Syst. Biol. 55(4):610–622, 2006

ISSN: 1063-5157 print / 1076-836X online
DOI: 10.1080/10635150600812619

Dating Dispersal and Radiation in the Gymnosperm *Gnetum* (Gnetales)—Clock Calibration When Outgroup Relationships Are Uncertain

HYOSIG WON[1,2] AND SUSANNE S. RENNER[1,3]

[1]*University of Missouri-St. Louis, Department of Biology, 8001 Natural Bridge Road, St. Louis, Missouri 63121, USA; and The Missouri Botanical Garden, St. Louis, Missouri 63166, USA; E-mail: wonhs@snu.ac.kr (H.W.); renner@umsl.edu (S.R.R.)*
[2]*Present Address: Seoul National University, College of Natural Sciences, School of Biological Sciences, Seoul, Korea 151-742*
[3]*Present Address: Department of Biology, Ludwig Maximilians University, D-80638 Munich, Germany; E-mail: renner@lrz.uni-muenchen.de*

Abstract.—Most implementations of molecular clocks require resolved topologies. However, one of the Bayesian relaxed clock approaches accepts input topologies that include polytomies. We explored the effects of resolved and polytomous input topologies in a rate-heterogeneous sequence data set for *Gnetum*, a member of the seed plant lineage Gnetales. *Gnetum* has 10 species in South America, 1 in tropical West Africa, and 20 to 25 in tropical Asia, and explanations for the ages of these disjunctions involve long-distance dispersal and/or the breakup of Gondwana. To resolve relationships within *Gnetum*, we sequenced most of its species for six loci from the chloroplast (*rbcL*, *matK*, and the *trnT-trnF* region), the nucleus (rITS/5.8S and the *LEAFY* gene second intron), and the mitochondrion (*nad1* gene second intron). Because *Gnetum* has no fossil record, we relied on fossils from other Gnetales and from the seed plant lineages conifers, *Ginkgo*, cycads, and angiosperms to constrain a molecular clock and obtain absolute times for within-*Gnetum* divergence events. Relationships among Gnetales and the other seed plant lineages are still unresolved, and we therefore used differently resolved topologies, including one that contained a basal polytomy among gymnosperms. For a small set of Gnetales exemplars ($n = 13$) in which *rbcL* and *matK* satisfied the clock assumption, we also obtained time estimates from a strict clock, calibrated with one outgroup fossil. The changing hierarchical relationships among seed plants (and accordingly changing placements of distant fossils) resulted in small changes of within-*Gnetum* estimates because topologically closest constraints overrode more distant constraints. Regardless of the seed plant topology assumed, relaxed clock estimates suggest that the extant clades of *Gnetum* began diverging from each other during the Upper Oligocene. Strict clock estimates imply a mid-Miocene divergence. These estimates, together with the phylogeny for *Gnetum* from the six combined data sets, imply that the single African species of *Gnetum* is not a remnant of a once Gondwanan distribution. Miocene and Pliocene range expansions are inferred for the Asian subclades of *Gnetum*, which stem from an ancestor that arrived from Africa. These findings fit with seed dispersal by water in several species of *Gnetum*, morphological similarities among apparently young species, and incomplete concerted evolution in the nuclear ITS region. [Bayesian relaxed clock; biogeography; *Ephedra*; *Gnetum*; long-distance water dispersal; polytomies in molecular clock dating.]

> "It is not at all clear—although often tacitly assumed—that an old lineage has occupied its current range for a long time. *Gnetum* in particular provides neat examples of the opposite." (Markgraf, 1929:429; translated from the German by SR)

Age estimation from molecular sequences has emerged as a powerful tool for inferring the time it took a plant lineage to radiate in a particular area. Analytical methods now try to model change in substitution rates along individual branches of a phylogeny by combining molecular data with multiple simultaneous time constraints, usually from fossils (Thorne et al., 1998; Rambaut and Bromham, 1998; Kishino et al., 2001; Thorne and Kishino, 2002; Aris-Brosou and Yang, 2002; Sanderson, 2002; Drummond and Rambaut, 2005). Simulations have shown that such "relaxed" clock approaches can recover known rates (and hence ages) only as long as several calibration points are accurate (Ho et al., 2005). In practice, uncertainties in calibrations due to misidentified or misdated fossils (Doyle and Donoghue, 1993; Sanderson and Doyle, 2001; Near et al., 2005), taxon-sampling effects (Linder et al., Rutschmann, 2005), and asymmetric tree shape often reduce the reliability of the temporal constraints employed. Given the difficulty of calibrating genetic distances, it is desirable to use multiple constraints wherever possible. Using multiple fossil constraints, however, usually requires knowing phylogenetic relationships outside the study group, with the corollary of having to use relatively conservative gene regions that are alignable across phylogenetically distant taxa. These conserved gene regions will often be uninformative in the group of interest. In the face of these difficulties, assuming a strict clock for a reduced (and then more or less clock-like) data set may still be the best way to improve the precision, and perhaps accuracy, of rate and date estimates (Sanderson, 2002; Ho et al., 2005). For data sets that behave in a less clock-like fashion, the safest approach may be to compare results from strict clocks with those from relaxed clocks, employing alternative input topologies and constrained with as many critical fossils as possible.

We here deal with the problems of rate heterogeneity and unclear outgroup relationships in the seed plant lineage Gnetales, focusing on *Gnetum*, for which we have near-complete species sampling. Many seed plant lineages, especially conifers, *Ginkgo*, and cycads, have excellent fossils that potentially can serve to constrain the timing of events within *Gnetum*. *Gnetum* itself has no fossil record. Using constraints from outgroup fossils in this group, however, is problematic for at least two reasons. First, *Gnetum* is highly divergent from its closest living relatives (below) and, second, the relationships of the Gnetales to the other four lineages of seed plants, cycads, *Ginkgo*, conifers, and flowering plants, are still unresolved. In what has been referred to as the single most surprising grouping discovered with molecular data (Palmer et al., 2004), Gnetales appear to be sister to Pinaceae, and both then sister to the remaining conifers. This is the favored topology in the most

comprehensive analysis to date (Burleigh and Mathews, 2004). Other analyses place Gnetales as sister to conifers or to all other seed plants (Quandt et al., 2006). The so-called anthophyte hypothesis, which sees Gnetales as sister to flowering plants, is rarely recovered with molecular data (e.g., Chaw et al., 2000; Bowe et al., 2000; Magallón and Sanderson, 2002; Rydin et al., 2002; Schmidt and Schneider-Poetsch, 2002; Soltis et al., 2002; Rai et al., 2003; see Burleigh and Mathews, 2004, for a review).

Obtaining estimates for divergence events within *Gnetum* is of interest mainly because of the group's disjunct geographic distribution, but also because of a lateral gene transfer event in one subclade for which the time to nonfunctionalization (pseudogenization) can be estimated (Won and Renner, 2003). *Gnetum* consists of probably fewer than 33 species; most are large woody climbers, two are trees (Markgraf, 1929; HW, monograph in preparation). Ten species occur in tropical South America, one in West Africa, and the remainder in tropical and subtropical Asia. Species relationships within Asian *Gnetum* cannot be resolved with the highly conserved gene regions required for seed plant-wide analyses and dating. Instead, we relied on a six-locus data set that includes nuclear, mitochondrial, and chloroplast genes, spacers, and introns.

The accepted hypothesis has long been that the pantropical range of *Gnetum* reflects a Gondwanan history (Markgraf, 1929). As explained above, *Gnetum* has no fossils, but the Gondwanan hypothesis receives indirect support from the fossil record of *Welwitschia*, the sister lineage of *Gnetum*. *Welwitschia mirabilis*, today restricted to the Namib dessert, is known from a 110-My-old seedling from the Lower Cretaceous (Aptian) Crato Formation in northeastern Brazil (Rydin et al., 2003). With its famously weird morphology that includes a deep tap root and two strap-shaped leaves that can grow for hundreds of years (Henschel and Seely, 2000), *Welwitschia* differs remarkably from the canopy climbers that are typical of the genus *Gnetum*. *Gnetum* and *Welwitschia* together are the sister clade to *Ephedra*, which consists of 35 to 45 species of shrubby plants with photosynthetic stems and reduced leaves that are widely distributed in temperate to arid areas of Eurasia, northern Africa, southwestern North America, and western South America (Ickert-Bond and Wojciechowski, 2004; Huang et al., 2005). The monophyly of these three genera, which make up the Gnetales, is supported by numerous morphological and molecular phylogenetic analyses (Kubitzki, 1990; Price, 1996).

The Crato Formation has also yielded male strobili similar to those of *Welwitschia* and vegetative stems with joints, whorled branchlets, reduced leaves, and sessile male strobili resembling those of *Ephedra* (Mohr et al., 2004; Dilcher et al., 2005, the last reviews the *Welwitschia* fossil record). Early Cretaceous floras from Buarcos in Portugal, Virginia, USA and from the Yixian Formation in northeast China contain additional fossils that probably represent *Ephedra* (Rydin et al., 2004, 2006; Yang et al., 2005), including *Chaoyangia liangii* (Duan, 1998), which, however has winged diaspores that superficially resemble those of *Welwitschia*. Similar fossils combining *Ephedra*-like stems with winged diaspores have been described as *Gurvanella dictyoptera* from the Early Cretaceous Gurvan-Eren Formation of Mongolia (Krassilov, 1982, 1997) and as *Gurvanella exquisita*, again from the Yixian Formation (Sun et al., 2001). Further fossils associated with the *Welwitschia* lineage include Aptian Eastern North American male cones (*Drewria*; Crane and Upchurch, 1987) and a Barremian-Aptian female cone from the Lake Baikal area (*Eoantha*; Krassilov, 1986). The earliest Gnetales pollen is much older, from the Late Triassic and Early Jurassic, and could represent either *Ephedra* or *Welwitschia* (Crane, 1996; Dilcher et al., 2005); a recent study even describes Permian (270 my old) cones with such pollen (Wang, 2004). *Gnetum* pollen has a much lower fossilization potential because of its thin exine (Hesse et al., 2000).

