Int. J. Plant Sci. 169(8):1086–1099. 2008.

1058-5893/2008/16908-0010$15.00 DOI: 10.1086/590472

USE OF SIMULTANEOUS ANALYSES TO GUIDE FOSSIL-BASED CALIBRATIONS OF PINACEAE PHYLOGENY

David S. Gernandt,[1,*] Susana Magallón,* Gretel Geada López,† Omar Zerón Flores,‡ Ann Willyard,§ and Aaron Liston‖

*Departamento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México, Apartado Postal 70-233, México, Distrito Federal 04510, Mexico; †Departamento de Biología, Universidad de Pinar del Río, Martí 270 final, esquina a 27 de Noviembre, Pinar del Río, Cuba; ‡Centro de Investigaciones Biológicas, Universidad Autónoma del Estado de Hidalgo, Apartado Postal 1-69, Plaza Juárez, Pachuca, Hidalgo 42001, Mexico; §Department of Biology, University of South Dakota, 414 East Clark Street, Vermillion, South Dakota 57069, U.S.A; and ‖Department of Botany and Plant Pathology, 2082 Cordley Hall, Oregon State University, Corvallis, Oregon 97331-2902, U.S.A.

Uncertainties in the age and phylogenetic position of Pinaceae fossils present significant obstacles to our understanding of the timing of diversification in the family. We demonstrate that simultaneous phylogenetic analyses of chloroplast DNA (*matK* and *rbcL*) and nonmolecular characters that include both extant genera and a limited number of fossil taxa provide useful hypotheses for calibrating molecular trees. Root placements varied for Pinaceae, with Bayesian analyses recovering mutually monophyletic subfamilies Pinoideae and Abietoideae and parsimony analyses recovering Abietoideae as paraphyletic by placing the root between *Cedrus* and the remaining genera. The inferred phylogenetic positions of fossil taxa *Pityostrobus bernissartensis* as the sister group to *Pinus* and *Pseudolarix erensis* as the sister group to extant *Pseudolarix* were used to guide divergence-time calibrations; these calibrations yielded an Early Cretaceous and an Early Jurassic age for crown-group Pinaceae, respectively. The older age estimates based on *Pseudolarix erensis* are supported by weaker evidence from the fossil record but are consistent with recent reports of Early Cretaceous leaf fossils that appear to coincide with extant genera. There remains a great need to characterize the anatomy of extant and fossil species and to code additional nonmolecular characters.

Keywords: phylogenetics, Pinaceae, *Pinus*, *Pityostrobus*, *Pseudoaraucaria*, *Pseudolarix*.

Online enhancements: appendixes.

Introduction

Pinaceae, with 11 genera and ~225 species distributed almost exclusively in the Northern Hemisphere, is the largest extant conifer family (Farjon 2001). Genera such as *Pinus*, *Picea*, *Larix*, *Pseudotsuga*, and *Abies* are important components of Northern Hemisphere forests, and many species are of exceptional economic importance as sources of lumber, pulp, and resins. The family is characterized by the presence of linear leaves (needles), ovulate cones with independent cone scales (each of which has two ovaries), bisaccate pollen, a specialized proembryogeny, and an absence of biflavonoids (Hart 1987; Price 1989; Farjon 1990). Phylogenetic studies have recovered Pinaceae as sister to other extant conifers (Hart 1987; Chaw et al. 1997; Stefanovic et al. 1998; Quinn et al. 2002; Hilton and Bateman 2006; Rai et al. 2008), although molecular analyses have often found the conifers to be rendered paraphyletic by the inclusion of Gnetales, sometimes as sister to Pinaceae (Bowe et al. 2000; Gugerli et al. 2001; Magallón and Sanderson 2002).

Phylogenetic studies of extant Pinaceae genera using nonmolecular (Hart 1987) and molecular characters (Price et al. 1987; Wang et al. 2000; Eckert and Hall 2006) have been fairly consistent with respect to generic relationships within the family. In contrast, simultaneous analyses of fossil taxa and extant genera using nonmolecular characters have yet to demonstrate robust relationships (Alvin 1988; Smith and Stockey 2001, 2002). Our limited understanding of the relationship between fossil and extant Pinaceae is a timely subject because fossils have been increasingly used to calibrate molecular phylogenies used, in turn, to estimate absolute divergence times and molecular rates (reviewed by Willyard et al. [2007]).

An important step in improving divergence-time calibration is to evaluate fossils that may be relevant in providing minimum ages for clades. The fossil record of Pinaceae extends at least to the Cretaceous (Miller 1976) and possibly to the Late Jurassic (LePage 2003). *Pinus belgica*, described from a single ovulate cone (Alvin 1960), has long been considered as the oldest unambiguous evidence of an extant genus. Although *P. belgica* was attributed to the Wealden of Belgium on the basis of the appearance of the rock matrix (Alvin 1960), its provenance is unconfirmed. Nevertheless, this fossil taxon has been used widely but inconsistently in the calibration of molecular phylogenies for *Pinus* and Pinaceae (Wang et al. 2000; Eckert and Hall 2006; reviewed by Willyard et al. [2007]). A competing claim for the oldest record of an extant genus has been proposed for fossil material from localities in Mongolia,

[1] Author for correspondence; e-mail: dgernandt@ibiologia.unam.mx.

Manuscript received December 2007; revised manuscript received January 2008.

described as *Pseudolarix* (Krassilov 1982) and attributed to the Late Jurassic (Keller and Hendrix 1997; LePage 2003). Finally, extinct conifer taxa from the Cretaceous have been anatomically characterized, particularly ovulate cones attributed to the organ genera *Pseudoaraucaria*, *Obirastrobus*, and *Pityostrobus* (Smith and Stockey 2001, 2002). The provenance and age of many of these fossils are better known than those of *P. belgica*.

Phylogenetic analyses using combined morphological and molecular data matrices that code most fossil characters as missing may reveal better alternatives to *P. belgica* for calibrating Pinaceae phylogenies. The objectives of this study are (1) to evaluate Pinaceae fossil taxa for their utility in providing minimum-age calibrations within Pinaceae and within the genus *Pinus*, (2) to perform simultaneous analyses of extant and fossil Pinaceae using nonmolecular and molecular chloroplast (cp) *matK* and *rbcL* data sets, (3) to compare the resulting trees and evaluate each fossil for usefulness in calibration, and (4) to propose fossil-calibrated phylogenetic hypotheses for Pinaceae.

Material and Methods

Taxon Sampling

Nonmolecular and molecular data sets were constructed for all 11 extant genera of Pinaceae, including *Nothotsuga*. *Pinus* was represented by four extant species encompassing its sectional diversity because considerable uncertainty surrounds whether the fossil taxon *Pinus belgica* is sister to all living species or derived from an extant lineage of subgenus *Pinus*. The nonmolecular matrix included fossil taxa *P. belgica* Alvin (1960), *Pseudolarix erensis* Krassilov (1982), *Pseudoaraucaria heeri* Alvin (1957), *Pityostrobus bernissartensis* Alvin (1953, 1957), *Pityostrobus californiensis* Smith and Stockey (2001), and *Pityostrobus corneti* Alvin (1953). *Pinus belgica* and *Pseudolarix erensis* were chosen because they are hypothesized to represent the oldest records for modern Pinaceae genera (Miller 1976; LePage 2003). The choice of specific taxa of *Pityostrobus* and *Pseudoaraucaria* was guided by three criteria: (1) they were among the oldest fossils available (all from the Early Cretaceous), (2) all had relatively precise locality information, and (3) descriptions of these taxa were sufficiently detailed to score a limited number of additional morphological characters. Three species were selected for *Pityostrobus* because it is apparently polyphyletic (Miller 1976; Smith and Stockey 2001, 2002).

A cpDNA sequence matrix was assembled from a combination of GenBank accessions (Hipkins et al. 1990; Brunsfeld et al. 1994; Wakasugi et al. 1994; Wang et al. 1999, 2000; Conran et al. 2000; Gadek et al. 2000; Kusumi et al. 2000; Quinn et al. 2002; Rydin et al. 2002; Leebens-Mack et al. 2005) and new data (app. A in the online edition of the *International Journal of Plant Sciences*). For a subsequent cpDNA calibration, sampling of *Pinus* was increased from four to 33 species (one to five representatives for each subsection) to capture the major lineages of *matK* and *rbcL* found in a previous analysis of 101 species (Gernandt et al. 2005). Outgroups were selected from other gymnosperms. Results from Hart (1987), Smith and Stockey (2001, 2002), and Quinn et al. (2002) were used to guide outgroup selection, with taxa chosen from Araucariaceae, Cupressaceae, Podocarpaceae, Sciadopityaceae, Taxaceae, and Ginkgoaceae. Alignment of Gnetales *matK* (*Gnetum gnemonoides*, AY449625; *Gnetum gnemon*, AY449621; and *Welwitschia mirabilis*, AY492030) was unsatisfactory because of a large number of indels and high sequence divergence, and these taxa were not considered further. The matrix used for analyzing the phylogenetic position of extant and fossil Pinaceae included a total of 26 taxa, and the matrix used for dating the age of Pinaceae diversification included 49.

Nonmolecular Character Matrix

A nonmolecular character matrix was compiled in MacClade 4.08 (Maddison and Maddison 2000) for Pinaceae and six outgroup taxa previously included in cladistic analyses of extant (Hart 1987; Farjon 1990) or extant and fossil Pinaceae (Smith and Stockey 2001, 2002). A total of 105 parsimony-informative characters were chosen: 46 first codified by Hart (1987), 11 added by Farjon (1990), four from Miller (1988), 33 from Alvin (1988) and Smith and Stockey (2002), five recognized as diagnostic for *Pseudolarix* by LePage and Basinger (1995), and six additional characters (app. B in the online edition of the *International Journal of Plant Sciences*). Hart (1987) and Smith and Stockey (2001, 2002) did not separate *Nothotsuga* from *Tsuga*. To code characters separately for these taxa, we relied on morphological descriptions by Frankis (1989) and Farjon (1990) and anatomical descriptions of wood and cones (Hu et al. 1989; Napp-Zinn and Hu 1989; Lin et al. 1995). Further modifications to the matrix were made on the basis of previous studies (Greguss 1955, 1972; Little and Critchfield 1969; Miller 1976, 1992; Alvin 1988; Tomlinson 1994; Takaso and Owens 1995; Biswas and Johri 1997; Wu and Hu 1997; Lin et al. 2002). Thirty-two characters were multistate. The resulting matrix of 26 taxa included 20 characters from wood, roots, and shoots, 15 from leaves, three from biochemistry, eight from pollen and pollen cones, and 59 from seed cones, seeds, and cotyledons.

