American Journal of Botany 85(5): 688–697. 1998.

PHYLOGENETIC RELATIONSHIPS OF CONIFERS INFERRED FROM PARTIAL 28S rRNA GENE SEQUENCES[1]

SAŠA STEFANOVIĆ,[2,6] MURIEL JAGER,[2] JEAN DEUTSCH,[2,3] JEAN BROUTIN,[4] AND MONIQUE MASSELOT[2,5]

[2]Service Commun de Bio-Systématique, Université Pierre et Marie Curie, 9, quai St-Bernard, case 241, 75252 Paris, France;
[3]Groupe Développement et Évolution, CNRS URA 1135 et Université Pierre et Marie Curie, 9, quai St-Bernard, case 241, 75252 Paris, France; and
[4]Laboratoire de Paléobotanique et Paléoécologie, Université Pierre et Marie Curie, 12, rue Cuvier, 75005 Paris, France

The conifers, which traditionally comprise seven families, are the largest and most diverse group of living gymnosperms. Efforts to systematize this diversity without a cladistic phylogenetic framework have often resulted in the segregation of certain genera and/or families from the conifers. In order to understand better the relationships between the families, we performed cladistic analyses using a new data set obtained from 28S rRNA gene sequences. These analyses strongly support the monophyly of conifers including Taxaceae. Within the conifers, the Pinaceae are the first to diverge, being the sister group of the rest of conifers. A recently discovered Australian genus *Wollemia* is confirmed to be a natural member of the Araucariaceae. The Taxaceae are nested within the conifer clade, being the most closely related to the Cephalotaxaceae. The Taxodiaceae and Cupressaceae together form a monophyletic group. *Sciadopitys* should be considered as constituting a separate family. These relationships are consistent with previous cladistic analyses of morphological and molecular (18S rRNA, *rbcL*) data. Furthermore, the well-supported clade linking the Araucariaceae and Podocarpaceae, which has not been previously reported, suggests that the common ancestor of these families, both having the greatest diversity in the Southern Hemisphere, inhabited Gondwanaland.

Key words: cladistic analysis; conifers; *Ginkgo*; phylogeny; rRNA (28S) gene sequences.

The conifers, a group that is widely distributed throughout the world, comprise more than 600 species grouped within 60–65 genera. They are the largest and the most diverse group of living gymnosperms. Known with certitude from fossil register as far back as upper Carboniferous, they were the most abundant on the Jurassic-Cretaceous limit. The conifers have attracted much systematic attention because of their importance for seed plant phylogeny and their substantial role in many of the earth's biota.

The range of morphological variation found in conifers is the greatest of all extant gymnosperms. Previous efforts to accommodate this diversity have often been implemented without a phylogenetic framework. According to the commonly accepted precladistic view, the conifers are divided into seven families: Pinaceae, Podocarpaceae, Araucariaceae, Taxaceae, Cephalotaxaceae, Taxodiaceae, and Cupressaceae (Pilger, 1926). Nevertheless, it has been doubted that they form a natural group and other classifications have also been proposed. The phylogenetic relationships between these families (and the genera they include) were the focus of numerous debates, but they have frequently been based on very few characters considered as ''the most important'' by the author. The structure and the position of the ovule, as well as the leaf form and venation pattern, for example, have given rise to several different interpretations.

The phylogenetic position of Taxaceae is one of the oldest unsolved problems in gymnosperm systematics. Members of this family are unique because they are devoid of the ''classical'' cone that characterizes the majority of conifers. Basing his conclusions on the study of the ovule structure of *Taxus, Torreya,* and *Cephalotaxus*, Sahni (1920) claimed that these genera, like *Ginkgo,* are direct descendants of the extinct Cordaitales and proposed placing them in a distinct group, Taxales. Florin (1948, 1951) estimated that only Taxaceae showed the specific character, a simple, uniaxial cone, in contrast to the compound and biaxial one found in other conifers. Therefore, he segregated Taxaceae in the distinct order Taxales as a parallel evolutionary lineage with Ginkgoales, Cordaitales, and Coniferales (Florin, 1951). According to his interpretation, the terminal position of the uniovulate seed is a primitive feature that can be traced down back to the Devonian, linking this group to the extinct Psilophytales.