These macro- and microfossils demonstrate the diversity of Gnetales in the Mesozoic (Late Triassic to Cretaceous) and provide an overall temporal context. However, their assignment to nodes in the tree of living species of *Ephedra*, *Welwitschia*, or *Gnetum* is problematic. Because *Welwitschia* consists of a single species, the above-mentioned fossil seedling, *Cratonia cotyledon* (belonging to the *Welwitschia* lineage), can only constrain the split between *Gnetum* and *Welwitschia*, possibly resulting in a considerable underestimate of the true age of that split. Similarly, the best preserved *Ephedra* fossils, 125-My-old seeds (Rydin et al., 2004; Yang et al., 2005), cannot confidently be assigned to the crown group of *Ephedra*, but do constrain the split between *Ephedra* and the other two genera. (We nevertheless carried out an experiment to test the effect of assigning them to the *Ephedra* crown group.)

To overcome these calibration problems, we decided to employ additional constraints from well-dated fossils of conifers, cycads, and flowering plants. This required a relaxed clock approach (i.e., modeling substitution rate change along branches) because the larger data matrix rejected the clock assumption and added the problem of the uncertain phylogenetic placement of Gnetales among seed plants (above). As a way out, we employed the current best estimate seed plant tree (Burleigh and Mathews, 2004), three alternative seed plant trees, and a topology with a basal polytomy in the gymnosperms. The latter implies a prior belief that Gnetales, conifers, cycads, and *Ginkgo* present a hard polytomy that cannot be resolved. In most current implementations of Bayesian clock approaches, prior beliefs about the topology cannot be overwritten by the data (they can in BEAST; Drummond and Rambaut, 2005), and workers have therefore generally preferred to use completely resolved topologies even for parts of the tree that are not supported by the underlying (or other) data (but see Renner and Zhang, 2004). Such a prior polytomy, however, does not prevent branch length information in the data from strongly affecting the posterior branch lengths assigned to branches above or below the polytomy. Where there is strong signal, substitutions will sort

themselves out along the branches emanating from the multifurcating node. The situation is complicated by the specific position of temporal constraints employed in a particular study. While in strict clock approaches the effect of a calibration point is identical throughout the tree, in relaxed clocks, the effect of age constraints—whether minimal, maximal, or fixed—diminishes with each node further away from the constraint(s). Given that the behavior of relaxed clocks is still so poorly understood, we felt it worthwhile to explore the empirical effects of a multifurcation in the input topology.

Having determined the time horizon for the evolution of *Gnetum* with multiple approaches, we develop a hypothesis for the radiation of the climbing and tree forms of *Gnetum* that incorporates climatic, geological, and biological evidence and set out ways in which it might be tested.

Materials and Methods

Taxon Sampling

Leaf samples with voucher information and GenBank accession numbers for the sequences obtained from them are listed in Appendices 1 and 2 (www.systematicbiology.org), which also give author names for all taxa. Most *Gnetum* samples were collected in the field by the first author; a few are from botanical gardens or herbarium material. Thirty-one accessions of *Gnetum* were sequenced for six loci, usually all from the same DNA aliquot (for exceptions see Appendix 2). The seed plant data set included representatives of all major clades of *Gnetum*, *Welwitschia mirabilis*, *Ginkgo biloba*, three species of *Ephedra* chosen to span the root of that genus (Ickert-Bond and Wojciechowski, 2004; Rydin et al., 2004; Huang et al., 2005), and representatives of conifers and cycads chosen based on the results of Gugerli et al. (2001), Quinn et al. (2002), and Rai et al. (2003). Seed plant-wide analyses were rooted with *Psilotum* because the *matK* sequence of *Marchantia* (liverwort) is much shorter than those of other vascular plants (1113 bp versus 1497–1662 bp in seed plants; it lacks the conserved domain X; see Hilu and Liang [1997]) and has a lower G+C content (18.1%; see Appendix 1), strongly suggesting pseudogenization.

Gene Sequencing

DNA extraction, cloning, and sequencing followed the methods described in Won and Renner (2005a, 2005b). Forward and reverse strands were sequenced. Figure S1 (www.systematicbiology.org) shows the approximate primer positions and organization of the six markers (cp *rbcL*, *matK*, and the $tRNA^{Leu}$ intron/spacer region; nuclear (nu) rITS/5.8S and the second intron of the *LEAFY* gene; and mt *nad1* second intron and partial exons) and Appendix 3 lists all primer sequences. To amplify *rbcL*, we used the *rbcL*-1352R primer, because the commonly used combination of *rbcL*-724F and 1460R failed to work in *Gnetum*. The *matK* regions of *G. woodsonianum* (= *G. leyboldii* var. *woodsonianum*) and *G. gnemon* Won 514 were cloned and provided by S.-M. Chaw (Institute of Botany, Academia Sinica, Taiwan). *Gnetum*-specific internal primers for *matK* were then designed, based on the sequences of *G. gnemon*, *G. montanum*, and *G. woodsonianum*. Except for these clones, the remaining *matK* sequences were obtained directly from PCR products.

Sequence Analyses and Phylogenetic Analyses

While *rbcL* sequences were constant in length (1428 base pairs [bp]), *matK* sequences showed large nucleotide composition and length variation. They were therefore translated into amino acid sequences to guide sequence alignment with Clustal X (Thompson et al., 1997). In addition, a combined six-marker data matrix (Appendix 2) was constructed by concatenating the data matrices of Won and Renner (2005a, 2005b: cp $tRNA^{Leu}$ intron/spacer; nu rITS/5.8S; nu *LEAFY* second intron; and mt *nad1* second intron/partial exons) with the *rbcL* and *matK* seed plant matrices generated here. Final alignments have been submitted to TreeBase (accession numbers S1493, M2681, M2682).

Length and G+C content of *rbcL* and *matK* were calculated from aligned sequences, and sequence divergences were obtained under the K-2-P model (Kimura, 1980). The number of substitutions at synonymous sites (*Ks*) and non-synonymous sites (*Ka*) was calculated using MEGA2 (Kumar et al., 2001). Analyses included full-length *matK* sequences and base positions 77 to 1324 of *rbcL*.

The concatenated matrix of six loci used to resolve intra-*Gnetum* relationships (Appendix 2) did not include outgroups because large length and nucleotide differences in several of the spacers and introns prevent alignment across *Gnetum*, *Welwitschia*, and *Ephedra*. (Separate studies have analyzed the phylogenetic signal contained in the secondary structure of these loci in Gnetales; Won and Renner, 2005a, 2005b.) Based on the results of the seed plant-wide analysis, within-*Gnetum* trees were rooted on the South American *Gnetum* clade.

Parsimony and maximum likelihood analyses were conduced using PAUP 4.0b10 (Swofford, 2002). Data matrices were analyzed separately under parsimony optimization and in the absence of statistically supported conflict were then combined. Parsimony analyses employed heuristic searches that used 100 random taxon-addition replicates, holding 100 trees at each step, tree bisection-reconnection (TBR) branch swapping, and the options MulTrees, Collapse, and Steepest Descent, without an upper limit for trees held in memory. Nonparametric bootstrap support was obtained by resampling the data 1000 times with the same search options and model, except using closest taxon addition.

Best-fitting substitution models for the combined *matK*/*rbcL* data sets, one comprising 13 taxa, the other 38 taxa, were found with ModelTest (Posada and Crandall, 1998), applying the Akaike information criterion. The best-fit model was the general time-reversible model

with a gamma-shape parameter and a proportion of sites modeled as invariable; i.e., the GTR+G+I model. The best-fit model for the 6-locus-31-taxa data set also was the GTR+G+I model.

Bayesian analyses of the 6-locus-31-taxa data set relied on MrBayes 3.1 (Huelsenbeck and Ronquist, 2001) and were conducted under the GTR+G+I model, employing the flat prior beliefs that are the default in MrBayes and four Markov chain Monte Carlo (MCMC) chains, one of them heated (the default in MrBayes). Markov chains were run twice for 1 million generations, sampling every 100th generation for a total of 10,000 trees. The first 1000 trees were discarded as burn-in (data points sampled before the chain reaches stationarity), and the remaining 9000 samples were combined into a single file and analyzed using the ''sumt'' command in MrBayes. Both independent runs found essentially identical tree topologies and posterior probabilities, indicating that the sample number was sufficient to permit the algorithm to converge on a global solution.

Molecular Clock Divergence Time Estimation

For age estimation, most gapped positions were excluded from the matK data set, so that the gene matrix used for dating consisted of 2756 bp, 1428 bases of *rbcL* and 1328 bp of *matK*. Likelihood-ratio tests were used to assess rate heterogeneity in the concatenated *rbcL* + *matK* data sets of 38 seed plant exemplars and of 13 exemplars (each set including *Psilotum* for rooting purposes as explained under Taxon Sampling). When the 13-taxon data set justified the assumption of equal rates between sister groups, we used it to estimate divergence times from a strict clock by choosing the "enforce clock" option in PAUP. When the 38-taxon data rejected the assumption of equal rates in sister groups, we used a Bayesian approach that attempts to model variation in the rate of substitution along a tree. The software package used was that of Thorne and Kishino (2002; available at http://statgen.ncsu.edu/thorne/). Branch lengths were estimated with PAML's baseml component (ver. 3.14; Yang, 1997) under the F84+G model (with five rate categories), that being the only model implemented in Thorne's estbranches software, which calculates a variance/covariance matrix of branch length estimates, given sequence data and a specified rooted topology that may include polytomies. The five rooted topologies we used as input are described below. Output branch lengths from estbranches became priors for MCMC searches in multidivtime that sought to find the most likely model of rate change (with rate change assumed to be log-normally distributed), given the branch lengths, topology, time constraints on nodes, and a Brownian motion parameter nu that controls the magnitude of autocorrelation per million years along the descending branches of the tree. Prior gamma distributions on parameters of the relaxed clock model were as follows: The mean of the prior distribution for the root age was set to 3 time units (i.e., 300 My, based on seed plant fossils; below); the standard deviation of this prior was also set to 3. The mean and SD of the prior distribution for ingroup root rate were set to 0.0035 substitutions/site/My by dividing the median of the distances between the ingroup root and the tips by the time unit. The prior for nu was set to 0.4, following the manual's recommendation that the time units between root and tips multiplied by nu be about 1. The standard deviation on nu was also set to 0.4. Markov chains in multidivtime were run for 1 million generations, sampling every 100th generation for a total of 10,000 trees, with a burn-in of 1000 trees before the first sampling of the Markov chain. In the course of this study, 19 runs of 1 million generations were carried out on the *matK/rbcL* 38-taxon matrix, with consistent results that reflected the experimentally changed fossil constraints.