DNA Character Matrix

New exemplars for this study were extracted from leaves (*Pinus pinea*, *Pinus krempfii*, *Pinus cembra*, *Abies hidalgensis*, *Cathaya argyrophylla*, *Keteleeria davidiana*, and *Picea chihuahuana*) with a modified CTAB protocol (Doyle and Doyle 1987) or from seed tissue (*Pinus nigra*) with a FastDNA Kit (Qbiogene). PCR amplification was performed in 50-μL volumes with the following concentrations: 1× *Taq* buffer (Invitrogen), 1.5 mM $MgCl_2$, 0.2 mM for each dNTP (Invitrogen), 1 μM for each primer, 1.5 U recombinant *Taq* polymerase, and ~0.2 ng of genomic DNA. PCR fragments were ~500 bp each, with three separate reactions per exemplar performed for *matK* and *rbcL*. Primers for *rbcL* were from Wang et al. (1999), and primers for *matK* were from Wang et al. (1999), Gadek et al. (2000), and Gernandt et al. (2005). Thermal-cycle conditions for PCR reactions were as follows: 94°C for 3 min; then 30 cycles of 94°C for 1 min, 55°C for 50 s, and 72°C for 80 s; and a final step of 72°C for 5 min. PCR products were purified with a Geneclean III kit (Qbiogene).

Purified PCR products were sequenced with a DYEnamic ET cycle sequencing kit (Amersham Biosciences) in 10-μL

volumes as follows: 6 μL of purified PCR product, 1 μL of 2 μM primer, 1 μL of DYEnamic Terminator dilution buffer, and 2 μL of DYEnamic reaction mix. Thermal-cycle conditions for the sequencing reaction were 70 cycles of 96 °C for 30 s, 50 °C for 15 s, and 60 °C for 4 min. Sequencing reactions were precipitated in final concentrations of 66% ethanol and 0.08 M sodium acetate (pH = 4.6), and then eluted in 25 μL of MegaBACE loading solution (Amersham Biosciences) for sequencing on an ABI 310 automated sequencer with short (47 cm × 50 μm) capillaries (PE Applied Biosystems).

Sequence reads were edited in BioEdit 6.0.7 (Hall 1999). Alignments were performed with CLUSTAL W (Thompson et al. 1994) and adjusted by eye.

Phylogenetic Analysis

Molecular and nonmolecular matrices are deposited in Treebase (study accession number S1973; matrix accession numbers M3674–M3676). The nonmolecular and cpDNA matrices were analyzed separately and combined using parsimony in PAUP* 4.0b10 for Unix (Swofford 2003). Characters were equally weighted and unordered. The molecular characters were scored as missing for the fossil taxa. Heuristic searches included 10,000 random-addition sequence replicates and tree-bisection-reconnection (TBR) branch swapping. Bootstrap values were based on 10,000 replicates of a heuristic search, each with 10 random-addition sequence replicates and TBR branch swapping, saving all trees. Decay values were determined using converse constraints in PAUP*, with a 10,000-replicate heuristic search using random-addition sequence and TBR branch swapping.

Bayesian inference using separate molecular and morphological data partitions has been offered as a useful alternative to parsimony, for example, when the partitions provide heterogeneous signal (Nylander et al. 2004). For this reason, we performed separate analyses in MRBAYES 3.1.2 (Huelsenbeck and Ronquist 2001; Ronquist and Huelsenbeck 2003). Separate models were applied for nonmolecular characters, *matK*, and *rbcL*. The nonmolecular characters were treated as standard discrete data, with all changes between states equally probable (Lewis 2001). The Akaike Information Criterion (AIC), as implemented in ModelTest 3.7 (Posada and Crandall 1998; Posada and Buckley 2004), was used to choose the best-fitting models for *matK* and *rbcL* separately. The nonmolecular and molecular partitions were analyzed separately and then combined using two independent Markov chain Monte Carlo (MCMC) runs performed on random starting trees with a temperature of 0.2. Each run included five million generations with four incrementally heated chains. A tree was sampled every 200 generations, with the number of trees to be excluded for the burn-in (20%) verified by identifying the stable phase of a plot of log-likelihood values across generations. The first 5000 trees, corresponding to one million generations, were excluded before calculation of the majority-rule consensus, which was based on 20,000 trees corresponding to the stable phase of the chain. Posterior probabilities (PPs) for each MCMC run were calculated from a majority rule of the sampled trees after burn-in.

To evaluate topological congruence between the nonmolecular and cpDNA data sets, parsimony trees were inspected for the presence of conflicting clades receiving >70% bootstrap support (Mason-Gamer and Kellogg 1996). Two tests of topological congruence also were performed in PAUP* (excluding the fossil taxa). For the incongruence length difference (ILD) test (Farris et al. 1995), the sums of tree lengths resulting from searches on 1000 randomly assigned partitions were compared to the tree lengths of the actual partition and tested for significant differences. The Templeton (1983) test was also applied to determine whether trees resulting from either matrix resulted in a significantly poorer fit for the character state data from the other matrix.

Age Estimation

Age estimation was performed on the 49-taxon molecular matrix. The tree with the best likelihood score was chosen from the sample of Bayesian replicates, and its branch lengths were estimated in PAUP* using maximum likelihood. Before model selection and branch length estimation, a 130-bp segment was trimmed from the 3′ end of the *matK* partition because it was incomplete in the outgroups. A GTR + Γ + I model was applied to the unpartitioned cpDNA sequence matrix, with parameters estimated from the data.

Absolute ages and molecular rates were estimated in r8s, version 1.7 for Unix (Sanderson 2002). All outgroups were pruned except for *Araucaria*; retaining that genus permitted inclusion of the crown-group node of extant crown conifers if Pinaceae is considered the sister group to the remaining families. Polytomies were treated as collapsed. The χ^2 test indicated that the Langley-Fitch molecular clock was rejected (results not shown). As a result, only age estimates using penalized likelihood are presented. The magnitude of the penalty for rate changes across branches, the smoothing parameter (λ), was optimized for each calibration with the branch-pruning cross-validation procedure, which used five guesses, three restarts, and a perturbation factor of 0.05. The error was estimated for $\log_{10}$ cross-validation values ranging from 0 to 8 in increments of 0.2.

The r8s program allows for one or more nodes to be fixed in a calibration and for others to be constrained to a minimum or a maximum age. The fixed calibration points on the cpDNA tree were determined by the inferred phylogenetic position of fossil taxa with respect to extant genera. Additional fossils were taken into account to apply minimum-age constraints, which accommodate rate variability and place a lower limit on the age of specific nodes. When radiometric dates were unavailable, fossil ages were assigned on the basis of a recent geologic timescale (Gradstein and Ogg 2004). *Pseudolarix erensis* was described from the Gurvan-Eren, Manlaj, and Bon-Tsagaan localities associated with the Tsagaan-Tsav Formation, Mongolia, with a radiometric date of 156 ± 0.76 Ma (Late Oxfordian; Keller and Hendrix 1997; LePage 2003). The age of *Pityostrobus bernissartensis* (Alvin 1957) was reported from the Puits Négresse, Bernissart, Belgium (as *Pityostrobus bommeri*; Alvin 1953). The Sainte Barbe Formation at Bernissart was recently dated as mid-Barremian to earliest Aptian (Yans et al. 2005; range: 127.5–118.5 Ma, midpoint: 123 Ma). For the remaining Wealden taxa (*Pityostrobus corneti, Pseudoaraucaria heeri,* and *Pinus belgica*), several recent publications were taken into account to establish an age range of 127.3–89.5 Ma (Robaszynski et al. 2001; Yans et al. 2005; Quinif et al. 2006; Dejax et al. 2007). However, after the phylogenetic results were considered

(see "Phylogenetic Analysis"), these taxa were ultimately not used for calibration or to apply constraints. A similar situation applied to *Pityostrobus californiensis*, although a fairly precise age could be assigned to this taxon, corresponding to the midpoint of the Bedoulian stage from which it was described (Smith and Stockey 2001; middle Aptian, 118.5 Ma).

Additional minimum-age constraints were applied to extant genera and a *Pinus* section based on younger fossils not included in the phylogenetic analyses. Eocene records exist for *Abies* (46–45 Ma; Axelrod 1976; Erwin and Schorn 2005), *Larix* (44.85 Ma: LePage and Basinger 1991; LePage 2003; 45 Ma, based on radiometric dating: Schorn 1994), *Picea* (46–45 Ma; LePage 2003; Erwin and Schorn 2005), *Keteleeria* (44.4 Ma; LePage 2003), and *Tsuga* (44.85 Ma; LePage 2003). Oligocene fossils have been attributed to *Pseudotsuga* (32 Ma; Axelrod 1976; Schorn 1994). *Pinus* subsection *Balfourianae* and probably subsection *Cembroides* are present in the Oligocene (27.2 Ma; Wolfe and Schorn 1990), indicating that the crown group of section *Parrya* was present at the time. Age estimates were also evaluated with and without applying maximum-age constraints for crown-group conifers (the most recent common ancestor of Pinaceae-Araucariaceae) based on the first unequivocal appearance of conifers fossils in the Late Carboniferous (Pennsylvanian, 299–318 Ma, midpoint: 308.5 Ma; Rothwell et al. 1997).

Substitution rates were estimated for *rbcL* and *matK* separately with penalized likelihood and two calibration points. A silent rate was estimated for each gene separately on the basis of third-codon positions. An overall rate was also calculated using all sequence positions. The same tree topology was used for age estimation and substitution rates (see "Age Estimates"), but to avoid zero-length branches, only one species per section was used for *Pinus*. Nucleotide substitution models were chosen separately for each partition on the basis of the AIC test, and branch lengths were calculated with maximum likelihood.

Results

Phylogenetic Analysis

The 26-taxon matrix included 105 parsimony-informative nonmolecular characters. The fossil taxa had a high number of cells scored as missing or inapplicable: 82 for *Pseudolarix erensis*, 58 for *Pinus belgica*, 58 for *Pseudoaraucaria heeri*, 61 for *Pityostrobus bernissartensis*, 66 for *Pityostrobus californiensis*, and 66 for *Pityostrobus corneti*. Among extant taxa, those with the most missing or inapplicable cells were *Ginkgo* (53), *Amentotaxus* (48), *Podocarpus* (36), *Araucaria* (31), and *Nothotsuga* (39). Fossil taxa were missing many characters related to wood, leaves, and pollen reproductive structures, while outgroups *Ginkgo*, *Amentotaxus*, and *Podocarpus* had many inapplicable characters for seed reproductive structures.

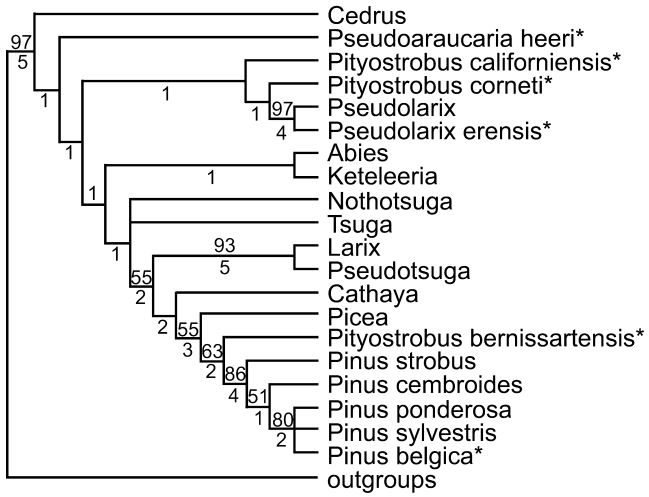
The heuristic search of the nonmolecular matrix recovered three most parsimonious trees (MPTs) with a length (L) of 303 steps, a consistency index (CI) of 0.498, and a retention index (RI) of 0.675. For the Bayesian analysis of the same data, the average standard deviation of split frequencies was 0.00608 at the end of the MCMC run. More nodes were resolved in the parsimony strict consensus tree (fig. 1*A*) than in the Bayesian majority-rule tree (fig. 1*B*).

