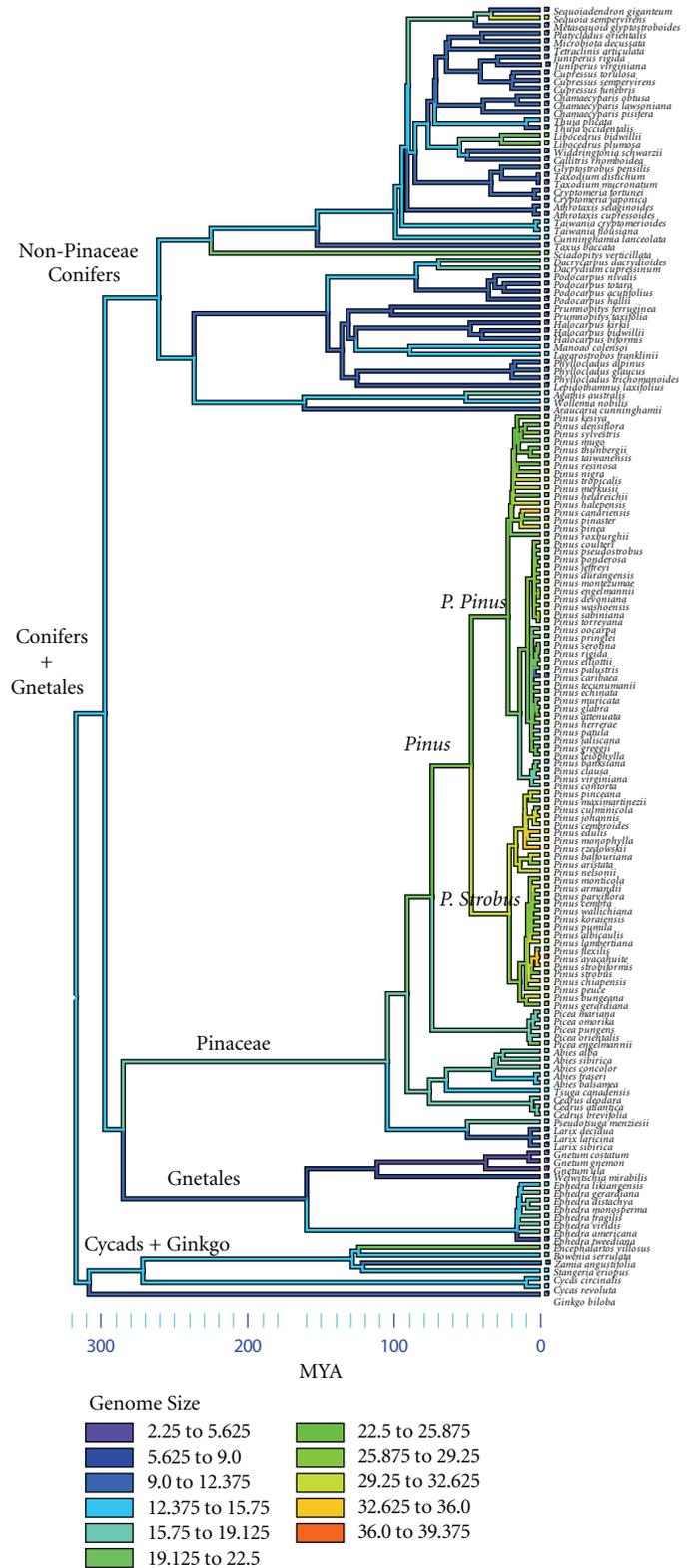


FIGURE 3: Ancestral state reconstruction of genome size (in pg DNA) on a chronogram 165 gymnosperm taxa. Different genome sizes are represented by different colors, with the ancestral genome sizes estimated with squared change parsimony.

TABLE 1: Rate estimates from the two rate parameter models from Brownie. *Indicate that the single rate model was rejected based on the Chi-squared P value (** $P < 0.005$; *** $P < 0.0005$). Significance was also assessed using AIC.