Because of the unresolved placement of Gnetales among seed plants, we used the following alternative input topologies for the 38-taxon data set: (1) The preferred topology of Burleigh and Mathews (2004; our Fig. 1), which has Gnetales embedded in conifers as sister to Pinaceae and is referred to as the *Gnepine* hypothesis; (2) Gnetales as sister to conifers, the Gnetifer hypothesis (Fig. S1A; www.systematicbiology.org); (3) Gnetales as sister to all other seed plants, the Gnetales sister hypothesis (Fig. S1B); (4) Gnetales as sister to angiosperms, the Anthophyte hypothesis (Fig. S1C); and (5) a topology with a basal polytomy in the gymnosperms (Fig. S1D).

For absolute ages we relied on the geologic time scale of Gradstein et al. (2004) and the following calibrations (Fig. 1): The root node, i.e., the crown group of seed plants, was constrained to maximally 385 My old based on the oldest seed precursors (Gerrienne et al., 2004). This may be an overestimate because these early seeds, which have lobed integuments, probably reflecting their origin from fused sporangia or fertile telomes, are quite different from the seeds of modern seed plants. Modern seed types with completely fused integuments are not known until the late Early Carboniferous (ca. 325 My; J. Doyle, personal communication, July 2005). Crown group seed plants thus arose between 325 and 385 Mya. Node 2, the gymnosperm crown group, was constrained to 315 Mya, based on the appearance of stem relatives of living conifers in the middle Late Carboniferous (ca. 310 Mya) and the appearance of Cordaitales, which are probably still more basal stem relatives of conifers, in the earliest Late Carboniferous (ca. 318 Mya). Node 3 was constrained to minimally 125 My old based on earliest *Ephedra*-type seeds (Barrett, 2000; Rydin et al., 2004; Yang et al., 2005); the minimal age of 125 My for the appearance of *Ephedra* is further supported by *Ephedra*-like macrofossil from the Brazilian Crato Formation (Mohr et al., 2004). Node 4, the split between *Gnetum* and *Welwitschia*, was constrained to 110 My based on the *Cratonia cotyledon* fossil (Rydin et al., 2003). Node 5, crown conifers, was constrained to minimally 225 My old based on the oldest Pinaceae-type cones (Miller, 1999; however, most genera of Pinaceae first appear during the Early Tertiary; LePage, 2003a). The split of Pinaceae from other conifers is bracketed not only by these Triassic cones, but also by

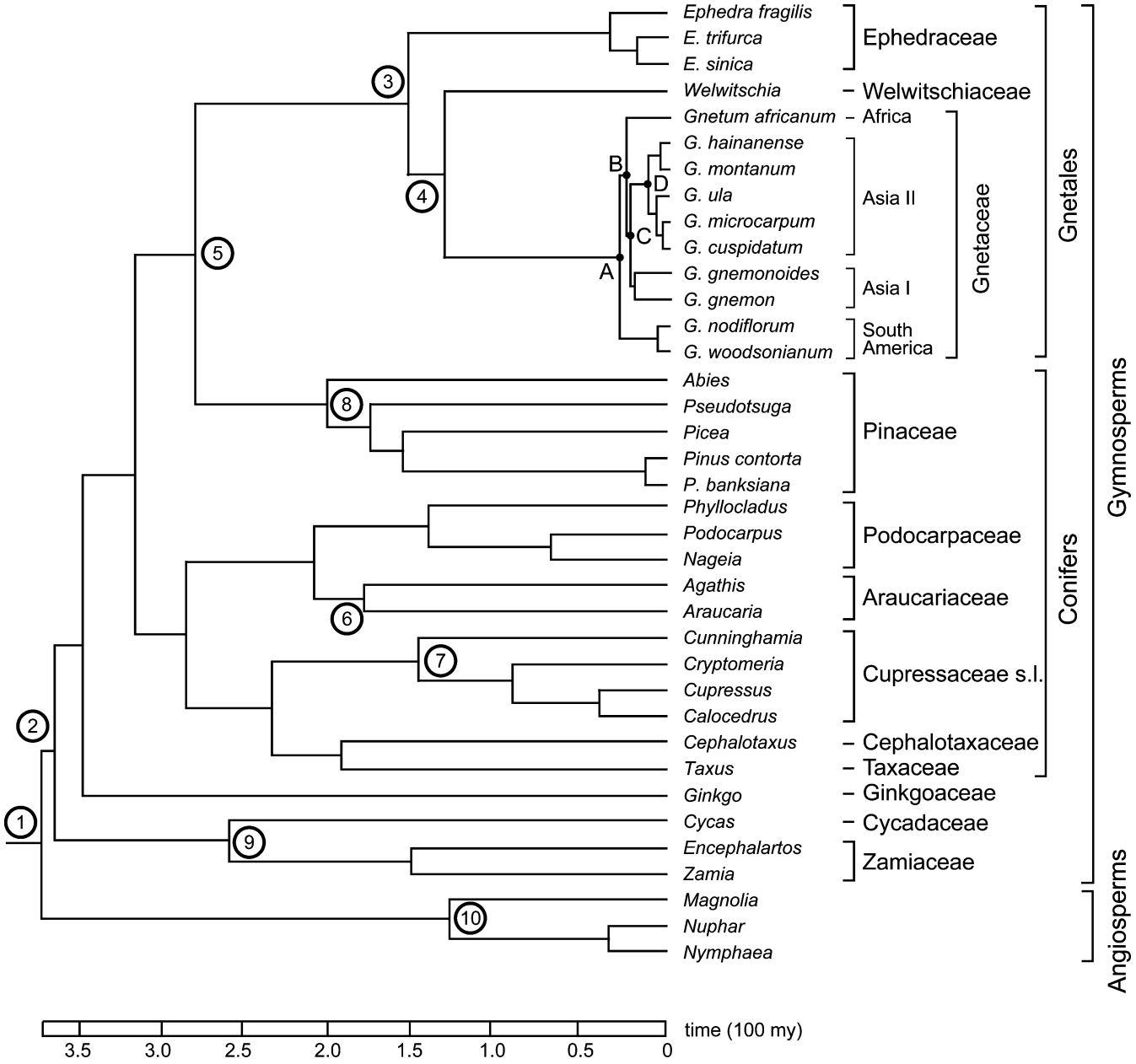

FIGURE 1. Divergence events in *Gnetum* estimated from the chloroplast genes *rbcL* and *matK* analyzed under a Bayesian relaxed clock, constrained by fossil-based minimal ages at nodes 2 to 9 (see Materials and Methods) and assuming that Gnetales are nested in the conifers (the so-called Gnepine topology). For results with four other seed plant topologies see Figures S1A–D. All analyses were rooted with *Psilotum* because the GenBank *matK* sequence of *Marchantia* appears to be a pseudogene. Nodes of particular biogeographic interest are A, the split between the South American clade of *Gnetum* and the remainder of the genus; B, the split between African and Asian *Gnetum*; C, the split between the two main Asian clades of *Gnetum*; and D, the onset of radiation of the most species-rich Asian clade.

records of Taxaceae (*Palaeotaxus*) from the Late Triassic and probable stem relatives of Podocarpaceae (*Rissikia*) from the (mid?) Triassic (J. Doyle, personal communication, July, 2005). Node 6 was constrained to minimally 160 My based on Middle Jurassic (160 to 175 My) Araucariaceae cones from Argentina (Stockey, 1982). Node 7 was constrained to minimally 90 My old based on Late Cretaceous (Turonian) Cupressaceae *Thuja*-like fossils (LePage, 2003b). The node just below could also have been constrained to the Late Cretaceous based on *Cunninghamia*-like cones (Farjon and Ortiz Garcia, 2003). Node 8 was constrained to minimally 90 My based on Turonian cones that clearly belong to crown Pinaceae (Gandolfo et al., 2001). Node 9, crown cycads, was constrained to minimally 270 My old based on *Crossozamia* from the Early Permian Lower Shihhotse Formation (Gao and Thomas, 1989), which appears to be nested among crown group cycads (Brenner et al., 2003). The

270-My age was chosen because the age of the top of the Lower Shihhotse Formation, though poorly constrained, is probably between from 275 and 265 My old (J. Hill, School of Geography, Earth and Environmental Sciences, University of Birmingham, personal communication, 9 Nov. 2005). Node 10 was constrained to minimally 125 My old, based on oldest Nymphaeaceae flowers (Friis et al., 2001). Gandolfo et al. (2004) do not accept these flowers as Nymphaeaceae, but other stem relatives or crown group members of crown angiosperms, such as Chloranthaceae, Winteraceae, monocots, and even eudicots also are known from 120- to 125-My-old fossils.

To allow for the possibility that 125-My-old *Ephedra* seed fossils may represent living species (as implied by Rydin et al., 2004, and Yang et al., 2005), we used these fossils to constrain the *Ephedra* crown group to 125 My old; i.e., we moved the constraint at node 3 in Figure 1 up to the node where *E. fragilis* diverges from *E. sinica* and *E. trifurca*.

Calibration for the strict clock (in the 13-taxon data set) came from the *Cratonia* fossil, which was used to fix the split between *Gnetum* and *Welwitschia* as 110 My. This split may have occurred considerably earlier since *Cratonia* is almost as apomorphic as *Welwitschia* (J. Doyle, personal communication, 2005), and this was one of the reasons why we resorted to calibration with fossils from more distant outgroups.