The cpDNA partition of the same matrix (with the six fossil taxa deleted) was 2838 bp in length (*rbcL*: 1262 bp; *matK* plus 289 bp of 3′ spacer: 1576 bp). No gaps were needed for *rbcL*, but 14 insertion-deletion events were hypothesized to align *matK*. The matrix included 1042 variable characters, 569 of which were parsimony informative. Parsimony analysis recovered one MPT ($L = 1892$ steps, CI = 0.695, $CI_{exc} = 0.571$ [excluding variable but parsimony-uninformative characters], RI = 0.684; fig. 1*C*). Bayesian analysis (fig. 1*D*) used the Hasegawa-Kishino-Yano model (HKY + Γ + I) for *rbcL* and the general time-reversible model (GTR + Γ) for *matK*. The average standard deviation of split frequencies was 0.000444 at the end of the MCMC run. For both parsimony and Bayesian analyses, branch support on the cpDNA trees was equal to or higher than that on the nonmolecular trees.

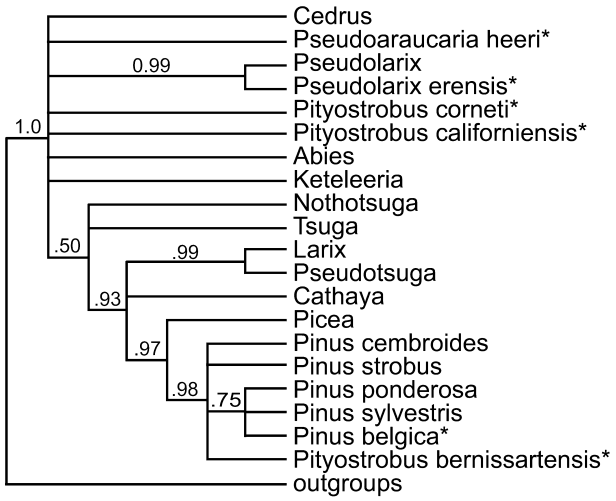
The 26-taxon combined nonmolecular and cpDNA data matrix included 1147 variable characters, 674 of which were parsimony informative. The parsimony search yielded three MPTs ($L = 2211$ steps, CI = 0.663, $CI_{exc} = 0.552$, RI = 0.675). For the Bayesian analysis, the average standard deviation of split frequencies was 0.00866 at the end of the MCMC run. All branch-support measures were lower than in the analyses of cpDNA alone (fig. 1*C*, 1*D*), with the most marked reductions in decay indices, partly because of the unstable position of the fossil taxa in slightly suboptimal trees.

Parsimony trees based on nonmolecular, cpDNA-only, and combined matrices resolved subfamily Abietoideae as paraphyletic to subfamily Pinoideae, placing the root between *Cedrus* and the remaining members of Pinaceae (fig. 1*A*, 1*C*, 1*E*). In contrast, the Bayesian majority-rule trees for the nonmolecular matrix failed to resolve Abietoideae (fig. 1*B*), while both subfamilies were monophyletic for the cpDNA and combined matrices (fig. 1*D*, 1*F*). The fossil taxa *Pseudoaraucaria heeri*, *Pityostrobus corneti*, *Pityostrobus californiensis*, and *Pseudolarix* occurred within the Abietoideae clade or grade, but their positions within the group varied. In contrast, *Pseudolarix erensis* consistently grouped with extant *Pseudolarix*. Relationships among the Pinoideae clade were more stable among analyses. *Cathaya*, *Picea*, (*Pityostrobus bernissartensis*), and *Pinus* formed a paraphyletic grade with nonmolecular characters, while *Cathaya* and *Picea* formed a monophyletic sister group to *Pinus* with cpDNA. *Pityostrobus bernissartensis* was sister to *Pinus* in the parsimony analyses (fig. 1*A*, 1*E*) and formed a trichotomy with two lineages of *Pinus* in the Bayesian analyses (fig. 1*B*, 1*F*). *Pinus belgica* formed a clade with *Pinus ponderosa* and *Pinus sylvestris* in all analyses.

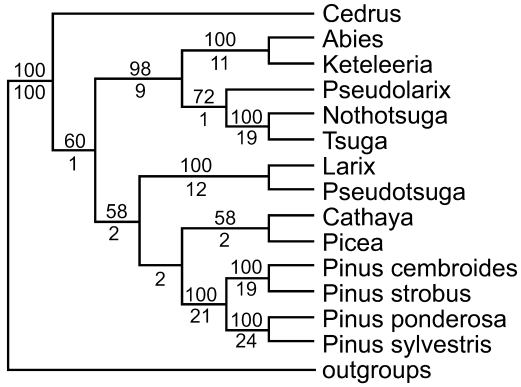
Topological differences were recovered between the nonmolecular and cpDNA-based matrices, although conflicting clades did not receive high branch support. After deletion of the six fossil taxa present only in the nonmolecular matrix, the ILD test found significant conflict between the nonmolecular and cpDNA matrices ($P = 0.002$). The Templeton (1983) test found that the nonmolecular matrix gave a significantly poorer fit on the single most parsimonious molecular tree than on the single fossil-pruned nonmolecular tree (292 vs. 275 steps; $P = 0.009$) and, conversely, that the molecular data gave a significantly worse fit on the nonmolecular tree (2007 vs. 1892 steps; $P < 0.0001$).



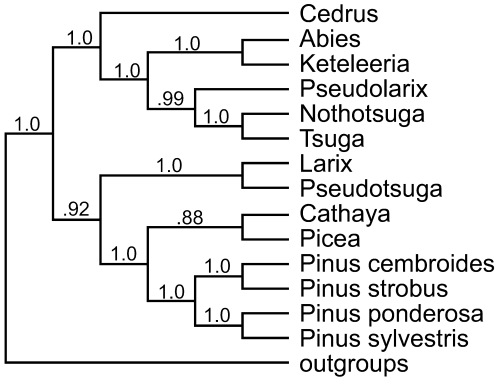
A. non-molecular parsimony



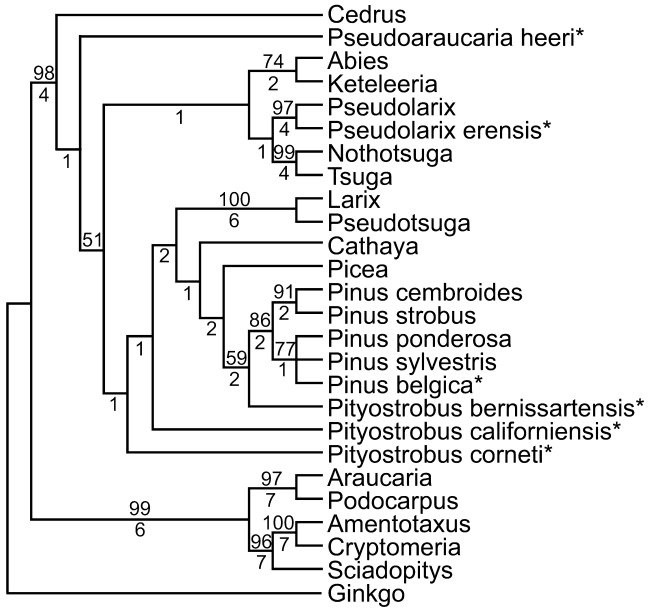
B. non-molecular Bayesian



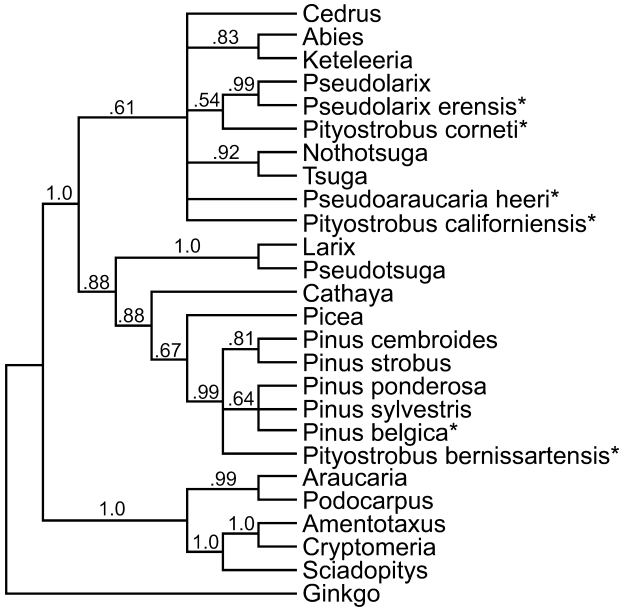
C. cpDNA parsimony



D. cpDNA Bayesian



E. combined parsimony



F. combined Bayesian

Diagnostic, nonmolecular synapomorphies for Pinaceae were identified by using the "trace all changes" option in MacClade to display the near-maximum number of changes on the combined parsimony tree (fig. 1*E*). The monophyly of Pinaceae was supported by 11 diagnostic synapomorphies: protein-accumulating sieve elements, absence of phloem fiber sclereids, tracheids of leaf transfusion tissue surrounding the vascular bundle, absence of biflavonoids, absence of sperm nuclei cell walls, ventral canal cell without a distinct wall, megaspore membrane thin at the micropylar end, primary type proembryo wall formation, four-tiered proembryo, two ovules per scale, and seed wing originating from scale epidermis. All but the last two characters were scored as missing in fossil taxa. In the parsimony tree, *Pseudoaraucaria* was the sister group to all members of Pinaceae except *Cedrus*, but only homoplasious character state changes supported its position. Extant *Pseudolarix* and *Pseudolarix erensis* were united by brachioblasts with close annual rings, deltoid triangular cone scales, and semitrullate, pointed seed wings. *Nothotsuga* and *Tsuga* were united by the presence of a pendulous primary shoot, and support for *Pityostrobus californiensis* as the sister group to subfamily Pinoideae was given by the absence of resin vesicles in the seed coat. Constraining subfamily Abietoideae to monophyly to reflect the Bayesian result revealed four diagnostic characters: abietoid immunology, seed scales with a pedicillate base, presence of a seed trace (scored as missing in *Nothotsuga*, fossil *Pityostrobus* spp., and *P. erensis*), and seed wings attached as a deep cup.