Conversely, Chamberlain (1935), Takhtajan (1953), and Harris (1976) suggested that the simple, terminal-ovule strobile was derived from the compound, biaxial one by its reduction followed by a shift from the lateral to the apical position. A recent cladistic analysis based on 18S rRNA data (Chaw et al., 1993) as well as that based on chloroplast DNA structural mutations (Raubeson and Jansen, 1992) strongly supported this view.

Buchholz (1934) considered the conifers to be a natural group, but he made a distinction between those families

[1] Manuscript received 19 March 1997; revision accepted 14 August 1997.

The authors thank M-C. Boisselier for allowing us to use her unpublished sequence; G. Callen, Chèvreloup Arboretum, MNHN, T. Jaffré, Centre ORSTOM, Nouméa, Nouvelle-Calédonie, and Ph. Morat, the curator of P for supplying plant material; and Denise Pons for critical comments on the manuscript. This work was supported by the Reseau National de Biosystématique (ACC SV-7). SS was supported in part by Open Society Institute SGP FR-134-96/97YU.

[5] Author for correspondence (e-mail: Monique.Masselot@snv.jussieu.fr).

[6] Current address: Department of Botany, University of Washington, Seattle, Washington 98195-5325.

presenting an evident ovulate cone: Cupressaceae, Taxodiaceae, Araucariaceae, and Pinaceae (suborder Phanerostrobilares) and those without it: Podocarpaceae, Taxaceae, and Cephalotaxace (suborder Aphanostrobilares). This classification was subsequently accepted by many authors (e.g., Chamberlain, 1935; Li, 1953).

The form of the leaves was also used to question the monophyly of the conifers. The leaves of cycads, Cordaitales as well as those of *Ginkgo,* are broad and multiveined. On the other hand, conifer leaves are usually smaller, narrow, needle- or scale-like, and single-veined. Certain genera of Podocarpaceae and the Araucariaceae, however, also have broad and multiveined leaves without any midrib. On the basis of these features Fu (1992) separated the genus *Nageia* sensu stricto (s.s) from the Podocarpaceae and associated it with the Araucariaceae, the Ephedraceae, the Welwitschiaceae, and the Ginkgoaceae in an evolutionary lineage characterized by multiveined leaves, whose origin he sees in the Paleozoic seed plants Cordaitales.

Within the conifers, the monophyly of the Podocarpaceae and the Taxodiaceae was frequently disputed. The genera regrouped in these families show a wider range of morphological variation than any other conifer family, which led some authors to segregate them into several distinct families. For example, Keng (1974, 1978) isolated *Phyllocladus,* a genus with reduced leaves and photosynthetic cladodes, in the monogeneric family of the Phyllocladaceae. Traditional recognition of separate Cupressaceae and Taxodiaceae has been recently questioned on the basis of morphological (Eckenwalder, 1976; Hart, 1987) and molecular data (Price et al., 1993; Brunsfeld et al., 1994). The genus *Sciadopitys,* usually associated with Taxodiaceae, is highly divergent in its leaf and short shoot morphology and does not share any obvious synapomorphies with the other genera of that family (Eckenwalder, 1976; Price and Lowenstein, 1989). For these reasons, it was separated in a distinct family, the Sciadopityaceae (Hayata, 1931).

The distinction between primitive and derived characters is one of the critical problems in phylogenetic reconstructions and highlights the danger of using one single character for phylogenetic purposes. By employing the principle of parsimony, it is only a posteriori, in the interaction with other data that the preponderance of characters sorts out their phylogenetic implications and enables us to interpret certain characters as being autapomorphic or homoplasic. Several modern cladistic analyses (e.g., Crane, 1985; Doyle and Donoghue, 1986, 1987; Loconte and Stevenson, 1990; Nixon et al., 1994; Rothwell and Serbet, 1994; Doyle, 1996) were conducted on large data sets relevant to seed plant relationships. The majority of these analyses, however, were simplified in the choice of modern conifers by omitting one or more groups from their analysis or by condensing all conifers in a single terminal taxon. Consequently, a certain number of potentially informative characters within conifers were omitted, being autapomorphic on the higher level.

A recent cladistic analysis of living conifers based on morphology (Hart, 1987) provided an initial hypothesis for the relationships within the conifers. Although it was the most complete (63 genera and 123 characters), this study lacked certain available data and offered only preliminary results.