Results

Characteristics of Gnetum rbcL *and* matK *Sequences*

We generated 47 *rbcL* sequences from *Gnetum* as well as one sequence each from *Welwitschia mirabilis* and *Ephedra trifurca*. None of the *rbcL* sequences contained insertions or deletions within the exon stretch (bp 1 to 1352). The length of *rbcL* in Gnetales as in other seed plants is 1428 bp. For *matK*, we generated 12 new sequences (Appendix 1). Because a GenBank *matK* sequence identified as *G. parvifolium* (accession number AF280995) was indistinguishable from sequences of *G. gnemon*, it was excluded from further analyses. Two other *matK* sequences with single-base deletions were also excluded (Appendices 1 and 2). The length of *matK* is 1554 to 1557 bp for *Gnetum*, 1575 bp for *Welwitschia*, and 1662 bp for *Ephedra* (Appendix 1; also Huang et al., 2005). The length variation in *Gnetum matK* sequences was caused by a three-base-pair insertion between positions 10 and 12 in African *G. africanum* sequences. The G+C contents of *rbcL* sequences ranged from 41.8% to 45.2% and those of *matK* sequences from 31.1% to 33.1%, fitting well with G+C contents of other vascular plants (40.8 to 45.8% for *rbcL* and 27.1% to 37.7% for *matK*; Appendix 1).

Sequence divergences of *rbcL* and *matK* sequences within and among gnetalean genera are shown in Appendix 4 and neighbor-joining trees from these data are shown as Figure S3. Figure S4 plots the numbers of nucleotide substitution at synonymous (*Ks*) and nonsynonymous sites (*Ka*) in *rbcL* and *matK* across seed plants. Pairwise comparisons between nongnetalean vascular plants and comparisons between Gnetales and other vascular plants revealed no significant differences. Numbers of variable and parsimony-informative sites did not differ among codon positions (Appendix 5). The estimated number of nucleotide substitutions for *Gnetum rbcL* sequences was 0.058 (*Ks*) for synonymous sites and 0.031 (*Ka*) for nonsynonymous sites, on average. The averages and standard deviations of the observed *Ka*/*Ks* ratios were 0.057 ± 0.081 for *rbcL* and 0.524 ± 0.125 for *matK*. The *Ka*/*Ks* ratio in *rbcL* was thus about 1/10th that in *matK*, mainly because of much reduced substitutions at nonsynonymous sites (Fig. S4). Over all taxa, *matK* sequences showed about 1.5 times the maximum sequence divergence seen in *rbcL*, and in both genes, *Gnetum* had about six times the maximum divergence observed in *Ephedra* (in *rbcL* 0.048 versus 0.008; in *matK* 0.068 versus 0.012).

Phylogenetic Analyses and Divergence Time Estimation

Statistics for sequence variation, parsimony tree length, and consistency indices for the six data partitions are given in Appendix 6. The *rbcL* gene did not contain sufficient signal to reliably resolve relationships within *Gnetum,* but *matK* supported a sister-group relationship between Asian and African *Gnetum* and subdivision of Asian *Gnetum* into two clades (Fig. S5A). Bayesian analysis of the combined *rbcL* and *matK* data yielded a well-resolved and supported topology (Fig. S5B) in which Gnetales were sister to Pinaceae (with .86 posterior probability [PP]). However, *matK* alone placed Gnetales as sister to conifers (.96 PP), regardless of whether all positions, only 1st and 2nd positions, or only 3rd positions were included (trees not shown). It also weakly supported the monophyly of conifers (.80 PP).

A maximum likelihood tree obtained from the 6-locus-31-taxon data set shows well-resolved relationships within *Gnetum* (Fig. 2). Likelihood-ratio tests for this data set strongly rejected the clock assumption, and it was not used for dating purposes. The clock assumption was also rejected by just the *rbcL* and *matK* sequences (individually or combined) as long as data matrices included other seed plants besides Gnetales ($P \ll 0.001$). However, a *matK*/*rbcL* data set that included only 12 species of Gnetales plus *Psilotum* satisfied the clock ($P > 0.05$) and was analyzed under the assumption of a strict clock using the GTR+G+I model. Resulting genetic distances were calibrated by fixing the split between *Gnetum* and *Welwitschia* at 110 My based on the *Cratonia* fossil (see Materials and Methods).

Because all larger *rbcL/matK* matrices rejected a strict clock, we applied a Bayesian relaxed clock, constrained with the fossil-based minimal ages specified in Materials and Methods and shown in Figure 1; the root node was constrained by a maximal age (Materials and Methods). Figure S6 shows a GTR+G+I ML phylogram from these data to provide a sense of branch length heterogeneity.

Table 1 lists the divergence times obtained for key nodes within *Gnetum* (labeled A, B, C, and D in all figures) under the Gnepine, Gnetifer, Gnetales-sister, Anthophyte, and Polytomy input topologies shown in Figures 1 and S1A–D. The different hierarchical relationships of

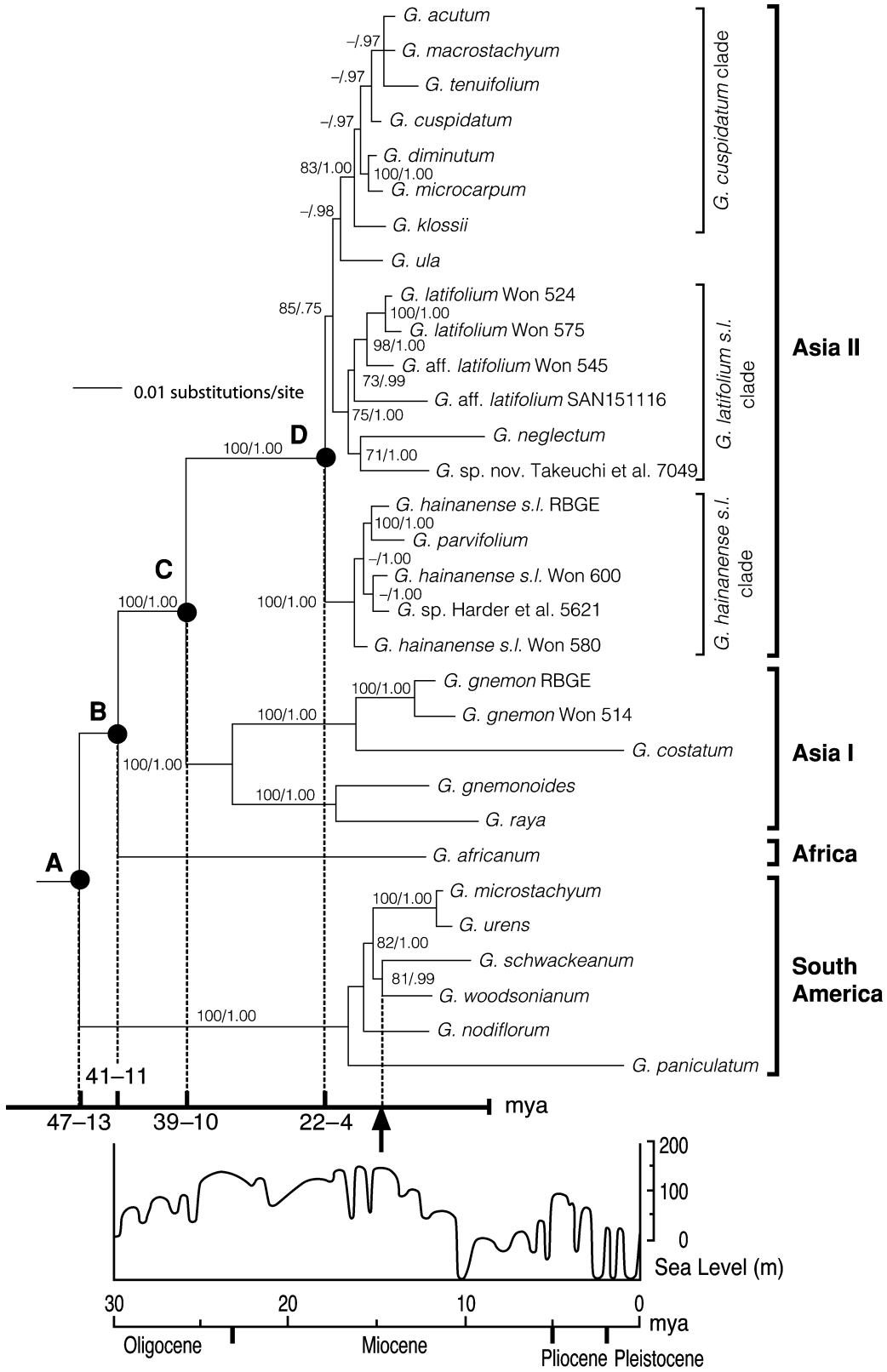

FIGURE 2. Maximum likelihood tree for *Gnetum* obtained from combined nuclear, chloroplast, and mitochondrial data, analyzed under the GTR+G+I model. Numbers above branches indicate bootstrap supports from parsimony analyses, followed by Bayesian posterior probabilities. Dashes indicate support values lower than 50%. The ages shown for nodes A to D (with 95% confidence intervals) are from the Bayesian relaxed clock analysis of *matK* and *rbcL* sequences (Fig. 1; see Table 1 for point age estimates and strict clock estimates). Geological periods and eustatic sea level changes are shown in the inset below. The arrow indicates the separation between and *G. schwackeanum* and *G. woodsonianum* (= *G. leyboldii* var. *woodsonianum*). The depiction of sea level changes is modified from Haq et al. (1987) and illustrates the dramatic fluctuations from the Miocene to the Pleistocene that correlate with most of the speciation in the Asia II clade.

TABLE 1. Time estimates (in million years, followed by 95% confidence intervals) for nodes A, B, C, D in Figures 1 and 2, and in Figures S1A–D, obtained from combined *matK* and *rbcL* sequences under a strict clock (row 1), calibrated with one fossil (Fig. 1, node 4), or a Bayesian relaxed clock (rows 2 to 7), constrained with 9 fossil-based minimal ages (Fig. 1 and Materials and Methods). Row 7 shows results obtained when 125-My-old *Ephedra* seeds were used as a minimal constraint for crown group *Ephedra*. Alternative input topologies are shown in Figure 1 and Figures S1A–D.