Monophyly of subfamily Pinoideae was supported by the presence of horizontal resin ducts in the rays; piceoid, pinoid, or fenestriform cross-field pits; two root resin canals; pinoid chemical immunology; absence of a seed trace; and seed wings attached as a shallow cup or a claw. Diagnostic characters supporting the sister relationship of *Larix* and *Pseudotsuga* were the presence of phloem fiber sclereids and an asymmetric micropyle. These two genera were in a sister position to a clade consisting of *Cathaya*, *Picea*, *Pinus*, and *Pityostrobus bernissartensis* and supported by the presence of plicate mesophyll parenchyma (scored as missing or polymorphic in *Picea* and the fossil taxa). Monophyly of *Pinus* and *P. bernissartensis* was supported by tongue-shaped ovulate cone scales and clawlike seed wing attachment. *Pityostrobus bernissartensis* formed a trichotomy with two *Pinus* lineages in the Bayesian analysis, whereas *Pinus* was monophyletic in the parsimony tree, as supported by the fusion of the vascular traces leading to the bract and scale in the ovulate cone. Extant *Pinus* species were united by needlelike leaves spirally arranged in fascicles on short shoots and by a thin epithelium of vertical resin canals (scored as missing in *P. belgica*). *Pinus cembroides* and *Pinus strobus* were sister groups, but no diagnostic character state was identified. *Pinus ponderosa*, *P. sylvestris*, and *P. belgica* were united by platelike thickenings of transverse tracheid walls.

The 49-taxon (with expanded *Pinus* sampling) cpDNA matrix had 1074 variable characters, 614 of which were parsimony informative. The heuristic search recovered 1480 MPTs ($L = 2045$ steps, CI = 0.668, $CI_{exc} = 0.551$, RI = 0.785). The family Pinaceae was monophyletic in the strict consensus tree (not shown). The root of Pinaceae in the strict consensus tree was unresolved, with 90% of the MPTs indicating that the root was placed between *Cedrus* and the remaining genera. Neither subfamily Pinoideae nor subfamily Abietoideae was resolved as monophyletic in the strict consensus tree. The only generic relationships resolved in the strict consensus tree for subfamily Pinoideae was the sister position of *Pseudotsuga* and *Larix*. Except for *Cedrus*, subfamily Abietoideae had the same topology as that recovered in the cpDNA analyses of 20 taxa (fig. 1*C*, 1*D*).

For the 49-taxon data set, the best *rbcL* model was K81uf + Γ + I, and the best *matK* model was TVM + Γ. The mean standard deviation of split frequencies between the two MCMC runs was 0.00362 after five million generations. The tree with the best log-likelihood value ($-\ln = 14,582.626$) was found in generation 4,199,800 of the first run (fig. 2; fig. C1, available in the online edition of the *International Journal of Plant Sciences*, shows a majority-rule tree with PP values). Pinaceae divided into two clades corresponding to subfamilies Pinoideae and Abietoideae. Generic relationships were identical to those found in the Bayesian analysis of 20 taxa (fig. 1*D*). PP support for intergeneric relationships was >0.95 in all cases, except for the sister relationship between *Picea* and *Cathaya* (PP = 0.79) and monophyly of subfamily Pinoideae (PP = 0.90). Within *Pinus*, the subgeneric, sectional, and subsectional relationships were identical to those in previous cpDNA studies (Gernandt et al. 2005; Eckert and Hall 2006).

Age Estimates

With the results from the combined analyses (fig. 1*E*, 1*F*), two calibration scenarios were evaluated. A conservative scenario (fig. 3) took into account the position of *Pityostrobus bernissartensis* as the sister group to *Pinus* by calibrating the node corresponding to the most recent common ancestor of *Pinus*, *Picea*, and *Cathaya* at 123 Ma (hereafter the *Pinus-Picea* calibration). A second scenario (fig. 4) took into account the position of *Pseudolarix erensis* as the sister group to extant *Pseudolarix* by calibrating the node for the most recent common ancestor of *Pseudolarix*, *Tsuga*, and *Nothotsuga* at 156 Ma (hereafter the *Pseudolarix-Tsuga* calibration). *Pinus belgica*, *Pityostrobus californiensis*, *Pityostrobus corneti*, and *Pseudoaraucaria heeri* were not used to set minimum-age constraints in the conservative *Pinus-Picea* calibration because their positions were not considered robust enough. The last three fossils were too young to be relevant in the *Pseudolarix-Tsuga* calibration.

The *Pinus-Picea* calibration (fig. 3) yielded a younger age at all nodes than did the *Pseudolarix-Tsuga* calibration (fig. 4). The Pinaceae crown group was estimated to have first diversified in the Early Cretaceous (*Pinus-Picea* calibration) or the

Fig. 1 Comparison of phylogenetic hypotheses in Pinaceae. *A*, Parsimony strict consensus of nonmolecular characters. *B*, Bayesian majority rule of nonmolecular characters. *C*, Parsimony strict consensus of chloroplast (cp) DNA. *D*, Bayesian majority rule of cpDNA. *E*, Parsimony strict consensus of combined characters. *F*, Bayesian majority rule of combined characters. Asterisks denote fossil taxa; values above branches are bootstrap values (*A*, *C*, and *E*) or posterior probability values (*B*, *D*, and *F*); values below branches (*A*, *C*, and *E*) are decay indices.

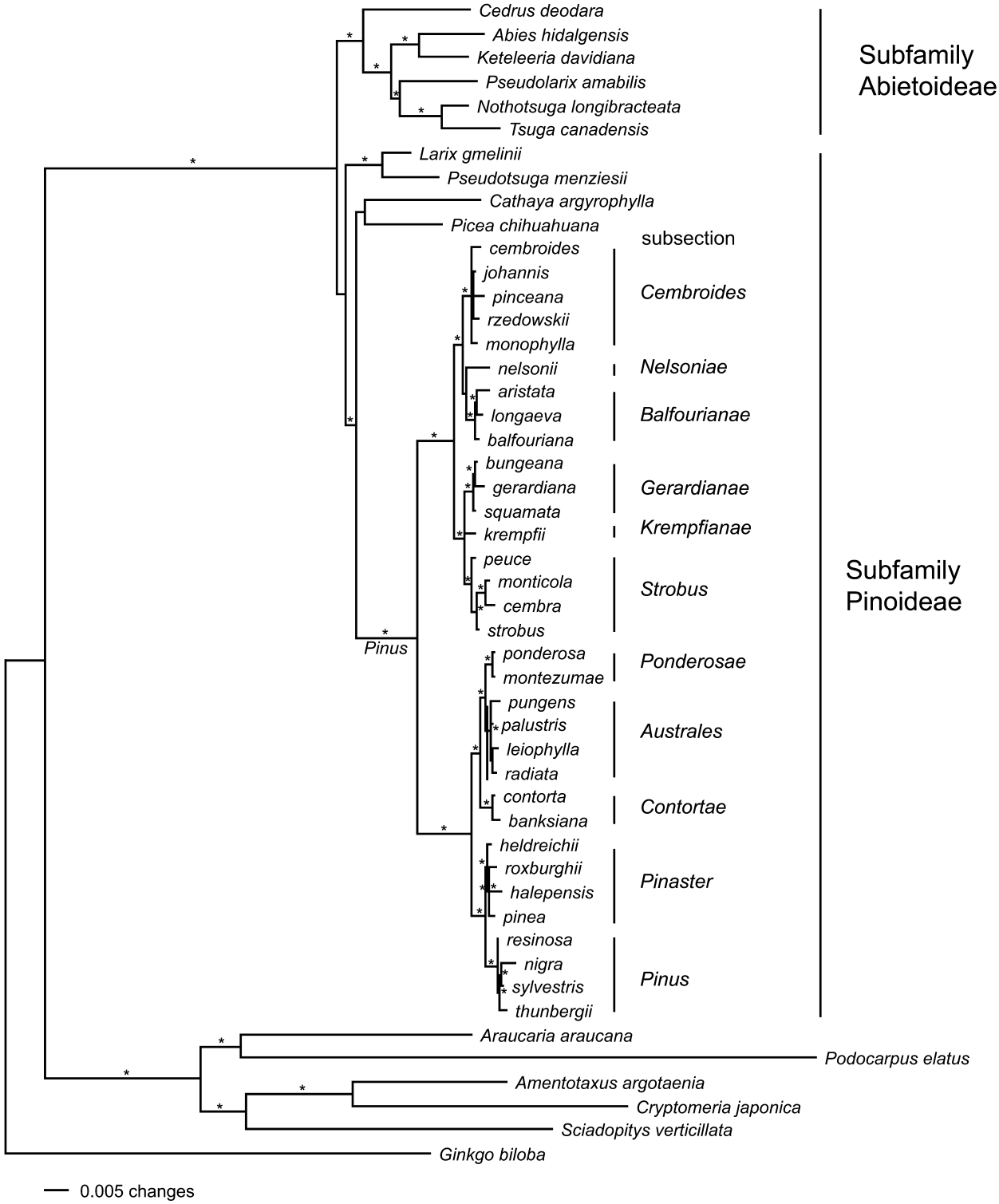


Fig. 2 Best Bayesian tree for Pinaceae. *Pinus* subsections and Pinaceae subfamilies are indicated on the right. Posterior probability values >0.95 are indicated by asterisks. The branch lengths are proportional to the number of nucleotide changes, inferred with maximum likelihood.

Early Jurassic (*Pseudolarix-Tsuga* calibration). Calibration decisions had a great influence on age estimates, particularly at the deepest nodes. For example, for the *Pinus-Picea* calibration (fig. 3), the *Pseudolarix-Tsuga* node was estimated at 74 Ma, the *Pinus-Picea* node was calibrated at 123 Ma, and the age of the subgenus *Pinus-Strobus* node was estimated at 72 Ma. In contrast, for the *Pseudolarix-Tsuga* calibration (fig. 4), the *Pseudolarix-Tsuga* node was calibrated at 156 Ma (82 Myr