An additional question requires some consideration. There is no clear consensus on the relationships between the conifers and the rest of the main seed plant groups despite the substantial importance of choosing appropriate outgroups. The abovementioned cladistic studies based on morphological characters and comprising both extinct and extant taxa of spermatophytes gave varying and at least partially inconsistent results at this level. In the resulting cladograms conifers are frequently found as a sister group to Anthophytes, alone (Nixon et al., 1994; Rothwell and Serbet, 1994) or together with *Ginkgo* (Crane, 1985). The same arrangement, grouping conifers plus ginkgos (coniferopsids), is found by Doyle and Donoghue (1986, 1987) and Doyle (1996) but with cycads in the position of sister group to the anthophytes. The analysis of Loconte and Stevenson (1990), based exclusively on extant taxa, yielded a tree in which conifers are linked with anthophytes, rather than with *Ginkgo.* The rRNA study of living spermatophytes (Hamby and Zimmer, 1992) contradicts all these results by linking conifers with cycads.

Current knowledge of conifers and the rest of the spermatophytes is based on more than a century of work in morphology, anatomy, embryology, cytology, phytochemistry, and paleobotany. In spite of this important database, proposed evolutionary relationships between and within major spermatophyte lineages remain controversial. Molecular data have the advantage of being independent of the different interpretations of the morphological characters. Therefore, we introduce here a new data set, DNA sequences from the nuclear-encoded 28S rRNA gene, in order to resolve some of long-standing questions: (1) the phylogenetic position of conifers within the orders of spermatophytes; (2) the monophyly of conifers; and (3) the relationships between the families and genera that constitute the conifers.

MATERIALS AND METHODS

The taxa used in our analysis are listed in Table 1, with information on sources of plant material or GenBank accession numbers for published DNA sequences. The 28 conifers on which our analyses are principally based include genera from all traditionally recognized families. We also included eight gymnosperm taxa, representing *Ginkgo* and cycads and eight Anthophytes taxa, representing Gnetales and angiosperms. As outgroups we used *Marchantia,* a nonvascular land plant, and two seedless tracheophytes, *Equisetum* and *Polypodium.*

Fresh or dried leaves (1–20 g) were washed in distilled water and in 50 mmol/L Tris-HCl and then ground in liquid nitrogen. Total DNA was isolated from the ground tissue by the modified CTAB method (Rogers and Bendich, 1985; Doyle and Doyle, 1987). The 28S rRNA gene was amplified from genomic DNA by polymerase chain reaction (PCR) (Mullis and Faloona, 1987; Saiki et al., 1988), using either Biotaq or Hitaq (Bioprobe) polymerases. The amplified sequences correspond to the regions C1 (partially), D1, C2, D2 (completely), and C3 (partially) of the *Mus musculus* gene (Hassouna, Michot, and Bachellerie, 1984). This gene fragment was amplified using six pairs of forward and reverse primers listed in Table 2. Their relative positions are shown in Fig. 1. This procedure yielded double-stranded segments of approximately: 650 bp (2617–2618), 700 bp (C1′–D2), 300 bp (2618–E1), 350 bp (F1–R1), 350 bp (C1′–E1), and 300 bp (F1–D2). All amplification products were directly sequenced on both strands using the Thermo Sequenase™ (Amersham) sequencing kit by the dideoxy chain termination technique (Sanger, Nicklen, and Coulson, 1977), with ^{33}P-dATP/dCTP. Sequencing products were resolved on 6% polyacrylamide

TABLE 1. List of analyzed species. Source information and GenBank accession numbers are given for sequences new to this study as well as GenBank accession numbers for previously published sequences. All voucher specimens cited are deposited at the Service Commun de Bio-Systématique, Université Pierre et Marie Curie, Paris.