Model	Node A (split: South American vs. remaining *Gnetum*)	Node B (split: African vs. Asian *Gnetum*)	Node C (split: Asia I and II clades of *Gnetum*)	Node D (split: basal divergence within Asia II clade of *Gnetum*)
Strict clock, 13-taxon *matK/rbcL* data set	14	12-11	11	3
Relaxed clock, 38-taxon *matK/rbcL* data				
Constrained as in Figure 1 under the Gnepine topology	26 (13, 47)	22 (11, 41)	21 (10, 39)	11 (4, 22)
Constrained as in Figure S1A under the Gnetifer topology	29 (14, 51)	25 (11, 44)	24 (11, 43)	12 (5, 25)
Constrained as in Figure S1B under the Gnetales-sister topology	38 (19, 66)	33 (15, 58)	31 (15, 56)	17 (7, 34)
Constrained as in Figure S1C under the Anthophyte topology	37 (18, 64)	32 (15, 58)	30 (14, 55)	16 (6, 33)
Constrained as in Figure S1D with basal 4-tomy in gymnosperms	30 (14, 53)	26 (12, 47)	24 (11, 45)	13 (5, 26)
Constrained as in Figure 1 under the Gnepine topology, but *Ephedra* crown min. age set to 125 My	44 (23, 71)	38 (19, 64)	36 (18, 62)	20 (8, 39)

seed plants, in particular whether conifers are monophyletic or paraphyletic and whether Gnetales are sister to all other seed plants or embedded in conifers, imply changes in the placements of two minimal constraints, namely the oldest gymnosperm (node 2) and the oldest conifer (node 5). This changed within-*Gnetum* time estimates by more or less 10 My (Table 1). For example, under the Gnepine seed plant topology, *Gnetum* first diverged at 26 My, under the Gnetifer topology it diverged at 29 My, under the Gnetales-sister hypothesis at 38 My, and under the Anthophyte hypothesis at 37 My. Estimates under the Polytomy input tree were closest to the estimates under the Gnetifer topology (Table 1). The experimental reassignment of the 125-My-old *Ephedra* seeds to the *Ephedra* crown instead of the stem (node 3 in Fig. 1), resulted in almost doubled within-*Gnetum* ages (Table 1).

Estimates for major divergence events in *Gnetum* from the 13-taxon strict clock analysis were younger by about half compared to those from the 38-taxon relaxed clock analysis (Table 1). For example, the time estimated for the divergence of South American *Gnetum* from the remainder of the genus under a relaxed clock and the Gnepine topology (Fig. 1) was 26 My (13 to 47 my, 95% confidence interval), whereas under a strict clock, the same event was dated to 14 My.

DISCUSSION

Effects of Tree Topologies, Constraints from Distant Outgroups, and Strict versus Relaxed Clock Inference

A polytomy in the input tree for a Bayesian analysis implies a prior belief that speciation events occurred at about the same time so that it cannot be estimated from substitutions in what order they occurred. This is hardly a well-founded hypothesis for the evolution of the four lineages of gymnospermous seed plants that make up the particular polytomy relevant to the present study. Empirically, however, this unconvincing prior hypothesis had little impact on the posterior distribution of within-*Gnetum* estimates, which under the polytomy input tree closely resembled estimates obtained under the Gnetifer input tree. The likely reason is that polytomies in dating trees might impact posteriors branch lengths only where there is branch length uncertainty throughout the tree (J. Thorne, personal communication, Oct. 2005), whereas there is strong branch length information in many parts of the seed plant tree, as evident in a phylogram of the underlying data (Fig. S6). Another caveat concerning current Bayesian relaxed clock implementations is that the impact of the other priors used, namely the root rate and the Brownian motion parameter (Materials and Methods), is poorly understood.

What is clear from our results is that the constraint(s) closest to the nodes being estimated exert(s) the strongest impact (also Wiegmann et al., 2003) and that clear signal in molecular branch lengths can push nodes far from their own minimal constraints. For example, node 7, crown Cupressaceae, and node 8, crown Pinaceae, are both constrained by fossils to minimally 90 My old, but inspection of Figures 1 and S1A–D shows that based on genetic distances and irrespective of the seed plant topology employed, crown Pinaceae are considerably older than crown Cupressaceae.

In terms of their likelihood scores, the unconstrained input topologies rank as follows: Gnepine seed plant topology ($-\ln = 27{,}964.2703$), Gnetifer topology ($-\ln = 27{,}964.9216$), Gnetales sister topology ($-\ln = 27{,}971.7141$), and, worst, the Anthophyte topology ($-\ln = 27{,}976.5747$). The most comprehensive seed plant analysis to date (Burleigh and Mathews, 2004) found that combined data from 13 loci (for 31 exemplar taxa) favored the Gnetales sister topology, whereas slower evolving data partitions favored the Gnepine tree. In the following section we concentrate on the estimates obtained under the Gnepine topology (Table 1, row 2). However, the within-*Gnetum* ages obtained under the alternative seed plant topologies, including the polytomy tree, differ by only 4 to 12 My, and our inferences concerning *Gnetum* therefore would change little if one of the other seed

plant trees turned out to be true. Our conclusions further incorporate strict clock estimates, which do not depend on seed plants outside Gnetales. The strict clock (for 13 Gnetales) yielded an estimate of 14 My for the onset of diversification within *Gnetum*, about half that from the 38-taxon relaxed clock (26 My; 95% confidence range 13–47 My). This is probably due to the well-known taxon density effect in molecular dating (e.g., Linder et al., 2005); more densely sampled clades generally yield older ages. An earlier strict clock estimate for the onset of diversification of extant *Gnetum* was 11 to 6 My (Won and Renner, 2003); the difference between our earlier analysis and the present is that we here included two additional species of *Ephedra* and deleted a stretch of about 400 bp from the *matK* data because of gapped *Ephedra* sequences.

Biogeography and Evolution of Gnetum

The major divergence events among extant clades of *Gnetum* are estimated as dating to the Upper Oligocene, Miocene, and Pliocene (Table 1). Under the strict clock, the main South American, African, and Asian divergence events occurred sometime during the Miocene and Pliocene, whereas the Bayesian relaxed clock implies divergence during the Upper Oligocene and Miocene. These estimates in no way contradict the Mesozoic fossil record of Gnetales (Introduction), which includes over 100-My-old relatives of extant species of *Ephedra* and *Welwitschia* (e.g., Mohr and Friis, 2000; Yang, 2002; Rydin et al., 2003, 2004, 2006; Mohr et al., 2004; Wang, 2004; Yang et al., 2005; Dilcher et al., 2005). Rather, the molecular estimates concern the recent-most common ancestor of living clades, while the fossils may represent members of stem lineages or extinct sister species. Gnetales were diverse and widespread during the Cretaceous, and the almost worldwide occurrence of their fossils compared to their present range indicates that they have suffered major extinctions. A new insight from this study is that *Gnetum* has undergone a geologically recent radiation, especially in the Malesian region and that the genus's present disjunct range is not Gondwanan. It is possible that the Asian radiation coincided with times of low sea levels because of the opportunities for overland seed dispersal they would have afforded (Fig. 2). Temporally, although not geographically, our findings for *Gnetum* parallel those for *Ephedra* (Huang and Price, 2003; Ickert-Bond and Wojciechowski, 2004). In the case of *Ephedra*, the evolution of Mediterranean climates, the rise of the Rockies and the Andes, and seed dispersal by birds and rodents (e.g., Ridley, 1930; Holmgren et al., 2003) all seem to have contributed to relatively recent radiations (as well as contributing to species formation per se).

Biogeographic analyses of other plant families with ancient and relatively good fossil records, such as Calycanthaceae, Chloranthaceae, *Nothofagus*, and Nymphaeaceae, have obtained similar results, namely that in spite of their impressive fossil records, these clades apparently reached some parts of their range quite recently (Zhang and Renner, 2003: Chloranthaceae; Knapp et al., 2005: *Nothofagus*; Zhou et al., 2006: Calycanthaceae; Yoo et al., 2006: Nymphaeaceae). In other words, old lineages do not necessarily stop dispersing and diversifying, as pointed out by Markgraf (1929), the last monographer of *Gnetum*, in the quote at the beginning of this paper. Similarly, *Ephedra* seems to have diversified in its current habitat during Miocene times (Huang and Price, 2003; Ickert-Bond and Wojciechowski, 2004), in spite of morphological stasis in some morphological traits, such as the seed coat in fossil *E. archaeorhytidosperma* and living *E. rhytidosperma* (Yang et al., 2005). Such morphological stasis does not imply that the respective fossil represents a member of continuous ancestor-descendent chains of populations without intervening speciation events. When we nevertheless used the 125-My-old *Ephedra* seed fossils to constrain the minimal age of living species of *Ephedra* (the experiment reported in Table 1, row 7), divergences between South American, African, and Asian *Gnetum* were still estimated as dating to the Eocene and Oligocene, rather than being Gondwanan.

Gnetum currently comprises ten species in South America, one in West Africa, and ~25 in tropical and subtropical Asia. If an Oligocene/Miocene age is accepted for the extant lineages in South America and Asia (Fig. 2), then *Gnetum* must have reached its disjunct pantropical range either through transoceanic dispersal or through a Laurasian expansion followed by southwards spread. If one considers the upper error bracket of ± 50 My (under any of the different seed plant topologies; Table 1), rather than the point estimates, *Gnetum* would be old enough to have dispersed across the North Atlantic land bridge to Africa and Asia. An Eocene boreotropical range, with subsequent fragmentation during the Oligocene climate cooling, is difficult to reject with available data, but does not fit well with the occurrence of it sister group, *Welwitschia*, in Africa and Brazil (living and as fossils, respectively).

In terms of dispersal biology, water dispersal in seawater is a strong possibility. Some South American species, for example, *G. venosum*, have a special middle layer in the seed coat that gives buoyancy (Kubitzki, 1985), and others, such as *G. gnemonoides*, have large, corky diaspores (Markgraf, 1951). The smallest diaspores in the genus, those of *G. africanum*, measure 1.2−1.5 cm × 0.8 cm; the largest *Gnetum* seeds measure 7 cm × 3–4 cm. In most species, the mature seed envelope turns red or yellow, and seeds are then attractive to large birds, such as toucans, but also to rodents and monkeys (e.g., Ridley, 1930; Markgraf, 1951; Kubitzki, 1985; Van Roosmalen, 1985; Forget at al., 2002). Fish eat the seeds when they fall into streams, and Amazon catfish are known to regurgitate them because of sclerenchyma needles in the seed envelope (Goulding, 1980; Kubitzki, 1985). Fish dispersal therefore does not destroy the embryo. *Gnetum* species commonly grow along rivers, and there are also observations of *Gnetum* seeds from sea drift; for example, from beaches in Malaysia (Ridley, 1930; Hemsley in Markgraf, 1951). Several months are required for embryo maturation (Maheshwari and Vasil, 1961).