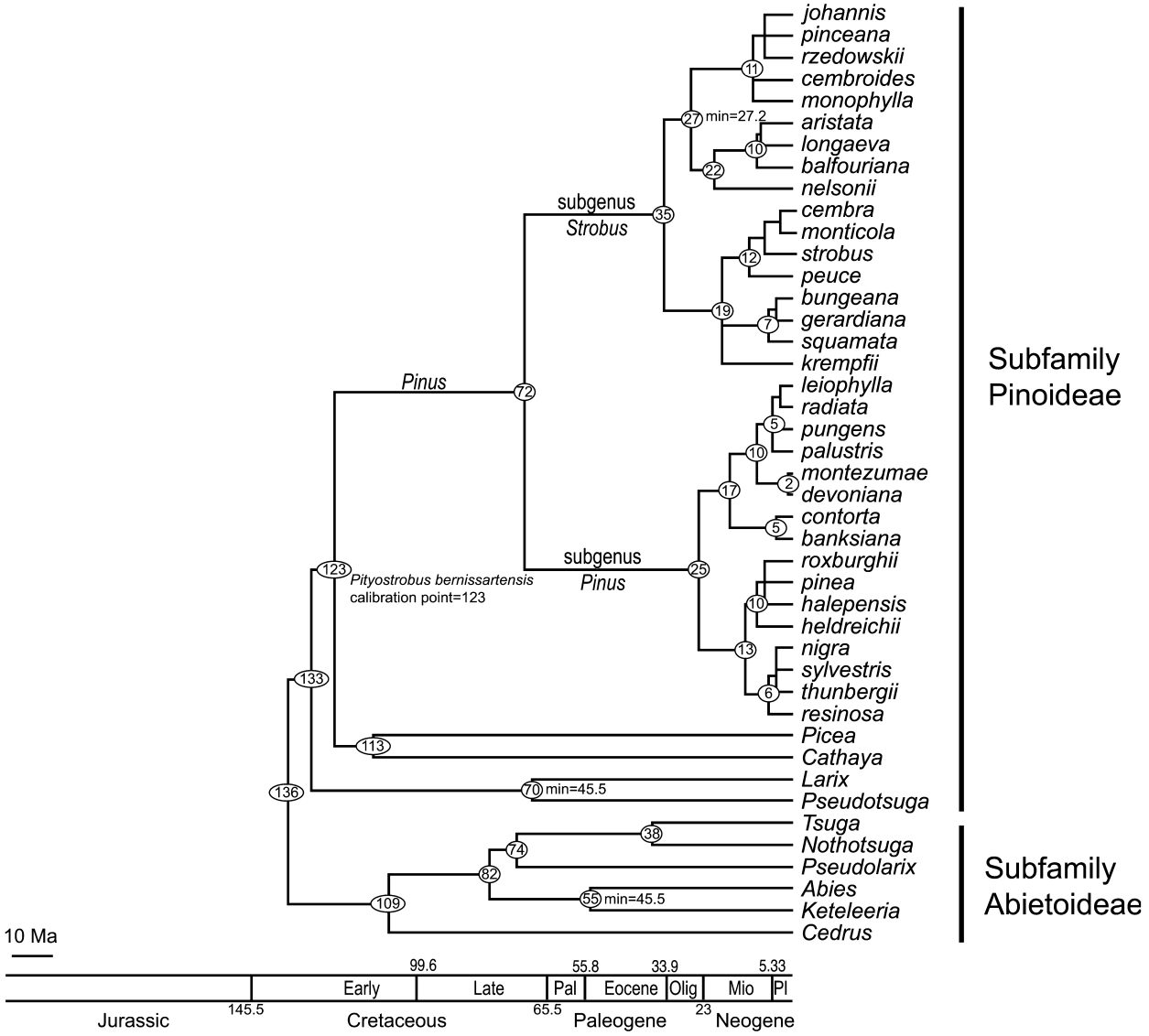


Fig. 3 Chronogram for Pinaceae based on a conservative fossil calibration with penalized likelihood. The fixed calibration point is the *Pinus-Picea* split at 123 Ma.

older), the *Pinus-Picea* node was estimated at 155 Ma (32 Myr older), and the subgenus *Pinus-Strobus* node was estimated at 87 Ma (15 Myr older).

To evaluate the effect of constraining the Pinaceae-Araucariaceae divergence to a maximum age of 308.5 Ma, ages for five representative nodes were compared with and without the age constraint and by increasing or decreasing the constrained value by 20% (table 1). The ages estimated for the deepest nodes in the tree were unrealistically old when maximum-age constraints were not enforced. The removal of the maximum-age constraint also changed the results of the cross-validation test, resulting in lower error rates at higher smoothing values (allowing for more clocklike substitution rates). Constraining the Pinaceae-Araucariaceae node to 370.2, 308.5, or 246.8 Ma had a minimal effect on the optimal smoothing parameter (λ), but a greater effect was observed on the age estimates, particularly at the deepest nodes (table 1).

With the *Pinus-Picea* calibration, the *rbcL* silent-substitution rate estimated with penalized likelihood was 3.28×10^{-10} substitutions/site/year (421 third-codon positions; range: 3.26×10^{-10} to 3.30×10^{-10}). The rate for all 1262 sites was 1.46×10^{-10} substitutions/site/year (range: 0.878×10^{-10} to 2.17×10^{-10}). For the *Pseudolarix-Tsuga* calibration, the *rbcL* silent-substitution rate was 2.61×10^{-10} substitutions/site/year (range: 1.91×10^{-10} to 3.75×10^{-10}). The rate for all 1262 sites was 0.962×10^{-10} substitutions/site/year (range: 0.537×10^{-10} to 1.84×10^{-10}). The *Pinus-Picea* calibration yielded a *matK* silent-substitution rate estimate of 4.02×10^{-10} substitutions/site/year (471 third-codon positions; range: 3.52×10^{-10} to 4.50×10^{-10}). The rate for 1414 sites, excluding 152 bp of a 3′ flanking region, was 2.80×10^{-10} substitutions/site/year (range: 1.51×10^{-10} to 3.84×10^{-10}). With the *Pseudolarix-Tsuga* calibration, the *matK* silent-substitution rate was 2.91×10^{-10} substitutions/

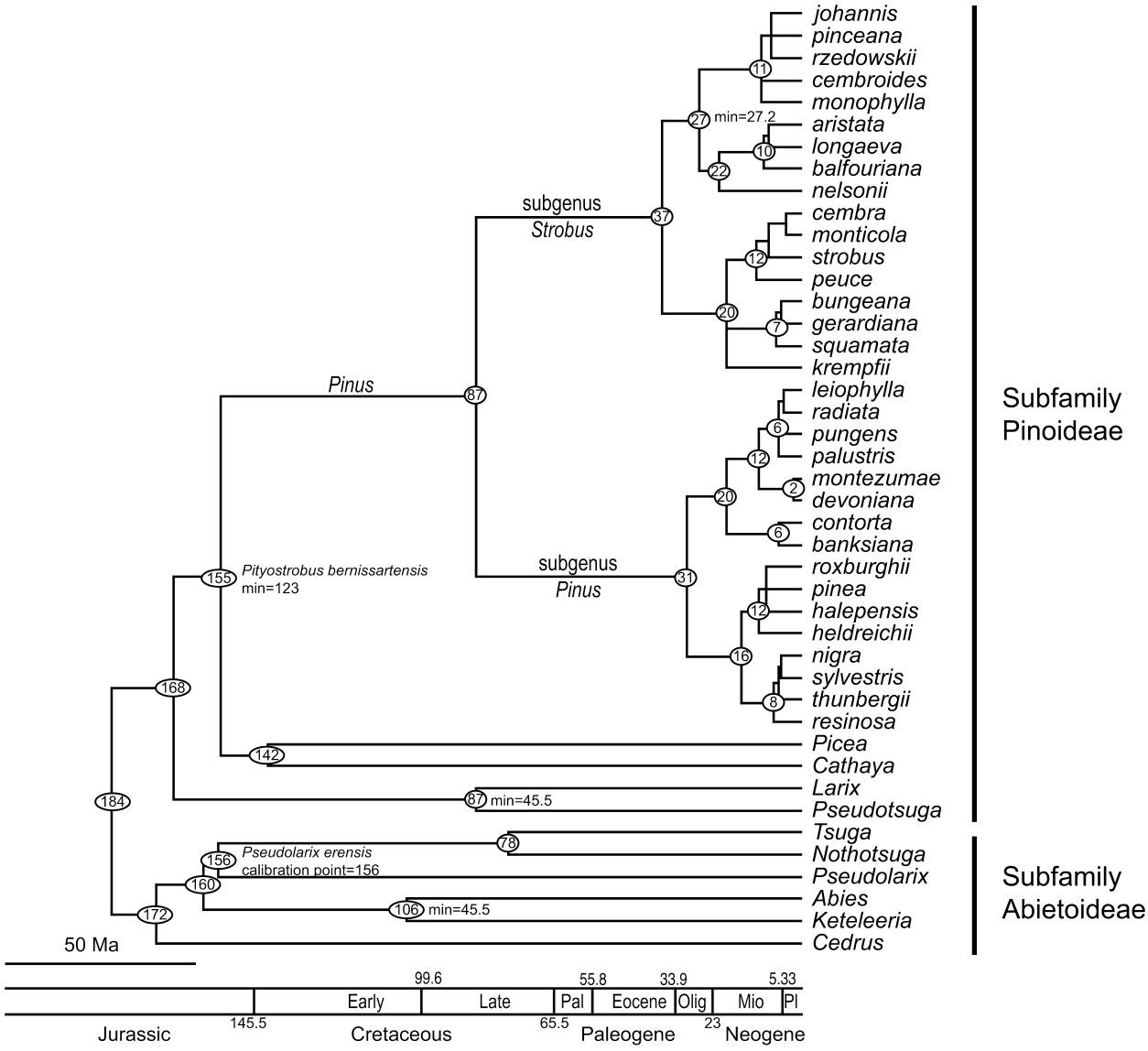


Fig. 4 Chronogram for Pinaceae based on a liberal fossil calibration with penalized likelihood. The fixed calibration point is the *Pseudolarix-Tsuga* split at 156 Ma.

site/year (range: 1.59×10^{-10} to 4.43×10^{-10}), and the rate for 1414 sites was 2.25×10^{-10} substitutions/site/year (range 1.69×10^{-10} to 3.22×10^{-10}).

Discussion

Assigning Ages to the Oldest Pinaceae Fossil Cones

Pinaceae is well represented in the fossil record throughout the Cretaceous, but with the exception of *Pinus* and *Pseudolarix*, ovulate cone fossils from that time do not belong to extant genera but are assigned to organ genera *Pseudoaraucaria*, *Pityostrobus*, and *Obirastrobus* (Smith and Stockey 2002). In Europe, the oldest Cretaceous fossils reported for *Pinus*, *Pseudoaraucaria*, and *Pityostrobus* are from the Wealden facies, mainly the Mons Basin of Belgium (Alvin 1953, 1957, 1960) but also sites in northern France and southern England (Alvin 1988). The ages for Cretaceous deposits in Belgium have been revised, and works on specific localities have offered an age range for the Belgian Wealden of 127.3–89.5 Ma (Barremian-Turonian), although the Wealden is highly complex and some formations may be found to extend to the Late Jurassic (Robaszynski et al. 2001). Pinaceae fossils from these sites usually lack collection information precise enough to link them to a specific excavation or a specific depth. The most problematic example is that of *Pinus belgica*, which was attributed to the Wealden on the basis of adhering particles but is of unknown provenance (Alvin 1960). Although *P. belgica* ultimately was not used for calibration, we note that it may not be accurate to assign the midpoint age of the Wealden (108.4 Ma) to this taxon. The upper age limit of the Wealden, 89.5 Ma, corresponds better to the next appearance of an anatomically characterized *Pinus* ovulate cone in the fossil record, *Pinus mutoi*

Table 1

Effect of Maximum-Age Constraints on the Ages of Five Pinaceae Nodes

	Pinus-Picea calibration				*Pseudolarix-Tsuga* calibration			
Node	None	370.2 Ma	308.5 Ma	246.8 Ma	None	370.2 Ma	308.5 Ma	246.8 Ma
Optimal smoothing (λ)	6.0	2.2	2	2	6.0	2.4	2.4	2.2
Pinaceae crown (Ma)	148.1	139.8	136.0	131.5	271.2	195.6	183.9	173.2
Pinus-Picea[a] (Ma)	123	123	123	123	227.7	165.3	155.1	148.1
Larix-Pseudotsuga (Ma)	62.3	69.4	70.5	68.7	113.8	94.8	87.0	85.8
Nothotsuga-Tsuga (Ma)	44.9	38.8	37.7	35.9	81.5	78.7	78.4	79.6
Subsect. *Cembroides* (Ma)	9.3	10.1	10.5	10.4	16.0	11.0	10.6	10.7

Note. Columns represent maximum-age constraints of 308.5 Ma, 370.2 Ma (308.5 + 20%), and 246.8 Ma (308.5 − 20%) and the absence of a constraint.