Species	Source
Ginkgoales	
Ginkgo biloba L.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90672
Cycadales	
Cycas revoluta Thunb.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90673
Cycas rumphii Miq.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90674
Stangeria eriopus Nash	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90675
Dioon edule Lindl.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90676
Zamia furfuracea Ait.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90677
Encephalartos lebomboensis I. Verd.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90678
Encephalartos laevifolius Stapf & Burtt-Davy	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90679
Coniferales	
Pinus nigra Arn.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90680
Pinus cembra L.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90681
Pseudotsuga menziesii (Mirb.) Franco	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90682
Abies grandis Lindl.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90683
Cedrus deodara Loud.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90684
Podocarpus macrophyllus (Thunb.) D. Don	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90685
Retrophyllum minor (Larr.) C. N. Page	Centre ORSTOM, Nouméa, Nouvelle-Calédonie; U90686
Falcatifolium taxoides (Brongn. & Gris) de Laub.	Centre ORSTOM, Nouméa, Nouvelle-Calédonie; U90687
Acmopyle pancheri (Brongn. & Gris) Pilg.	coll. Veillon 6156; MNHN Herbaria, P, France; U90688
Araucaria heterophylla (Salisb.) Franco	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90689
Araucaria araucana (Molina) K. Koch	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90690
Araucaria angustifolia (Bertol.) Kuntze	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90691
Agathis australis (D. Don) Salisb.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90692
Agathis palmerstoni F. v. Muell.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90693
Wollemia nobilis W. G. Jones, K. D. Hill & M. J. Allen	coll. Jones 362731; MNHN Herbaria, P, France; U90694
Taxus baccata L.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90695
Torreya grandis Fortune	coll. Cheng 2463; MNHN Herbaria, P, France; U90696
Cephalotaxus harringtonia (Forkes) K. Koch.	Cultivated; The Botanical Garden, MNHN, Paris, France; U90697
Sciadopitys verticillata (Thunb.) Siebold & Zucc.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90698
Cunninghamia lanceolata (Lamb.) Hook.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90699
Taiwania cryptomeroides Hayata	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90700
Sequoia sempervirens (D. Don) Endl.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90701
Taxodium distichum (L.) Rich.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90702
Cryptomeria japonica D. Don	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90703
Chamaecyparis lawsoniana (Murr.) Parl.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90704
Juniperus communis L.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90705
Thuja orientalis L.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90706
Calocedrus decurrens (Torrey) Florin	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90707
Gnetales	
Welwitschia mirabilis Hook. f.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90708
Ephedra nebrodensis Tineo	Cultivated; The Botanical Garden, MNHN, Paris, France; U90709
Gnetum gemon L.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90710
Angiosperms	
Nymphea stellata F. v. Muell.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90711
Citrus limon Risso	X05910
Lycopersicon esculentum Mill.	X13557
Daucus carota L.	X17534
Oryza sativa L.	M11585
Outgroups	
Polypodium vulgare L.	Cultivated; Chèvreloup Arboretum, MNHN, Rocquencourt, France; U90712
Equisetum hyemale L.	Cultivated; The Botanical Garden, MNHN, Paris, France; U90713
Marchantia polymorpha L.	M-C. Boisselier, MNHN, Paris, France

sequencing gels, visualized by autoradiography, and read at least twice by two different readers.

A data set of 47 28S rDNA sequences (42 new for this study) was analyzed. Since the questions were addressed on two different phylogenetic levels, two separate phylogenetic analyses were conducted. The first analysis, designed to verify the monophyly of conifers and to find their closest relatives (i.e., appropriate outgroups), was conducted on a 34-sequence set, comprising all major orders of spermatophytes. Once the monophyly of conifers (including Taxaceae) had been determined, with *Ginkgo* and cycads as sister groups (see Results), a second analysis including 32 sequences was performed to elucidate the familial relationships of the conifers.

Sequences were registered and handled, from database to tree analysis, with the MUST package (Philippe, 1993). The alignment was carried out manually without taking into account the secondary structure. Since information was missing at the ends of the molecule in some

TABLE 2. List of primers used (for their relative position, see Fig. 1.).

Name	Primer sequence (5′ to 3′)
Forward primers	
2618	CATATCAATAAGCGGAGGAAAAGAAAC
F1	AAAGATGAAAAGGACTTTGAAAAGAGAGT
C1′	ACCCGCTGAATTTAAGCAT
Reverse primers	
2617	TTACTCACCCGTTGACTCGCACAC
E1	CTCTCTTTTCAAAGTCCTTTTCATCTTT
R1	CATGTTAGACTCCTTGGT
D2	TCCGTGTTTCAAGACGGG

taxa, 48 and 68 sites were excluded from the aligned sequences at the 5′ and 3′ terminus, respectively. Additionally, we eliminated 42 sites in regions that could not be unambiguously aligned across all taxa. This left a total of 638 sites. In the ordinal-level matrix there were 385 variable sites, of which 248 were informative. The family-level matrix contained 232 variable sites, of which 149 were informative. Only the sites informative for parsimony were used in all phylogenetic analyses.