A review of molecular clock-dated instances of diaspore dispersal and how they relate to the direction of surface currents in the Atlantic (Renner, 2004) showed that there are several cases of likely water dispersal between South America and Africa. Dispersal between Africa, Madagascar, and India/Malaysia appears even more common (Thorne, 1973; additional references in Renner, 2005) and is more readily explained than that between South America and Africa because distances are shorter and there are regular monsoon wind systems as well as sea currents favorable for the transport of small objects between Africa, the Seychelles, the Comores, the Chagos archipelago (about half way between Africa and Indonesia), and India. Recently discovered examples of transmarine dispersal between Madagascar, Africa, and the Seychelles include multiple lineages of frogs, chameleons, rodents, Carnivora, and lemurs (DeQueiroz, 2004; Renner, 2004).

With the likely time horizon for the evolution of *Gnetum* being the Upper Oligocene, Miocene, and Pliocene (Table 1), what were the main forces behind the radiation of the genus in South American and Southeast Asia? Entry into new kinds of habits appears to have played a limited role; the climbing species of *Gnetum* as well as the two tree species occur in the understory of upland forests and in riparian vegetation. The tree habit appears to have evolved once, in the ancestor of *G. costatum* and *G. gnemon*, rather than being an ancestral state as previously assumed (Markgraf, 1929; Fig. 2). Besides growth form, there are few other morphological differences that would readily suggest different habitat niches; instead, species differences in *Gnetum* concern details of strobilus branching, and seed and leaf size.

Apparently, geographic isolation played the main role in species formation in *Gnetum*. Thus, in South America, the uplift of the northern Andes led to at least one species pair: *Gnetum woodsonianum* (= *G. leyboldii* var. *woodsonianum*) occurs from Costa Rica to Colombia and is sister to *G. schwackeanum* from Venezuela and the Guianas to northern Brazil (Stevenson and Zanoni, 1991; Fig. 2). The Colombian Andes were formed by rapid uplift of the Eastern Cordillera between 5 and 2 Mya (Gregory-Wodzicki, 2000), and Central and South America were not connected until about 3 Mya (Haug and Tiedemann, 1998). The ancestor of *G. woodsonianum* likely became separated by the uplift of the Colombian Andes and then expanded its distribution across the Panamanian land bridge into Costa Rica. The remaining eight South American species are distributed east of the Andes, mainly in the Amazonian lowland below 600 m altitude; only *G. camporum* occurs in the Guiana highland at up to 1800 m altitude.

The Asian subclades of *Gnetum* are between 4 and 22 My old (Table 1, Fig. 2) and result from multiple cycles of colonization, judging from the overlapping distributions of the largest subclades, *G. hainanense* s.l., *G. latifolium* s.l., and *G. cuspidatum*. All three subclades range from the Malay Peninsula to New Guinea across Borneo, Sulawesi, and the Philippines. Each subclade contains widely distributed and endemic species. Such pairs (Fig. 2) are *G. gnemon* with a wide distribution and its sister species *G. costatum*, endemic in New Guinea; *G. gnemonoides* with a wide distribution and its sister *G. raya*, endemic to Borneo; and *G. latifolium* s.l. with a wide distribution and its sister clade *G. neglectum*, endemic to Borneo, and *G*. sp. nov. Takeuchi et al. 7049, endemic in New Guinea. The *Gnetum cuspidatum* clade includes further pairs of wide-ranging species and narrow endemics (Fig. 2: *G. klossii* endemic to Borneo; *G. tenuifolium* endemic to the Malay Peninsula; *G. microcarpum, G. diminutum*, and *G. acutum* endemic to the Malay Peninsula and Borneo, *G. macrostachyum* ranging all the way from southern Indochina through the Malay Peninsula to Sumatra). These overlaid patterns of wide-ranging species related to local endemics suggest waves of expansion, local isolation or extinction, and reexpansion. Fluctuation in sea level (Haq et al., 1987; our Fig. 2) may have facilitated the dispersal of *Gnetum* over to the Malesian islands, while at other times acting as barriers to dispersal. To date, only one case of introgression has been detected among species of *Gnetum*: analyses of nuclear ribosomal internal transcribed spacer (ITS) sequences and of the second intron of the nuclear *LEAFY* gene suggest that *G. klossii*, a member of the *G. cuspidatum* clade in terms of its ITS sequences (as well as its chloroplast sequences), falls in the *G. latifolium* s.l. clade in terms of its *LEAFY* second intron sequences (Won and Renner, 2005a).

Dense sampling of population-level genetic variation in several of the suspected young species pairs would allow testing our scenario of recent waves of expansion and extinction in Indonesian and Malesian clades of *Gnetum*. A complicating aspect deserving further study in this context is the impact humans may have had on the distribution of *Gnetum*; *G. gnemon* is commonly cultivated for food in Indonesia and New Guinea, where the migrations of human populations have greatly influenced the distribution patterns of important tropical crop plants (Barrau, 1963). Currently, *G. gnemon* has the widest distribution of all Asian *Gnetum*, even having reached Fiji, either via over water dispersal or with the help of *Homo sapiens sapiens*.

ACKNOWLEDGMENTS

We thank Jim Doyle, an anonymous reviewer, and the associate editor, Peter Linder, for constructive criticism; Y.-C. Chan, N. V. Dzu, H. Gang, A. Hussein, Z. bin Mat Said, J. F. Maxwell, M. Postar, R. Wang, and K.-M. Wong for help with fieldwork; L. Gillespie, D. K. Harder, M. F. Prévost, S. Shi, P. Suksathan, W. Takeuchi, and the botanical gardens of Bogor, Mainz, New York, Edinburgh, and the University of California–Berkeley for plant material; S.-M. Chaw for *Gnetum matK* cloning; S. Ickert-Bond, C. Rydin, and J. Huang for *Ephedra* aliquots and pre-publication sequences; B. Mohr and Y. Yang for information about *Welwitschia* and *Ephedra* fossils; and M. Krings, J. Hill, and E. Hermsen for discussion of seed plant fossils. This paper represents part of a dissertation submitted by the first author in partial fulfillment of the requirements for a Ph.D. from University of Missouri–St. Louis. The American Society of Plant Taxonomy, the International Association of Plant Taxonomy, and the International Center for Tropical Ecology at the University of Missouri–St. Louis supported our work financially.

REFERENCES

Aris-Brosou, S., and Z. Yang. 2002. Effects of models of rate evolution on estimation of divergence dates with special reference to the metazoan 18S rRNA phylogeny. Syst. Biol. 51:703–714.

Barrau, J. 1963. Plants and the migration of Pacific peoples. Bishop Museum Press, Honolulu, Hawaii.

Barrett, P. M. 2000. Evolutionary consequences of dating the Yixian Formation. Trends Ecol. Evol. 15:99–103.

Bowe, L. M., G. Coat, and C. W. dePamphilis. 2000. Phylogeny of seed plants based on all three genomic compartments: extant gymnosperms are monophyletic and Gnetales' closest relatives are conifers. Proc. Natl. Acad. Sci. USA. 97:4092–4097.

Brenner, E. D., D. W. Stevenson, and R. W. Twigg. 2003. Cycads: evolutionary innovations and the role of plant-derived neurotoxins. Trends Plant Sci. 8:446–452.

Brunsfeld, S. J., P. S. Soltis, D. E. Soltis, P. A. Gadek, C. J. Quinn, D. D. Strenge, and T. A. Ranker. 1994. Phylogenetic relationships among the genera of Taxodiaceae and Cupressaceae: evidence from *rbcL* sequences. Syst. Bot. 19:253–262.

Burleigh, J. G., and S. Mathews. 2004. Phylogenetic signal in nucleotide data from seed plants: implications for resolving the seed plant tree of life. Am J. Bot. 91:1599–1613.

Chaw, S.-M., C. L. Parkinson, Y. Cheng, T. M. Vincent, and J. D. Palmer. 2000. Seed plant phylogeny inferred from all three plant genomes: monophyly of extant gymnosperms and origin of Gnetales from conifers. Proc. Natl. Acad. Sci. USA. 97:4086–4091.

Cheng, Y., R. G. Nicolson, K. Tripp, and S.-M. Chaw. 2000. Phylogeny of Taxaceae and Cephalotaxaceae genera inferred from chloroplast *matK* gene and nuclear rDNA ITS region. Mol. Phylogenet. Evol. 14:353–365.

Conran, J. G., G. A. Wood, P. G. Martin, J. M. Dowd, C. J. Quinn, P. A. Gadek, and R. A. Price. 2000. Generic relationships within and between the gymnosperm families Podocarpaceae and Phyllocladaceae based on an analysis of the chloroplast gene *rbcL*. Aust. Syst. Bot. 48:715–724.

Crane, P. R. 1996. The fossil history of the Gnetales. Int. J. Plant Sci. 157 (6 Suppl.):S50–S57.

Crane, P. R., and G. R. Upchurch. 1987. *Drewria potomacensis* gen. et sp. nov., an early Cretaceous member of Gnetales from the Potomac group of Virginia. Am. J. Bot. 74:1722–1776.

De Queiroz, A. 2004. The resurrection of oceanic dispersal in historical biogeography. Trends Ecol. Evol. 20:68–73.

Dilcher, D. L., M. E. Bernardes-de-Oliveira, D. Pons, and T. A. Lott. 2005. Welwitschiaceae from the Lower Cretaceous of Northeastern Brazil. Am. J. Bot. 92:1294–1310.

Doyle, J. A., and M. J. Donoghue. 1993. Phylogenies and angiosperm diversification. Paleobiol. 19:141–167.

Drummond, A. J., and A. Rambaut. 2005. BEAST v1.3, Available from http://evolve.zoo.ox.ac.uk/beast/.

Duan, S.-Y. 1998. The oldest angiosperm—a tricarpous female reproductive fossil from western Liaoning Province, NE China. Sci. China Ser. D 41:14–20.

Farjon, A., and S. Ortiz Garcia.2003. Cone and ovule development in *Cunninghamia* and *Taiwania* (Cupressaceae sensu lato) and its significance for conifer evolution. Am. J. Bot. 90:8–16.