[a] Constrained node.

(Saiki 1996) from the Coniacian (85.8–89.3 Ma) of Japan. Historical misassignment of the age of *P. belgica* may explain the perceived gap in the Cretaceous fossil record between the first and subsequent appearances of *Pinus* (Miller 1976).

A second relevant example is the age assigned to *Pityostrobus bernissartensis*. This taxon was reported from a Bernissart locality, Puits Négresse, that has not been precisely dated, but the age and palynological composition of a nearby Bernissart locality have been subject to several studies, particularly at the depth of 322 m, where many *Iguanadon* fossils were found (Yans et al. 2005; Quinif et al. 2006; Dejax et al. 2007). Seward (1900) reported that Pinaceae fossils from the *Iguanadon* pit at Bernissart were taken from a depth of greater than 250 m, and conifer pollen is reported from 322 m (Dejax et al. 2007). The age assigned here for *Pityostrobus bernissartensis*, based on the age of the *Iguanadon* site (range: 127.5–118.5 Ma, midpoint: 123 Ma), is slightly older than the midpoint age of the Wealden as a whole (range: 127.3–89.5 Ma, midpoint: 108.4 Ma).

In contrast to the relatively wide age intervals associated with the Wealden, the provenance of *Pityostrobus californiensis* was reported in detail (Smith and Stockey 2001), allowing for the assignment of a more precise age (middle Aptian, 118.5 Ma). The age and inferred phylogenetic position of this and the Wealden taxa considered here support the hypothesis that modern Pinaceae genera began to diverge no later than the Aptian.

Considering the diversity of Pinaceae in the Cretaceous, Alvin (1960) expressed surprise that extant Pinaceae genera were not known from the Jurassic. More recently, LePage and Basinger (1995) reviewed the fossil record of *Pseudolarix* and concluded that fossil seeds, cone scales, brachioblasts, and leaves extending throughout the Cretaceous are morphologically indistinguishable from those of the extant species *Pseudolarix amabilis*. Cretaceous fossils accepted by LePage and Basinger (1995) were recorded from the Portlandian/Berriasian, Berriasian, Neocomian, Neocomian/Aptian, Barremian, Albian, Cenomanian, and Turonian. Subsequent $^{40}Ar/^{39}Ar$ dating of fossil deposits correlated to localities bearing *Pseudolarix* in Mongolia (Keller and Hendrix 1997) would extend the fossil record of the genus to the Oxfordian (Late Jurassic; Krassilov 1982; LePage 2003). Dates for the specific formations where *Pseudolarix* was described are needed, but these findings suggest that *Pseudolarix* has the oldest and most continuous fossil record of any extant Pinaceae genus.

Miller (1976) interpreted *Compsostrobus neotericus* from the Late Triassic of North Carolina as suggestive that extant Pinaceae may have been present at that time. Age estimates for crown-group Pinaceae were heavily influenced by use of a maximum-age constraint one node removed (table 1), but only use of a higher maximum-age constraint (370.2 Ma) with the *Pseudolarix-Tsuga* calibration gives a crown-group Pinaceae age estimate that approaches support for the possibility suggested by Miller (195.6 Ma, Early Jurassic; table 1).

Phylogenetic Relationships among Extant Genera and Six Fossils

Integration of fossil taxa into phylogenetic hypotheses is a critical step in understanding the evolutionary history of Pinaceae. Analyses of nonmolecular data, alone (fig. 1*A*, 1*B*) and in combination with molecular characters (fig. 1*E*, 1*F*), yield phylogenies that resolve six fossil taxa within the crown of Pinaceae. Inferred tree topologies incorporating nonmolecular and molecular data that are unavailable for the fossil taxa bore a much stronger resemblance to molecular results than did trees inferred solely from ovulate cone characters (Smith and Stockey 2001, 2002). This and prior molecular studies are generally consistent with the classification by Van Tieghem (1891), who divided Pinaceae into Cédrées (subfamily Abietoideae: *Tsuga*, *Cedrus*, *Abies*, *Keteleeria*, and *Pseudolarix*), united by the presence of a single central resin canal in the taproot primary vascular region, and Pinées (subfamily Pinoideae: *Pinus*, *Picea*, *Larix*, and *Pseudotsuga*), united by resin canals positioned "adjacent to each protoxylem pole." *Cathaya* and *Nothotsuga* were described after Van Tieghem's classification (Chun and Kuang 1958; Page 1989); *Cathaya* has pinoid root anatomy (Hu and Wang 1984), but we have not located a description for the root anatomy of *Nothotsuga*. Wood resin canal characters support these groupings, as does the presence of resin vesicles in the seed coats of all members of subfamily Abietoideae (Price et al. 1987; Wang et al. 2000; Liston et al. 2003) and the distinctive form of their ovulate cone scales. However, placement of the root between *Cedrus* and the rest of Pinaceae often has rendered subfamily Abietoideae paraphyletic to subfamily Pinoideae; this is the case in all parsimony trees presented here (fig. 1*A*, 1*C*, 1*E*) and in cp *matK*, mitochondrial *nad5*, and nuclear *4CL* data sets analyzed separately (Wang et al. 2000). In contrast, the subfamilies were

mutually monophyletic in Bayesian analyses of cpDNA alone (figs. 1*D*, 2) and of combined cpDNA and nonmolecular data (fig. 1*F*), a result recovered in two prior studies with reduced generic sampling (Prager et al. 1976; Price et al. 1987). Other positions for the root of Pinaceae were found with nonmolecular characters (Hart 1987), with cpDNA restriction sites (Tsumura et al. 1995), and with a combined analysis of *rbcL* and nuclear ITS regions (Vining 1999). Paraphyly of subfamily Abietoideae is supported by the observation that the simpler wood anatomical characters in subfamily Abietoideae are considered plesiomorphic (Greguss 1955); the upright ovulate cones and their disarticulation to release seeds, features also observed in Araucariaceae, might also be plesiomorphic.

Several other generic relationships in Pinaceae are uncertain. *Cedrus*, besides having been recovered as sister to the remaining Pinaceae or sister to subfamily Abietoideae, has also been recovered as sister to *Abies* and *Keteleeria* (Price et al. 1987) or sister to *Pseudolarix* (Smith and Stockey 2001, 2002). In the three-genome data set of Wang et al. (2000), *Pseudolarix* was the sister group either to *Tsuga* and *Nothotsuga* (*matK* and *nad5*) or to *Tsuga*, *Nothotsuga*, *Keteleeria*, and *Abies* (*4CL*). The sister position of *Pseudolarix* spp. (with *Pityostrobus californiensis* and *Pityostrobus corneti*) to *Tsuga*, *Nothotsuga*, *Keteleeria*, *Abies*, and subfamily Pinoideae in the nonmolecular parsimony tree (fig. 1*A*) is consistent with the nuclear result of Wang et al. (2000) and better accommodates the older fossil record of *Pseudolarix* with respect to other genera in Pinaceae (LePage and Basinger 1995). In subfamily Pinoideae, a sister relationship between *Picea* and *Cathaya* has been found repeatedly in DNA-based studies, but branch support is not high (Wang et al. 2000; Eckert and Hall 2006; fig. 1*C*, 1*D*). Alternatively, *Cathaya* or *Picea* may be sister to *Pinus* (Hart 1987; Wang et al. 2000; fig. 1*A*, 1*B*, 1*E*, 1*F*).

The six fossil taxa analyzed in this study occupied three separate lineages within the Pinaceae crown group. *Pseudoaraucaria*, *Pityostrobus californiensis*, *Pityostrobus corneti*, and *Pseudolarix erensis* occurred within Abietoideae in the combined analyses (though at varying positions), and *Pityostrobus bernissartensis* and *Pinus* were monophyletic (fig. 1*E*, 1*F*). The position of *Pseudoaraucaria*, as either unresolved within subfamily Abietoideae in the Bayesian tree (fig. 1*F*) or sister to all members of Pinaceae except *Cedrus* in the parsimony tree (fig. 1*E*), is similar to that found by Alvin (1988) with a phenetic analysis of 22 characters, in which *Pseudoaraucaria* was the sister group to all extant members of Abietoideae. Whether *Pseudoaraucaria* had deciduous cone scales remains an open question, but we included two characters identified by Alvin (1988) referring to less developed, gelatinous "mechanical tissue" that could play a role in the shedding of ovulate cone scales and could point to a closer relationship between *Cedrus* and *Pseudoaraucaria* than was found by Smith and Stockey (2001, 2002).

Pityostrobus corneti was sister to *Pseudolarix* in the nonmolecular parsimony tree and the combined Bayesian tree (fig. 1*A*, 1*F*) but was sister to *Pityostrobus californiensis* and subfamily Pinoideae in the combined parsimony tree (fig. 1*E*). The taxon had been described previously as similar to two other members of subfamily Abietoideae, *Keteleeria* and *Cedrus* (Alvin 1953). *Pityostrobus californiensis* does not bear a strong resemblance to any particular extant genus (Smith and Stockey 2001), and although it seems to have an affinity for subfamily Abietoideae, it lacks resin vesicles in its seed coat and appears as sister to subfamily Pinoideae in the combined parsimony tree (fig. 1*E*).

Pseudolarix is well represented in the fossil record (LePage and Basinger 1995; LePage 2003). Inclusion of *Pseudolarix erensis* (*Pseudolarix* sp. sensu LePage and Basinger 1995) in a phylogenetic analysis supports both its interpretation as *Pseudolarix* (fig. 1) and its use as a Late Jurassic calibration point within crown Pinaceae (table 1; fig. 4). The interpretation of these fossils as *Pseudolarix* was described as "unequivocal" by LePage (2003). The high branch support for this relationship (fig. 1*E*, 1*F*) is intriguing but deceptive, because the anatomy of the fossil taxon is unknown; as a result, 78% of its cells were scored as missing in the nonmolecular matrix. Calibration based on the inferred age and phylogenetic position of *P. erensis* gives ages at most nodes that are much older than the appearance of ovulate cones attributed to extant genera in the fossil record but coincide well with recent findings of fossil leaves (Stockey and Wiebe 2006). Until detailed anatomical studies are published, these fossils must be treated with a greater degree of caution than anatomically characterized Pinaceae fossils from the Cretaceous.