In view of the large number of taxa under consideration in both studies it was not possible to use the exhaustive branch and bound search algorithm. Therefore, we performed the heuristic searches for most parsimonious trees using PAUP (Swofford, 1993) with MULPARS option and ACCTRAN optimization. All changes were weighted equally and all character state changes were unordered. In order to maximize the probability of discovering different islands of trees (Maddison, 1991) the analyses involved 100 replicates with stepwise random taxon addition and TBR branch swapping.

To infer the relative support for particular clades we performed both bootstrap and decay analyses. Bootstrap analyses (Felsenstein, 1985) using 250 replicates (each with five heuristic analyses with stepwise random addition of taxa and TBR branch swapping) were performed with PAUP. Decay analyses (Bremer, 1988; Donoghue et al., 1992) were done with PAUP by relaxing parsimony until a given clade collapsed. The decay index indicates the minimum number of additional steps needed for a branch to break down.

In order to detect possible homoplasies due to saturation in substitution we compared the distances computed from transversions alone to distances computed from transitions alone for each pair of taxa in data sets (MUST, COMP_MAT option). Each pair of species is characterized by a certain number of differences (transitions and transversions) and can be represented by one dot. In the resulting diagram the form of the cloud of points reflects its saturation; saturation is reached when the number of transitions remains constant while the number of transversions continues to increase (Philippe et al., 1994).

The left-hand skewness tests, i.e., g1 statistics (Hillis, 1991; Huelsenbeck, 1991) based on evaluation of the tree-length distribution of 10 000 randomly sampled trees (PAUP, RANDOM TREES option) were used as an indicator of nonrandom structure in the 28S rRNA data set. We evaluated the significance of g1 values using the tables of Hillis (1991).

Alternative topologies were assessed by implementing the TOPOLOGICAL CONSTRAINTS option of PAUP and the minimum number of steps required to produce the resulting topologies was recorded.

Tree statistics, such as consistency index (CI) (Kluge and Farris, 1969) and retention index (RI) (Farris, 1989), were also calculated with PAUP, as a measure of homoplasy.

RESULTS

Saturation—Homoplasy is not encountered solely in molecular data, but since each character has only five possible states it poses major problems in molecular phylogenetics. Assuming that molecular changes occur roughly in proportion to the time that has elapsed since divergence, after a certain period a saturation plateau is reached and no homology is left. Philippe et al. (1994) proposed to use the comparison between the phenetic distances calculated from transversions alone and the distances calculated from transitions alone as a method for detecting relative saturation. The results of these comparisons (MUST, COMP_MAT option) for ordinal- and family-level matrices are shown in Figs. 2 and 3.

For the ordinal-level data (Fig. 2) the dots were widely scattered, indicating a higher stochasticity of the data. We subsequently investigated the dot distribution for different pairs of taxa. Three subgroups were found: (1) a subgroup comprising anthophytes and/or outgroups; (2) a subgroup including cycads; and (3) a subgroup of conifers plus *Ginkgo biloba*. Transition and transversion distances clearly increase correlatively in the last subgroup only. Transition distances increase less than transversion distances for cycads and they do not increase at all in relation to transversions when anthophytes and/or outgroups are involved, indicating a high saturation of transitions between these taxa and the others. Consequently, the transversions only (124 informative sites) were used in further analyses concerning this data set. On the other hand, comparison of phenetic distance matrices for the family-level data set (Fig. 3) showed a better correlation between transitions and transversions. This enabled us to make use of all the differences for the phylogenetic analyses (149 informative sites).

Although a certain amount of phylogenetic information was lost due to the omission of the transitions from the ordinal-level data, the distribution of 10 000 random tree lengths was significantly left-skewed (Hillis, 1991). The values of g1 = −0.52 and −0.60 for the ordinal- and family-level analyses, respectively (with $P < 0.05$ in both cases), indicate the presence of substantial phylogenetic signals in both data matrices.