Forget, P.-M., D. S. Hammond, T. Milleron, and R. Thomas. 2002. Seasonality of fruiting and food hoarding by rodents in neotropical forests: consequences for seed dispersal and seedling recruitment. Pages 241–256 *in* Seed dispersal and frugivory: Ecology, evolution and conservation (D. J. Levey, W. R. Silva, and M. Galetti, eds.). CABI Publishing, U.K.

Friis, E. M., K. R. Pedersen, and P. R. Crane. 2001. Fossil evidence of water lilies (Nymphaeales) in the Early Cretaceous. Nature 410:357–360.

Gadek, P. A., D. L. Alpers, M. M. Heslewood, and C. J. Quinn. 2000. Relationships within Cupressaceae sensu lato: a combined morphological and molecular approach. Am. J. Bot. 87:1044–1057.

Gandolfo, M. A., K. C. Nixon, and W. L. Crepet. 2001. Turonian Pinaceae of New Jersey. Plant Syst. Evol. 226:187–203.

Gandolfo, M. A., K. C. Nixon, and W. L. Crepet. 2004. Cretaceous flowers of Nymphaeaceae and implications for complex insect entrapment pollination mechanisms in early angiosperms. Proc. Natl. Acad. Sci. USA. 101:8056–8060.

Gao, Z., and B. A. Thomas. 1989. A review of cycad megasporophylls, with new evidence of *Crossozamia* Pomel and its associated leaves from the Lower Permian of Taiyuan, China. Rev. Palaeobot. Palynol. 60:205–233.

Germano, J., and A. S. Klein. 1999. Species-specific nuclear and chloroplast single nucleotide polymorphisms to distinguish *Picea glauca*, *P. mariana* and *P. rubens*. Theor. Appl. Genet. 99:37–49.

Gerrienne, P., B. Meyer-Berthaud, M. Fairon-Demaret,M. Streel, and P. Steemans. 2004. *Runcaria*, a Middle Devonian seed plant precursor. Science 306:856–858

Goulding, M. 1980. The fishes and the forest. Berkeley, University of California Press.

Gradstein, F. M., J. G. Ogg, and A. G. Smith (eds.) 2004. A geologic time scale 2004. Cambridge University Press, Cambridge, UK.

Gregory-Wodzicki, K. M. 2000. Uplift history of Central and Northern Andes: a review. GSA Bulletin 112:1091–1105.

Gugerli, F., C. Sperisen, U. Büchler, I. Brunner, S. Brodbeck, J. D. Palmer, and Y.-L. Qiu. 2001. The evolutionary split of Pinaceae from other conifers: evidence from an intron loss and a multigene phylogeny. Mol. Phylogenet. Evol. 21:167–175.

Haq, B. U., J. Hardenbol, and P. R. Vail. 1987. Chronology of fluctuating sea levels since the Triassic. Science 235:1156–1167.

Hasebe, M., R. Kofuji, M. Ito, M. Kato, K. Iwatsuki, and K. Ueda. 1992. Phylogeny of gymnosperms inferred from *rbcL* gene sequences. Bot. Mag. Tokyo 105:673–679.

Hasebe, M., P. G. Wolf, K. M. Pryer, K. Ueda, M. Ito, R. Sano, G. J. Gastony, J. Yokoyama, J. R. Manhart, N. Murakami, E. H. Crane, C. H. Haufler, and W. D. Hauk. 1995. A global analysis of fern phylogeny based on *rbcL* nucleotide sequences. Am. Fern J. 35:134–181.

Haug, G. H., and R. Tiedemann. 1998. Effect of the formation of the Isthmus of Panama on Atlantic Ocean thermohaline circulation. Nature 393:673–676.

Henschel, J. R., and M. K. Seely. 2000. Long-term growth patterns of *Welwitschia mirabilis*, a long-lived plant of the Namib Desert (including a bibliography). Plant Ecology 150:7–26.

Hesse, M., M. Weber, and H. Halbritter. 2000. A comparative study of the polyplicate pollen types in Arales, Laurales, Zingiberales, and Gnetales. Pages 227-239 *in* Pollen and spores: morphology and biology (M. M. Harley, C. M. Morton, and S. Blackmore, eds.). Royal Botanic Gardens, Kew, UK.

Hilu, K. W., T. Borsch, K. Müller, D. E. Soltis, P. S. Soltis, V. Savolainen, M. W. Chase, M. P. Powell, L. A. Alice, R. Evans, H. Sauquet, C. Neinhuis, T. A. B. Slotta, J. G. Rohwer, T. A. B. Slotta, J. G. Rohwer, C. S. Campbell, and L. W. Chatrou. 2003. Angiosperm phylogeny based on *matK* sequence information. Am. J. Bot. 90:1758–1776.

Hilu, K. W., and H. Liang. 1997. The *matK* gene: sequence variation and application in plant systematics. Am. J. Bot. 84:830-839.

Hipkins, V. D., C. J. Tsai, and S. H. Strauss. 1990. Sequence of the gene for the large subunit of ribulose 1,5-bisphosphate carboxylase from a gymnosperm, Douglas fir. Plant Mol. Biol. 15:505–507.

Ho, S. Y. W., M. J. Phillips, A. J. Drummond, and A. Cooper. 2005. Accuracy of rate estimation using relaxed-clock models with a critical focus on the early metazoan radiation. Mol. Biol. Evol. 22:1355–1363.

Holmgren, C. A., M. C. Peñalba, K. A. Rylander, and J. L. Betancourt. 2003. A 16,000 14C yr B.P. packrat midden series from the USA–Mexico borderlands. Quatern. Res. 60:319–329.

Huang, J., D. E. Giannasi, and R. A. Price. 2005. Phylogenetic relationships in *Ephedra* (Ephedraceae) inferred from chloroplast and nuclear DNA sequences. Mol. Phylogenet. Evol. 35:48–59.

Huang, J., and R. A. Price. 2003. Estimation of the age of extant *Ephedra* using chloroplast *rbcL* sequence data. Mol. Biol. Evol. 20:435–440.

Huelsenbeck, J. P., and F. R. Ronquist. 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Ickert-Bond, S. M., and M. F. Wojciechowski. 2004. Phylogenetic relationships in *Ephedra* (Gnetales): Evidence from nuclear and chloroplast DNA sequence data. Syst. Bot. 29:834–849.

Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16:111–120.

Kishino, H., Thorne, J. L., and W. J. Bruno. 2001. Performance of a divergence time estimation method under a probabilistic model of rate evolution. Mol. Biol. Evol. 18:352–361.

Knapp, M., K. Stöckler, D. Havell, F. Delsuc, F. Sebastiani, P. J. Lockhart. 2005. Relaxed molecular clock provides evidence for long-distance dispersal of *Nothofagus* (Southern Beech). PLoS Bio. 3: 38–43.

Korall, P., and P. Kenrick. 2002. Phylogenetic relationships in Selaginellaceae based on *rbcL* sequences. Am. J. Bot. 89:506–517.

Krassilov, V. A. 1982. Early Cretaceous flora of Mongolia. Palaeontographica Abt. B. 181:1–43.

Krassilov, V. A. 1986. New floral structure from the Lower Cretaceous of Lake Baikal area. Rev. Palaeobot. Palynol. 47:9–16.

Krassilov, V. A. 1997. Pages 50–52 *in* Angiosperm origins: Morphological and ecological aspects. PENSOFT Publishers, Sofia, Bulgaria.

Kubitzki, K. 1985. Ichthyochory in *Gnetum venosum*. An. Acad. Barsil. Ciênc. 57:513–516.

Kubitzki, K. 1990. Gnetatae. Pages 378-391 *in* The families and genera of vascular plants I: Pteridophytes and Gymnosperms (K. U. Kramer and P. S. Green, eds.). Springer-Verlag, Berlin, Heidelberg.

Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei. 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17:1244–1245. (http://www.megasoftware.net)

LePage, B. A. 2003a. A new species of *Thuja* (Cupressaceae) from the Late Cretaceous of Alaska: implications of being evergreen in a polar environment. Am. J. Bot. 90:167–174.

LePage, B. A. 2003b. The evolution, biogeography and palaeoecology of the Pinaceae based on fossil and extant representatives. Acta Hort. 615 (IV Int. Conifer Conf.):29–52.

Les, D. H., D. K. Garvin, and C. F. Wimpee. 1991. Molecular evolutionary history of ancient aquatic angiosperms. Proc. Natl. Acad. Sci. USA. 88:10119–10123.

Les, D. H., E. L. Schneider, D. J. Padgett, P. S. Soltis, D. E. Soltis and M. Zanis. 1999. Phylogeny, classification and floral evolution of water lilies (Nymphaeales): a synthesis of non-molecular, *rbcL, matK* and 18S rDNA data. Syst. Bot. 24:28–46.

Lidholm, J., and P. Gustafsson. 1991. A three-step model for the rearrangement of the chloroplast *trnK-psbA* region of the gymnosperm *Pinus contorta*. Nucleic Acids Res. 19:2881–2887.

Linder, H. P., C. R. Hardy, and F. Rutschmann. 2005. Taxon sampling effects in molecular clock dating: an example from the African Restionaceae. Mol. Phyl. Evol. 35:569–582.

Magallón, S., and M. J. Sanderson. 2002. Relationships among seed plants inferred from highly conserved genes: sorting conflicting phylogenetic signals among ancient lineages. Am. J. Bot. 89:1991–2006.

Maheshwari, P., and V. Vasil. 1961. *Gnetum*. Council of Scientific & Industrial Research, New Delhi, India.

Manhart, J. R. 1994. Phylogenetic analysis of green plant *rbcL* sequences. Mol. Phylogenet. Evol. 3:114–127.

Markgraf, F. 1929. Monographie der Gattung *Gnetum*. Bull. Jard. Bot. Buitenzorg Ser. 3, 10:407–511.

Markgraf, F. 1951. Gnetaceae. Pages 336–347 *in* Flora Malesiana I, 4.

Miller, C. N. Jr. 1999. Implications of fossil conifers for the phylogenetic relationships of living families. Bot. Rev. 65:239–277.

Mohr, B. A., M. E. Bernardes-de-Oliveira, A. M. F. Barreto, and M. C. Castro-Fernandes. 2004. Gnetophyte preservation and diversity in the Early Cretaceous Crato Formation (Brazil). VII. International Organization of Paleobotany Conference, Bariloche, Argentina.