The sister position between *Pityostrobus bernissartensis* and *Pinus* supports the assertion by Miller (1976) that a "strong *Pinus* influence" is apparent in *Pityostrobus*. However, the artificial genus also includes taxa more closely related to genera in subfamily Abietoideae, such as *Keteleeria*, *Abies*, and *Cedrus* (fig. 1; Alvin 1953; Smith and Stockey 2001, 2002). Exploratory analyses of the combined nonmolecular and molecular matrix expanded to include all fossil taxa scored by Smith and Stockey (2002) suggest that at least 12 of the 25 *Pityostrobus* species belong to subfamily Pinoideae. Simultaneous analysis of an expanded version of the combined matrix presented here should greatly improve our understanding of the phylogenetic position of anatomically characterized fossil cones with respect to extant taxa.

The total-evidence analyses resolved *P. belgica* as a crown-group member of *Pinus* (fig. 1*E*, 1*F*). The taxon was described from a single ovulate cone with anatomical and morphological synapomorphies for *Pinus*, including scales with a raised area (apophyses) in a dorsal position and a vascular trace system leading to an ovuliferous scale that conforms to the pattern defined for *Pinus* rather than to the fossil organ genus *Pityostrobus* (Miller 1976). It is similar to extant species, such as *Pinus sylvestris* of section *Pinus* (Alvin 1960). Some studies have treated *P. belgica* as the earliest reliable evidence that *Pinus* had diverged from other extant genera of Pinaceae (Wang et al. 2000), while others have used the fossil to calibrate an internal node of *Pinus*, either the divergence of the subgenera *Pinus* and *Strobus* or the divergence of sections *Pinus* and *Pinaster* (Eckert and Hall 2006; reviewed by Willyard et al. [2007]).

Although our understanding of nonmolecular character evolution within *Pinus* is limited, fossils attributed to infrageneric groups of *Pinus* have been important in shaping hypotheses regarding the age of Pinaceae. The oldest fossil member plausibly assignable to *Pinus* subgenus *Strobus* appears in the Late Cretaceous Aachen Formation of Belgium (Santonian, 83.5–85.8 Ma; Meijer 2000): fossil wood with pinoid cross-field pits and conspicuously dentate ray tracheids representing the organ genus *Pinuxylon*. This fossil was considered an upper bound for the

age of the *Pinus* subgeneric split in the rate estimations of Willyard et al. (2007). Paleobotanists have varied in their willingness to attribute this and other fossil taxa to extant infrageneric ranks in *Pinus*. Axelrod (1986) argued that six extant *Pinus* subsections (*Strobus*, *Cembroides*, *Balfourianae*, *Pinus*, *Ponderosae*, and *Contortae*) first appeared in North America during the Eocene (55.8–33.9 Ma). This controversial hypothesis has never been supported by an explicit phylogenetic analysis, but these ages for extant subsections were generally accepted by Millar (1993, 1998), interpreting fossil taxa as having affinities to extant infrageneric categories. The subsectional age estimates here also conflict with the assignment of Middle Eocene taxa *Pinus arnoldii*, *Pinus princetonensis*, and *Pinus baileyi* to subsection *Pinus* (Erwin and Schorn 2006). These taxa, like *P. belgica*, possess symplesiomorphic characters of *Pinus* and at present are best considered unassignable to extant subsections. Similarly, assignment of Oligocene *Pinus escalentis* to subsection *Contortae* (Erwin and Schorn 2006) is inconsistent even with the *Pseudolarix-Tsuga* calibration (fig. 4), but the taxon may suggest the ancestral morphology of North American hard pines. These conflicts highlight the importance not only of revising other North American Tertiary floras but also of including both fossil and extant taxa in phylogenetic analyses.

The calibrations proposed here recover an Early Jurassic or Early Cretaceous crown-group diversification of Pinaceae, with the final generic divergence between *Nothotsuga* and *Tsuga* occurring during the Eocene (fig. 3) or the Late Cretaceous (fig. 4). For both calibrations, the cpDNA silent-substitution rate estimates (range: 2.61×10^{-10} to 4.02×10^{-10} substitutions/site/year) are within the range reported by Willyard et al. (2007) from an independent calibration of the *Pinus* subgenus *Pinus-Strobus* divergence (2.2×10^{-10} to 4.2×10^{-10} substitutions/site/year). Calibrating the *Pinus-Picea* split at 123 Ma from the position of *Pityostrobus bernissartensis* and avoiding calibration with *Pinus belgica* yield a reasonable minimum bound for ages in the family (fig. 3). The *Pinus-Picea* calibration at 123 Ma is conservative because it avoids assigning ages on the basis of fossils with a more tentatively phylogenetic place or age, such as *Pityostrobus californiensis*, *Pseudolarix erensis*, and *Pinus belgica*. This calibration yields a crown-group age for *Pinus* of 72 Ma, which requires that *P. belgica* and *P. mutoi* (both dated at >89 Ma) be interpreted as stem- rather than crown-group members of the genus. In the *Pseudolarix-Tsuga* calibration, the estimated age of the subgenus *Pinus-Strobus* node was 87 Ma, slightly younger than the minimum age of the Wealden considered here (89.5 Ma), which indicates that using *P. belgica* for calibrating any crown node of *Pinus* would yield ages older than those suggested by other Pinaceae fossils.

Significant character incongruence was found between nonmolecular and molecular matrices in this study. Nevertheless, no incongruent clades received high levels of branch support (fig. 1). Other studies have suggested that the Templeton (1983) and ILD tests can detect significant congruence between different data partitions as a result of factors related more to signal heterogeneity than to divergent genealogies (Cunningham 1997; Barker and Lutzoni 2002; Hipp et al. 2004). In this study, combining nonmolecular and molecular data for Pinaceae caused a reduction in phylogenetic resolution compared to separate analyses. This observation, together with the lower levels of branch support in the nonmolecular than in the molecular trees, indicates that, beyond the need to acquire more DNA characters from all three genomes, there is a need for more-refined scoring and analysis of nonmolecular characters, as has been done in Cupressaceae (see Little 2006). Functional correlation is widely thought to exist among certain morphological characters scored separately, such as the presence of winged pollen and pollination drops, and between cone scale and seed morphology (Hart 1987; Tomlinson 1994). Furthermore, the ontogeny of highly homoplasious nonmolecular characters could be examined in more detail.

Acknowledgments

We thank Paul Spagna and Johan Yans for helpful comments on the age of the Belgian Wealden and Ruth Stockey and Gar Rothwell for comments on the fossil record of Pinaceae. Esmeralda Salgado Hernández provided laboratory assistance, and Manuel González Ledesma, Jorge Perez de la Rosa, Sol Ortiz García, Arturo Sánchez González, and Denisse Téllez Mazzocco assisted with field collecting. The following people and institutions generously donated samples: Adrian Quijada and Robert Bye (Universidad Nacional Autónoma de México), Martin Gardner, Fiona Inches, and Philip Thomas (Royal Botanic Garden, Edinburgh), Santiago C. González-Martínez (Spanish Forest Research Centre), Martin Scanlon (United States National Arboretum), Roman Businský (Silva Tarouca Research Institute for Landscape and Ornamental Gardening), Zhang Zhi-Yong (Chinese Academy of Sciences), Rich Cronn and John Syring (Oregon State University), and Amy Tuininga and Dennis Gray (Rutgers University). This study was financed by a grant to D. S. Gernandt from the Secretaria de Educación Pública-Promep and the Consejo Nacional de Ciencia y Tecnología (SEP-2003-C02-45164).

Literature Cited

Alvin KL 1953 Three abietaceous cones from the Wealden of Belgium. Mem Inst R Sci Nat Belg 125:1–42.

——— 1957 On the two cones *Pseudoaraucaria heeri* (Coemans) nov. comb. and *Pityostrobus villerotensis* nov. sp. from the Wealden of Belgium. Mem Inst R Sci Nat Belg 135:1–27.

——— 1960 Further conifers of the Pinaceae from the Wealden Formation of Belgium. Mem Inst R Sci Nat Belg 146:1–39.

——— 1988 On a new specimen of *Pseudoaraucaria major* Fliche (Pinaceae) from the Cretaceous of the Isle of Wight. Bot J Linn Soc 97:159–170.

Axelrod DI 1976 History of the coniferous forests, California and Nevada. Univ Calif Publ Bot 70:1–62.

——— 1986 Cenozoic history of some western American pines. Ann Mo Bot Gard 73:565–641.

Barker FK, FM Lutzoni 2002 The utility of the incongruence length difference test. Syst Biol 51:625–637.

Biswas C, BM Johri 1997 The gymnosperms. Springer, Berlin.

Bowe LM, G Coat, CW 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.

Brunsfeld SJ, PS Soltis, DE Soltis, PA Gedek, CJ Quinn, DD Strenge, TA Ranker 1994 Phylogenetic relationships among the genera of Tax-

odiaceae and Cupressaceae: evidence from *rbcL* sequences. Syst Bot 19: 253–262.

Chaw S-M, A Zharkikh, H-M Sung, T-C Lau, W-H Li 1997 Molecular phylogeny of extant gymnosperms and seed plant evolution: analysis of nuclear 18S rRNA sequences. Mol Biol Evol 14:56–68.

Chun WH, KZ Kuang 1958 Genus novum Pinacearum ex sina australi et occidentali. Bot Zh 43:461–476.

Conran JG, GM Wood, PG Martin, JM Dowd, CJ Quinn, PA Gadek, RA Price 2000 Generic relationships within and between the gymnosperm families Podocarpaceae and Phyllocladaceae based on an analysis of the chloroplast gene *rbcL*. Aust J Bot 48:715–724.

Cunningham CW 1997 Can three incongruence tests predict when data should be combined? Mol Biol Evol 14:733–740.

Dejax J, D Pons, J Yans 2007 Palynology of the dinosaur-bearing Wealden facies in the natural pit of Bernissart (Belgium). Rev Palaeobot Palynol 144:25–38.

Doyle JJ, JL Doyle 1987 A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15.

Eckert AJ, BD Hall 2006 Phylogeny, historical biogeography, and patterns of diversification for *Pinus* (Pinaceae): phylogenetic tests of fossil-based hypotheses. Mol Phylogenet Evol 40:166–182.

Erwin DM, HE Schorn 2005 Revision of the conifers from the Eocene Thunder Mountain flora, Idaho, U.S.A. Rev Palaeobot Palynol 137:125–145.

——— 2006 *Pinus baileyi* (section *Pinus*, Pinaceae) from the Paleogene of Idaho, USA. Am J Bot 93:197–205.

Farjon A 1990 Pinaceae: drawings and descriptions of the genera *Abies*, *Cedrus*, *Pseudolarix*, *Keteleeria*, *Nothotsuga*, *Tsuga*, *Cathaya*, *Pseudotsuga*, *Larix* and *Picea*. Költz Scientific, Königstein.

——— 2001 World checklist and bibliography of conifers. 2nd ed. Royal Botanic Gardens, Kew.

Farris JS, M Källersjö, AG Kluge, C Bult 1995 Testing significance of incongruence. Cladistics 10:315–319.

Frankis MP 1989 Generic inter-relationships in Pinaceae. Notes R Bot Gard Edinb 45:527–548.

Gadek PA, DL Alpers, MM Heslewood, CJ Quinn 2000 Relationships within Cupressaceae sensu lato: a combined morphological and molecular approach. Am J Bot 87:1044–1057.