Comparison	Rates of Genome Evolution
Conifers + Gnetales	1.878***
Cycads + <i>Ginkgo</i>	0.095
Pinaceae	2.715***
Other conifers + Gnetales	0.178
Other conifers + Gnetales	0.178
Cycads + <i>Ginkgo</i>	0.095
<i>Pinus</i>	3.234**
Other Pinaceae	0.431
Other Pinaceae	0.431
Other conifers + Gnetales	0.178
<i>Pinus</i> Strobis subgenus	2.66
<i>Pinus</i> Pinus subgenus	3.56

gymnosperms enables sophisticated and comprehensive tests of evolutionary and ecological hypotheses. We demonstrate this with our diversification analysis, the results of which emphasize numerous, independent shifts in diversification rate throughout gymnosperms and apparently recent, high rates of diversification (Figure 2). Estimates of diversification may be affected by taxonomic sampling and inaccurate branch length estimates. However, we might expect that adding the remaining species, which would likely fit near the tips of the tree, would result in increased estimates of recent diversification. Thus, our analyses suggest the intriguing perspective that the extant gymnosperms are a vibrant, growing clade, and not simply the sole survivors of ancient diversity. Greater sampling and a more robust tree will provide a more complete view of gymnosperm diversification. With better branch length estimates, it will be possible to use more powerful likelihood-based approaches to identify clades with increasing and decreasing diversification rates [34]. With more complete taxon sampling, it may be possible to identify characters associated with changing speciation and extinction rates ([35], but see [36]).

One of the great challenges of evolutionary genomics is to identify the mechanisms of genome evolution that drive diversification. Some of the mechanisms that cause changes in genome size, such as whole-genome duplications or activity of retrotransposons, can have implications on diversification rates. Our analysis of the rates of genome size evolution demonstrate that *Pinus* is unique among gymnosperms. That is, the highly elevated rates of change in genome size appear to be limited to *Pinus*. However, in gymnosperms, we find no evidence of increases in diversification associated with *Pinus*, which displays a significantly elevated rate of genome size evolution. Furthermore, we find no obvious evidence for increase in rates of genome size evolution in clades associated with shifts in diversification. While our analysis failed to link genome size and diversification, this comparative approach for identifying shifts in genome size can inform our search for the specific drivers of the increased genomic complexity

in *Pinus*, and this ultimately can help inform strategies for sequencing and assembling the first *Pinus* genomes.

Supplementary Materials

The nucleotide and C-value data matrices along with all trees are available on the Dryad data repository (<http://datadryad.org/>).

Acknowledgment

This work was funded by a University of Florida Research Opportunity Fund Seed Grant.

References

- [1] S. Magallón and A. Castillo, “Angiosperm diversification through time,” *American Journal of Botany*, vol. 96, no. 1, pp. 349–365, 2009.
- [2] C. D. Bell, D. E. Soltis, and P. S. Soltis, “The age and diversification of the angiosperms re-revisited,” *American Journal of Botany*, vol. 97, no. 8, pp. 1296–1303, 2010.
- [3] S. A. Smith, J. M. Beaulieu, A. Stamatakis, and M. J. Donoghue, “Understanding angiosperm diversification using small and large phylogenetic trees,” *American Journal of Botany*, vol. 98, no. 3, pp. 404–414, 2011.
- [4] P. D. Cantino, J. A. Doyle, S. W. Graham et al., “Towards a phylogenetic nomenclature of Tracheophyta,” *Taxon*, vol. 56, no. 3, pp. 822–846, 2007.
- [5] D. Ohri and T. N. Khoshoo, “Genome size in gymnosperms,” *Plant Systematics and Evolution*, vol. 153, no. 1-2, pp. 119–132, 1986.
- [6] B. G. Murray, “Nuclear DNA amounts in gymnosperms,” *Annals of Botany*, vol. 82, pp. 3–15, 1998.
- [7] M. R. Ahuja and D. B. Neale, “Evolution of genome size in conifers,” *Silvae Genetica*, vol. 54, no. 3, pp. 126–137, 2005.
- [8] T. N. Khoshoo, “Polyploidy in gymnosperms,” *Evolution*, vol. 13, no. 1, pp. 24–39, 1958.
- [9] A. M. Morse, D. G. Peterson, M. N. Islam-Faridi et al., “Evolution of genome size and complexity in *Pinus*,” *PLoS ONE*, vol. 4, no. 2, Article ID e4332, 2009.
- [10] A. Kovach, J. L. Wegrzyn, G. Parra et al., “The *Pinus taeda* genome is characterized by diverse and highly diverged repetitive sequences,” *BMC Genomics*, vol. 11, no. 1, article 420, 2010.
- [11] K. L. Joyner, X.-R. Wang, J. S. Johnston, H. J. Price, and C. G. Williams, “DNA content for Asian pines parallels new world relatives,” *Canadian Journal of Botany*, vol. 79, no. 2, pp. 192–196, 2001.
- [12] S. E. Hall, W. S. Dvorak, J. S. Johnston, H. J. Price, and C. G. Williams, “Flow cytometric analysis of DNA content for tropical and temperate new world pines,” *Annals of Botany*, vol. 86, no. 6, pp. 1081–1086, 2000.
- [13] I. Wakamiya, R. J. Newton, J. S. Johnston, and H. J. Price, “Genome size and environmental factors in the genus *Pinus*,” *American Journal of Botany*, vol. 80, no. 11, pp. 1235–1241, 1993.
- [14] E. Grotkopp, M. Rejmánek, M. J. Sanderson, and T. L. Rost, “Evolution of genome size in pines (*Pinus*) and its life-history correlates: supertree analyses,” *Evolution*, vol. 58, no. 8, pp. 1705–1729, 2004.