Mohr, B. A., and E. M. Friis. 2000. Early angiosperms from the Lower Cretaceous Crato Formation (Brazil), a preliminary report. Int. J. Plant Sci. 161 (6 Suppl.):S155–S167.

Near, T. J., P. A. Meylan, and H. B. Shaffer. 2005. Assessing concordance of fossil calibration points in molecular clock studies: an example using turtles. Am. Nat. 165:137–146.

Ohyama, K., H. Fukuzawa, T. Kohchi, H. Shirai, T. Sano, S. Sano, K. Umesono, Y. Shiki, M. Takeuchi, Z. Chang, S. Aota, H. Inokuchi, and H. Ozeki. 1986. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature 322:572–574.

Palmer, J. D., D. E. Soltis, and M. W. Chase. 2004. The plant tree of life: an overview and some points of view Am. J. Bot. 91:1437–1445.

Posada, D., and K. A. Crandall 1998. ModelTest: testing the model of DNA substitution. Bioinformatics 14:817–818.

Price, R. A. 1996. Systematics of the Gnetales: a review of morphological and molecular evidence. Int. J. Plant Sci. 157 (6 Suppl.):S40–S49.

Pryer, K. M., H. Schneider, A. R. Smith, R. Cranfill, P. G. Wolf, J. S. Hunt, and S. D. Sipes. 2001. Horsetails and ferns are a monophyletic group and the closest living relatives to seed plants. Nature 409:618–622.

Quandt, D., T. Borsch, S. Wicke, H. Won, S. S. Renner, and K. W. Hilu. 2006. Land plant evolution: a perspective from fast-evolving chloroplast regions. BAS abstract 835, http://2006.botanyconference.org/engine/search/index.php?func=detail&aid=835.

Qiu, Y.-L., M. W. Chase, D. H. Les, H. G. Hills, and C. R. Parks. 1993. Molecular phylogenetics of the Magnoliidae: A cladistic analysis of nucleotide sequences of the plastid gene *rbcL*. Ann. Mo. Bot. Gard. 80:587–606.

Qiu, Y.-L., J. Lee, F. Bernasconi-Quadroni, D. E. Soltis, P. S. Soltis, M. Zanis, E. A. Zimmer, Z. Chen, V. Savolainen, and M. W. Chase. 1999. The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. Nature 402:404–407.

Quinn, C. J., R. A. Price, and P. A. Gadek. 2002. Familial concepts and relationships in the conifers based on *rbcL* and *matK* sequence comparisons. Kew Bull. 57:513–531.

Rai, H. S., H. E. O'Brien, P. A. Reeves, R. G. Olmstead, and S. W. Graham. 2003. Inference of higher-order relationships in the cycads from a large chloroplast data set. Mol. Phylogenet. Evol. 29:350–359.

Rambaut, A., and L. Bromham. 1998. Estimating divergence dates from molecular sequences. Mol. Biol. Evol. 15:442–448.

Renner, S. S. 1999. Circumscription and phylogeny of the Laurales: evidence from molecular and morphological data. Am. J. Bot. 86:1301–1315.

Renner, S. S. 2004. Plant dispersal across the tropical Atlantic by wind and sea currents. Int. J. Plant Sciences 165(4 Suppl.):S23–S33.

Renner, S. S. 2005. Relaxed molecular clocks for dating historical plant dispersal events. Trends Plant Sci. 10:550–558.

Renner, S. S., and L-B. Zhang. 2004. Biogeography of the *Pistia* clade (Araceae): Based on Chloroplast and mitochondrial DNA sequences and Bayesian divergence time inference. Syst. Biol.53:422–432.

Ridley, H. N. 1930. The dispersal of plants throughout the world. L. Reeve & Co., Ltd., Ashford.

Rydin, C., M. Källersjö, and E. M. Friis. 2002. Seed plant relationships and the systematic position of Gnetales based on nuclear and chloroplast DNA: conflicting data, rooting problems and the monophyly of conifers. Int. J. Plant Sci. 163:197–214.

Rydin, C., B. A. Mohr, and E. M. Friis. 2003. *Cratonia cotyledon* gen. et sp. nov.: a unique Cretaceous seedling related to *Welwitschia*. Proc. R. Soc. London Ser. B. 270 (Suppl.):29–32.

Rydin, C., K. R. Pedersen, and E. M. Friis. 2004. On the evolutionary history of *Ephedra*: Cretaceous fossils and extant molecules. Proc. Natl. Acad. Sci. USA. 101:16571–16576.

Rydin, C., and N. Wikström. 2002. Phylogeny of *Isoetes* (Lycopsida): resolving basal relationships using *rbcL* sequences. Taxon 51:83–89.

Rydin, C., K. R. Pederson, P. R. Crane, and E. M. Friis. 2006. Former diversity of Ephedra (Gnetales): Evidence from Early cretaceous seed from Portugal and North America. Ann. Bot. 98:123–140.

Sanderson, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. Mol. Biol. Evol. 19:101–109.

Sanderson, M. J., and J. A. Doyle. 2001. Source of error and confidence intervals in estimating the age of angiosperms from *rbcL* and 18S rDNA data. Am. J. Bot. 88:1499–1516.

Schmidt, M., and H. A. W. Schneider-Poetsch. 2002. The evolution of gymnosperms redrawn by phytochrome genes: the Gnetatae appear at the base of the gymnosperms. J. Mol. Evol. 54:715–724.

Setoguchi, H., T. A. Osawa, J.-C. Pintaud, T. Jaffre, and J.-M. Veillon. 1998. Phylogenetic relationships within Araucariaceae based on *rbcL* gene sequences. Am. J. Bot. 85:1507–1516.

Shi, S., H. Jin, Y. Zhong, X. He, Y. Huang, F. Tan, and D. E. Boufford. 2000. Phylogenetic relationships of the Magnoliaceae inferred from cpDNA *matK* sequences. Theor. Appl. Genet. 101:925–930.

Soltis, D. E., P. S. Soltis, and M. J. Zanis. 2002. Phylogeny of seed plants based on evidence from eight genes. Am. J. Bot. 89:1670–1681.

Stevenson, D., and T. Zanoni. 1991. Gnetaceae. Pages 12–18 *in* Flora of the Guianas (series A), vol. 9 (A.R.A. Gorts-van Rijn, ed.).

Stockey, R. A. 1982. The Araucariaceae: an evolutionary perspective. Rev. Palaeobot. Palynol. 37:133–154.

Sun, G., S.-L. Zheng, D. L. Dilcher, Y.-D. Wang, and S.-W. Mei. 2001. Early angiosperms and their associated plants from western Liaoning, China. Shanghai Scientific and Technological Education Publishing House, Shanghai, China.
Swofford, D. L. 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer Associates, Sunderland, Massachusetts.
Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The ClustalX Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24:4876–4882.
Thorne, J. L., and H. Kishino. 2002. Divergence time estimation and rate evolution with multilocus data sets. Syst. Biol. 51:689–702.
Thorne, J. L., Kishino, H., and I. S. Painter. 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol. Biol. Evol. 15:1647–1657.
Treutlein, J., and M. Wink. 2002. Molecular phylogeny of cycads inferred from *rbcL* sequences. Naturwissenschaften 89:221–225.
Tsumura,Y., and Y. Suyama. 1998. Differentiation of mitochondrial DNA polymorphisms in populations of five Japanese *Abies* species. Evolution 52:1031–1042.
Van Roosmalen, M. G. M. 1985. Fruits of the Guianan Flora. Institute of Systematic Botany, Utrecht University.
Wang, X. Q., D. C. Tank, and T. Sang. 2000. Phylogeny and divergence times in Pinaceae: evidence from three genomes. Mol. Biol. Evol. 17:773–781.
Wang, Z.-Q. 2004. A new Permian gnetalean cone as fossil evidence for supporting current molecular phylogeny. Ann. Bot. 94:281–288.
Wiegmann, B. M., D. K. Yeates, J. L. Thorne, and H. Kishino. 2003. Time flies, a new molecular time-scale for brachyceran flye evolution without a clock. Syst. Biol. 52:745–756.
Wikström, N., and P. Kenrick. 2000. Phylogeny of epiphytic *Huperzia* (Lycopodiaceae): paleotropical and neotropical clades corroborated by *rbcL* sequences. Nord. J. Bot. 20:165–171.
Won, H., and S. S. Renner. 2003. Horizontal gene transfer from flowering plants to *Gnetum*. Proc. Natl. Acad. Sci. USA. 100:10824–10829.
Won, H., and S. S. Renner. 2005a. The internal transcribed spacer of nuclear ribosomal DNA in the gymnosperm *Gnetum*. Mol. Phylogenet. Evol. 36: 581–597.
Won, H., and S. S. Renner. 2005b. The chloroplast *trnT-trnF* region in the seed plant lineage Gnetales. J. Mol. Evol. 60: 1–12.
Yang, Y., B.-Y. Geng, D. L. Dilcher, Z.-D. Chen, and T. A. Lott. 2005. Morphology and affinities of an Early Cretaceous *Ephedra* (Ephedraceae) from China. Am. J. Bot. 922:231–241.
Yang, Z. 1997. PAML: A program package for phylogenetic analysis by maximum likelihood *Computer Appl. Biosci.* 13:555–556 (http://abacus.gene.ucl.ac.uk/software/paml.html).
Yoo, M.-J., C. D. Bell, P. S. Soltis, and D. E. Soltis. 2005. Divergence times and biogeography of Nymphaeales. Syst. Bot. 30:693–704.
Zhang, L-B., and S. S. Renner. 2003. The deepest splits in Chloranthaceae as resolved by chloroplast sequences. Int. J. Plant Sci. 164(5 Suppl.):S383–S392.
Zhou, S., S. S. Renner, and J. Wen. 2006. Molecular phylogeny and inter- and intracontinental biogeography of Calycanthaceae. Molecular Phylogenetics and Evolution 39:1–15.

First submitted 23 June 2005; reviews returned 2 September 2005; final acceptance 13 March 2006
Associate Editor: Peter Linder

Male strobilus of *Gnetum gnemon*. The white structures are fertile stamens, the glistening droplets are nectar that is exuded from sterile ovules (Photo T. Stuetzel).