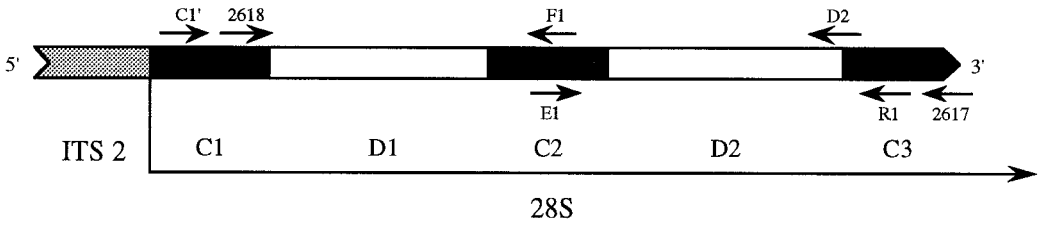

Fig. 1. Ribosome gene portion comprising the 3′ end of internal transcribed spacer 2 (ITS 2) and the 5′ end of the 28S rRNA gene showing regions sequenced in this study (C1 partially; D1, C2, D2 completely; C3 partially). The arrows refer to the primers used and show their relative locations. The different regions are designated according to Hassouna, Michot, and Bachellerie (1984).

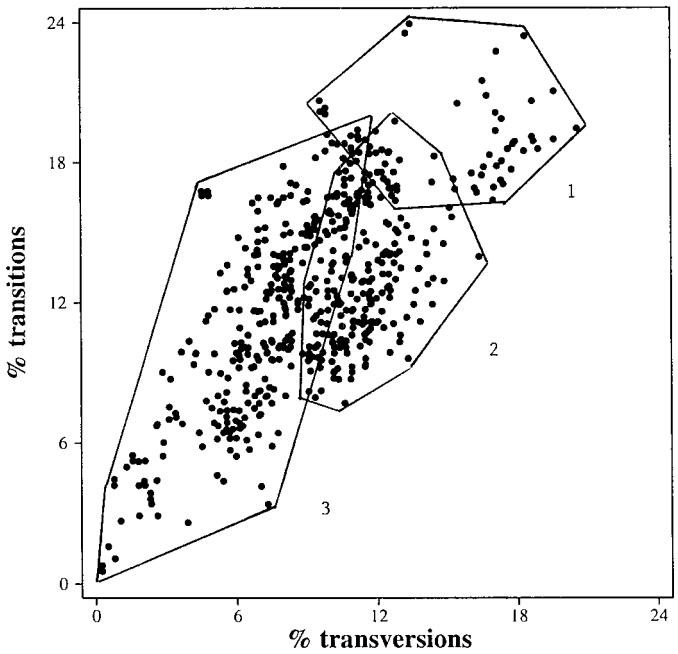

Fig. 2. Comparison of phenetic distances in the percentage of differences (MUST, COMP_MAT option), calculated from transversions alone (abscissa) and from transitions alone (ordinate) between pairs of sequences for ordinal-level data set (34 species). Dispersion of dots for which at least one pair member comes from: (1) anthophytes and/or outgroups; (2) cycads; and (3) conifers and/or *Ginkgo*. The dispersion indicates an important level of saturation in substitution. Correlation coefficient $r^2 = 0.63$.