- [15] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman, "Basic local alignment search tool," *Journal of Molecular Biology*, vol. 215, no. 3, pp. 403–410, 1990.
- [16] J. D. Thompson, D. G. Higgins, and T. J. Gibson, "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice," *Nucleic Acids Research*, vol. 22, no. 22, pp. 4673–4680, 1994.
- [17] A. de Queiroz and J. Gatesy, "The supermatrix approach to systematics," *Trends in Ecology and Evolution*, vol. 22, no. 1, pp. 34–41, 2007.
- [18] A. Stamatakis, "RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models," *Bioinformatics*, vol. 22, no. 21, pp. 2688–2690, 2006.
- [19] J. Felsenstein, "Confidence limits on phylogenies: an approach using the bootstrap," *Evolution*, vol. 39, no. 4, pp. 783–791, 1985.
- [20] M. J. Sanderson, "Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach," *Molecular Biology and Evolution*, vol. 19, no. 1, pp. 101–109, 2002.
- [21] M. J. Sanderson, "R8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock," *Bioinformatics*, vol. 19, no. 2, pp. 301–302, 2003.
- [22] H. Won and S. S. Renner, "Dating dispersal and radiation in the gymnosperm *Gnetum* (Gnetales)—clock calibration when outgroup relationships are uncertain," *Systematic Biology*, vol. 55, no. 4, pp. 610–622, 2006.
- [23] E. Paradis, J. Claude, and K. Strimmer, "APE: analyses of phylogenetics and evolution in R language," *Bioinformatics*, vol. 20, no. 2, pp. 289–290, 2004.
- [24] B. R. Moore, K. M. A. Chan, and M. J. Donoghue, "Detecting diversification rate variation in supertrees," in *Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life*, O. R. P. Bininda-Emonds, Ed., pp. 487–533, Kluwer Academic, Dodrecht, The Netherlands, 2004.
- [25] N. Bortolussi, E. Durand, M. G. B. Blum, and O. François, "Aptreeshape: statistical analysis of phylogenetic treeshape," *Bioinformatics*, vol. 22, no. 3, pp. 363–364, 2006.
- [26] M. D. Bennett and I. J. Leitch, "Plant DNA C-values database," 2005, <http://data.kew.org/cvalues/>.
- [27] B. C. O'Meara, C. Ané, M. J. Sanderson, and P. C. Wainwright, "Testing for different rates of continuous trait evolution using likelihood," *Evolution*, vol. 60, no. 5, pp. 922–933, 2006.
- [28] H. Wang, M. J. Moore, P. S. Soltis et al., "Rosid radiation and the rapid rise of angiosperm-dominated forests," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 10, pp. 3853–3858, 2009.
- [29] H. Schneider, E. Schuettpelz, K. M. Pryer, R. Cranfill, S. Magallón, and R. Lupia, "Ferns diversified in the shadow of angiosperms," *Nature*, vol. 428, no. 6982, pp. 553–557, 2004.
- [30] B. Kolaczkowski and J. W. Thornton, "Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous," *Nature*, vol. 431, no. 7011, pp. 980–984, 2004.
- [31] A. R. Lemmon, J. M. Brown, K. Stanger-Hall, and E. M. Lemmon, "The effect of ambiguous data on phylogenetic estimates obtained by maximum likelihood and bayesian inference," *Systematic Biology*, vol. 58, no. 1, pp. 130–145, 2009.
- [32] W. P. Maddison, "Gene trees in species trees," *Systematic Biology*, vol. 46, no. 3, pp. 523–536, 1997.
- [33] E. A. Stone, "Why the phylogenetic regression appears robust to tree misspecification," *Systematic Biology*, vol. 60, no. 3, pp. 245–260, 2011.
- [34] D. L. Rabosky, "LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates," *Evolutionary Bioinformatics*, vol. 2, pp. 247–250, 2006.
- [35] W. P. Maddison, P. E. Midford, and S. P. Otto, "Estimating a binary character's effect on speciation and extinction," *Systematic Biology*, vol. 56, no. 5, pp. 701–710, 2007.
- [36] D. L. Rabosky, "Extinction rates should not be estimated from molecular phylogenies," *Evolution*, vol. 64, no. 6, pp. 1816–1824, 2010.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