Gernandt DS, G Geada López, S Ortiz García, A Liston 2005 Phylogeny and classification of *Pinus*. Taxon 54:29–42.

Gradstein FM, JG Ogg 2004 Geologic time scale 2004: why, how, and where next! Lethaia 37:175–181.

Greguss P 1955 Identification of living gymnosperms on the basis of xylotomy. Académiai Kiadó, Budapest.

——— 1972 Xylotomy of living conifers. Académiai Kiadó, Budapest.

Gugerli F, C Sperisen, U Buchler, I Brunner, S Brodbeck, JD Palmer, 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.

Hall TA 1999 BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98.

Hart JA 1987 A cladistic analysis of conifers: preliminary results. J Arnold Arbor Harv Univ 68:269–307.

Hilton J, RM Bateman 2006 Pteridosperms are the backbone of seed-plant phylogeny. J Torrey Bot Soc 133:119–168.

Hipkins VD, CH Tsai, SH 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.

Hipp AL, JC Hall, KJ Sytsma 2004 Congruence versus phylogenetic accuracy: revisiting the incongruence length difference test. Syst Biol 53:81–89.

Hu Y-S, K Napp-Zinn, D Winne 1989 Comparative anatomy of seed-scales of female cones of Pinaceae. Bot Jahrb Syst 111:63–85.

Hu Y-S, FH Wang 1984 Anatomical studies of *Cathaya* (Pinaceae). Am J Bot 71:727–735.

Huelsenbeck JP, FR Ronquist 2001 MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Keller AM, MS Hendrix 1997 Paleoclimatologic analysis of a Late Jurassic petrified forest, southeastern Mongolia. Palaios 12:282–291.

Krassilov V 1982 Early Cretaceous flora of Mongolia. Palaeontogr Abt B Palaeophytol 181:1–43.

Kusumi J, Y Tsumura, H Yoshimaru, H Tachida 2000 Phylogenetic relationships in Taxodiaceae and Cupressaceae sensu stricto based on *matK* gene, *chlL* gene, *trnL-trnF* IGS region, and *trnL* intron sequences. Am J Bot 87:1480–1488.

Leebens-Mack J, LA Raubeson, L Cui, JV Kuehl, MH Fourcade, TW Chumley, JL Boore, RK Jansen, CW dePamphilis 2005 Identifying the basal angiosperm node in chloroplast genome phylogenies: sampling one's way out of the Felsenstein Zone. Mol Biol Evol 22: 1948–1963.

LePage BA 2003 The evolution, biogeography and palaeoecology of the Pinaceae based on fossil and extant representatives. Acta Hortic 615:29–52.

LePage BA, JF Basinger 1991 A new species of *Larix* (Pinaceae) from the early Tertiary of Axel Heiberg Island, Arctic Canada. Rev Palaeobot Palynol 70:89–111.

——— 1995 Evolutionary history of the genus *Pseudolarix* Gordon (Pinaceae). Int J Plant Sci 156:910–950.

Lewis PO 2001 A likelihood approach for estimating phylogeny from discrete morphological character data. Syst Biol 50:913–925.

Lin J, Y Hu, X He, R Ceulemans 2002 Systematic survey of resin canals in Pinaceae. Belg J Bot 135:3–14.

Lin J, Y Hu, FH Wang 1995 Wood and bark anatomy of *Nothotsuga*. Ann Mo Bot Gard 82:603–609.

Liston A, DS Gernandt, TF Vining, CS Campbell, D Piñero 2003 Molecular phylogeny of Pinaceae and *Pinus*. Acta Hortic 615:107–114.

Little DP 2006 Evolution and circumscription of the true cypresses (Cupressaceae: *Cupressus*). Syst Bot 31:461–480.

Little EL, Jr, WB Critchfield 1969 Subdivisions of the genus *Pinus* (pines). USDA Forest Service Miscellaneous Publication 1144. Washington, DC.

Maddison DR, WP Maddison 2000 MacClade, version 4: analysis of phylogeny and character evolution. Sinauer, Sunderland, MA.

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

Mason-Gamer RJ, EA Kellogg 1996 Testing for phylogenetic conflict among molecular data sets in the tribe Triticeae (Gramineae). Syst Biol 45:524–545.

Meijer JJF 2000 Fossil woods from the Late Cretaceous Aachen Formation. Rev Palaeobot Palynol 112:297–336.

Millar CI 1993 Impact of the Eocene on the evolution of *Pinus* L. Ann Mo Bot Gard 80:471–498.

——— 1998 Early evolution of pines. Pages 69–91 *in* DM Richardson, ed. Ecology and biogeography of *Pinus*. Cambridge University Press, Cambridge.

Miller CN, Jr 1976 Early evolution in the Pinaceae. Rev Palaeobot Palynol 21:101–117.

——— 1988 The origin of modern conifer families. Pages 448–486 *in* CB Beck, ed. Origin and evolution of gymnosperms. Columbia University Press, New York.

——— 1992 Structurally preserved cones of *Pinus* from the Neogene of Idaho and Oregon. Int J Plant Sci 153:147–154.

Napp-Zinn K, YS Hu 1989 Anatomical studies on the bracts in pinaceous female cones. III. Comparative study of (mostly Chinese) representatives of all genera. Bot Jahrb Syst 110:461–477.

Nylander JAA, F Ronquist, JP Huelsenbeck, JL Nieves-Aldrey 2004 Bayesian phylogenetic analysis of combined data. Syst Biol 53:47–67.

Page CN 1989 New and maintained genera in the conifer families Podocarpaceae and Pinaceae. Notes R Bot Gard Edinb 45:377–395.

Posada D, TR Buckley 2004 Model selection and model averaging in phylogenetics: advantages of Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. Syst Biol 53:793–808.

Posada D, KA Crandall 1998 Modeltest: testing the model of DNA substitution. Bioinformatics 14:817–818.

Prager EM, DP Fowler, AC Wilson 1976 Rates of evolution in conifers (Pinaceae). Evolution 30:637–649.

Price RA 1989 The genera of Pinaceae in the southeastern United States. J Arnold Arbor Harv Univ 70:247–305.

Price RA, J Olsen-Stojkovich, JM Lowenstein 1987 Relationships among the genera of Pinaceae: an immunological comparison. Syst Bot 12:91–97.

Quinif Y, H Meon, J Yans 2006 Nature and dating of karstic filling in the Hainaut Province (Belgium): karstic, geodynamic and paleogeographic implications. Geodin Acta 19:73–85.

Quinn CJ, RA Price, PA Gadek 2002 Familial concepts and relationships in the conifers based on *rbcL* and *matK* sequence comparisons. Kew Bull 57:513–531.

Rai HS, PA Reeves, R Peakall, RG Olmstead, SW Graham 2008 Inference of higher-order conifer relationships from a multi-locus plastid data set. Botany 86:658–669.

Robaszynski F, AV Dhondt, JWM Jagt 2001 Cretaceous lithostratigraphic units (Belgium). Geol Belg 4:121–134.

Ronquist FR, JP Huelsenbeck 2003 MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574.

Rothwell GW, G Mapes, RH Mapes 1997 Late Paleozoic conifers of North America: structure, diversity and occurrences. Rev Palaeobot Palynol 95:95–113.

Rydin C, M Källersjö, EM 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.

Saiki K 1996 *Pinus mutoi* (Pinaceae), a new species of permineralized seed cone from the upper Cretaceous of Hokkaido, Japan. Am J Bot 83:1630–1636.

Sanderson MJ 2002 Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol Biol Evol 19:101–109.

Schorn HE 1994 A preliminary discussion of fossil larches (*Larix*, Pinaceae) from the Arctic. Quat Intl 22/23:173–183.

Seward AC 1900 La flore wealdienne de Bernissart. Mem Mus R Hist Nat Belg 1:1–37.

Smith SY, RA Stockey 2001 A new species of *Pityostrobus* from the lower Cretaceous of California and its bearing on the evolution of Pinaceae. Int J Plant Sci 162:669–681.

——— 2002 Permineralized pine cones from the Cretaceous of Vancouver Island, British Columbia. Int J Plant Sci 163:185–196.

Stefanovic S, M Jager, J Deutsch, J Broutin, M Masselot 1998 Phylogenetic relationships of conifers inferred from partial 28S rRNA gene sequences. Am J Bot 85:688–697.

Stockey RA, NJP Wiebe 2006 Mosaic evolution of Early Cretaceous Pinaceae and the origins of extant genera. Botany 2006 abstract 569, http://www.2006.botanyconference.org/engine/search/index.php?func=detail&aid=569.

Swofford DL 2003 PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer, Sunderland, MA.

Takaso T, JN Owens 1995 Ovulate cone morphology and pollination in *Pseudotsuga* and *Cedrus*. Int J Plant Sci 156:630–639.

Templeton AR 1983 Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. Evolution 37:221–244.

Thompson JD, DG Higgins, TJ Gibson 1994 CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680.

Tomlinson PB 1994 Functional morphology of saccate pollen in conifers with special reference to Podocarpaceae. Int J Plant Sci 155:699–715.

Tsumura Y, K Yoshimura, N Tomaru, K Ohba 1995 Molecular phylogeny of conifers using RFLP analysis of PCR-amplified specific chloroplast genes. Theor Appl Genet 91:1222–1236.

Van Tieghem P 1891 Structure et affinités des *Abies* et des genres les plus voisins. Bull Soc Bot Fr 38:406–415.

Vining TF 1999 Molecular phylogenetics of Pinaceae. PhD diss. University of Maine, Orono.

Wakasugi T, J Tsudzuki, S Ito, K Nakashima, T Tsudzuki, M Sugiura 1994 Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. Proc Natl Acad Sci USA 91:9794–9798.

Wang X-Q, DC Tank, T Sang 2000 Phylogeny and divergence times in Pinaceae: evidence from three genomes. Mol Biol Evol 17:773–781.

Wang X-R, Y Tsumura, H Yoshimaru, K Nagasaka, AE Szmidt 1999 Phylogenetic relationships of Eurasian pines (*Pinus*, Pinaceae) based on chloroplast *rbcL*, *matK*, *rpl20-rps18* spacer, and *trnV* intron sequences. Am J Bot 86:1742–1753.

Willyard A, J Syring, DS Gernandt, A Liston, R Cronn 2007 Fossil calibration of molecular divergence infers a moderate mutation rate and recent radiations for *Pinus*. Mol Biol Evol 24:90–101.

Wolfe JA, HE Schorn 1990 Taxonomic revision of the Spermatopsida of the Oligocene Creede flora, southern Colorado. U.S. Geological Survey Bulletin 1923. Department of the Interior, Washington, DC.

Wu H, Z Hu 1997 Comparative anatomy of resin ducts of the Pinaceae. Trees 11:135–143.

Yans J, J Dejax, D Pons, C Dupuis, P Taquet 2005 Implications paléontologiques et géodynamiques de la datation palynologique des sédiments à faciès wealdien de Bernissart (bassin de Mons, Belgique). C R Palevol 4:135–150.