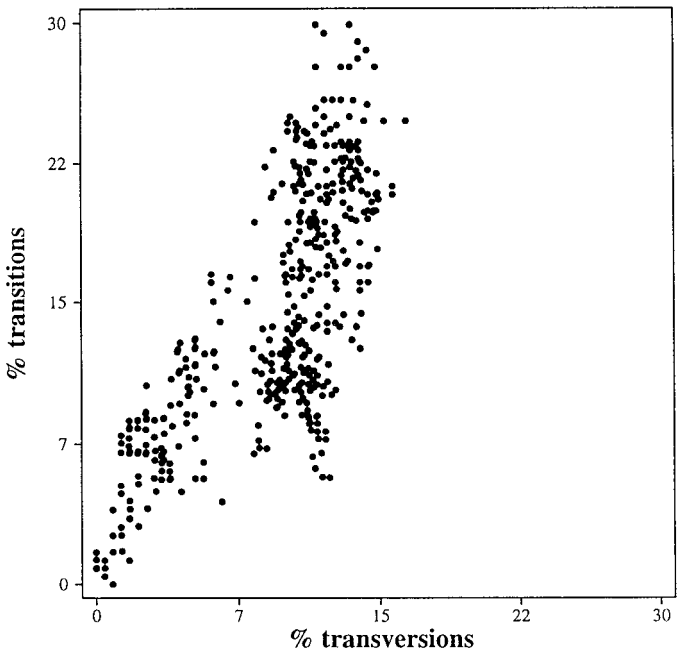

Fig. 3. Comparison of phenetic distances in the percentage of differences (MUST, COMP_MAT option), calculated from transversions alone (abscissa) and from transitions alone (ordinate) between pairs of sequences for family-level data set (32 species). Correlation coefficient $r^2 = 0.74$.

Unconstrained topologies—*Ordinal-level analysis*—The phylogenetic analysis of 34 sequences from conifers and other major spermatophyte groups resulted in eight most parsimonious trees, 323 steps each, all belonging to one island (Maddison, 1991). The shortest trees differed only in the resolution of one polytomy involving *Taxus, Cephalotaxus, Cunninghamia,* and *Sequoia.* The strict consensus of these trees is presented in Fig. 4 with decay values and bootstrap frequencies. The CI of 0.38 (RI = 0.65), excluding uninformative characters, reflects a somewhat higher level of homoplasy for a data matrix of 34 terminal taxa, but there is no statistically significant departure from the expected value of CI = 0.40, calculated by the formula of Sanderson and Donoghue (1989).

According to our results all conifer taxa (including *Taxus*) are found to be monophyletic as well as the cycads, the angiosperms, and the Gnetales. The taxa traditionally regarded as gymnosperms are polyphyletic. They are divided into two separate clades. One clade includes all members of Gnetales, with *Ephedra* as the sister group to [*Gnetum* plus *Welwitschia*]. The Gnetales and angiosperms are resolved as sister groups in the shortest trees, and together form the clade of Anthophytes. This placement, however, is only weakly supported (decay 1), most probably because of the relatively low number of sites. The other gymnosperm clade (indicated with an arrow in Fig. 4.) is supported by two steps of decay and regroups the cycads, *Ginkgo,* and the conifers. Within this clade, conifers and *Ginkgo* are more closely related to each other than either one is to the cycads. They form a monophyletic group, coniferopsids, but this assemblage of taxa is less well supported (one step of decay).

Family-level analysis—Taking the results of the ordinal-level analysis as a basis, we used *Ginkgo* and/or *Cycas* and *Dioon* as outgroup taxa for the family-level parsimony analysis. The phylogenetic analysis of 32 sequences from conifers and their putative closest relatives yielded six most parsimonious trees, 442 steps each, belonging to two different islands (Maddison, 1991). The topologies of the shortest trees from the two islands differed principally in the resolution of polytomies involving Taxaceae, Cephalotaxaceae, and Taxodiaceae (results not shown). In the first island, comprising two trees, Taxaceae were resolved as being monophyletic, with *Cephalotaxus* as their sister group. In the second island, comprising four trees, Taxaceae were found to be paraphyletic and *Cephalotaxus* was positioned as the sister group to the Cupressaceae sensu lato (s.l.). The strict consensus of all the most parsimonious trees is presented in Fig. 5 with indicated decay values and bootstrap frequencies. The CI of 0.53 (RI = 0.78) is significantly higher ($P < 0.01$) than the expected value of CI = 0.41 for 32 taxa, based on an analysis of published data sets (Sanderson and Donoghue, 1989).

These results confirm the monophyly of conifers, this time with higher decay values and bootstrap percentages (Fig. 5). Five families, the Pinaceae, the Podocarpaceae, the Araucariaceae, the Cephalotaxaceae, and the Cupressaceae, are found to be monophyletic. The Taxodiaceae (minus *Sciadopitys*) and Cupressaceae s.s. together form a well-supported clade and the genus *Sciadopitys* forms the monogeneric family. The Pinaceae form an isolated basal group positioned as a sister group to the remainder

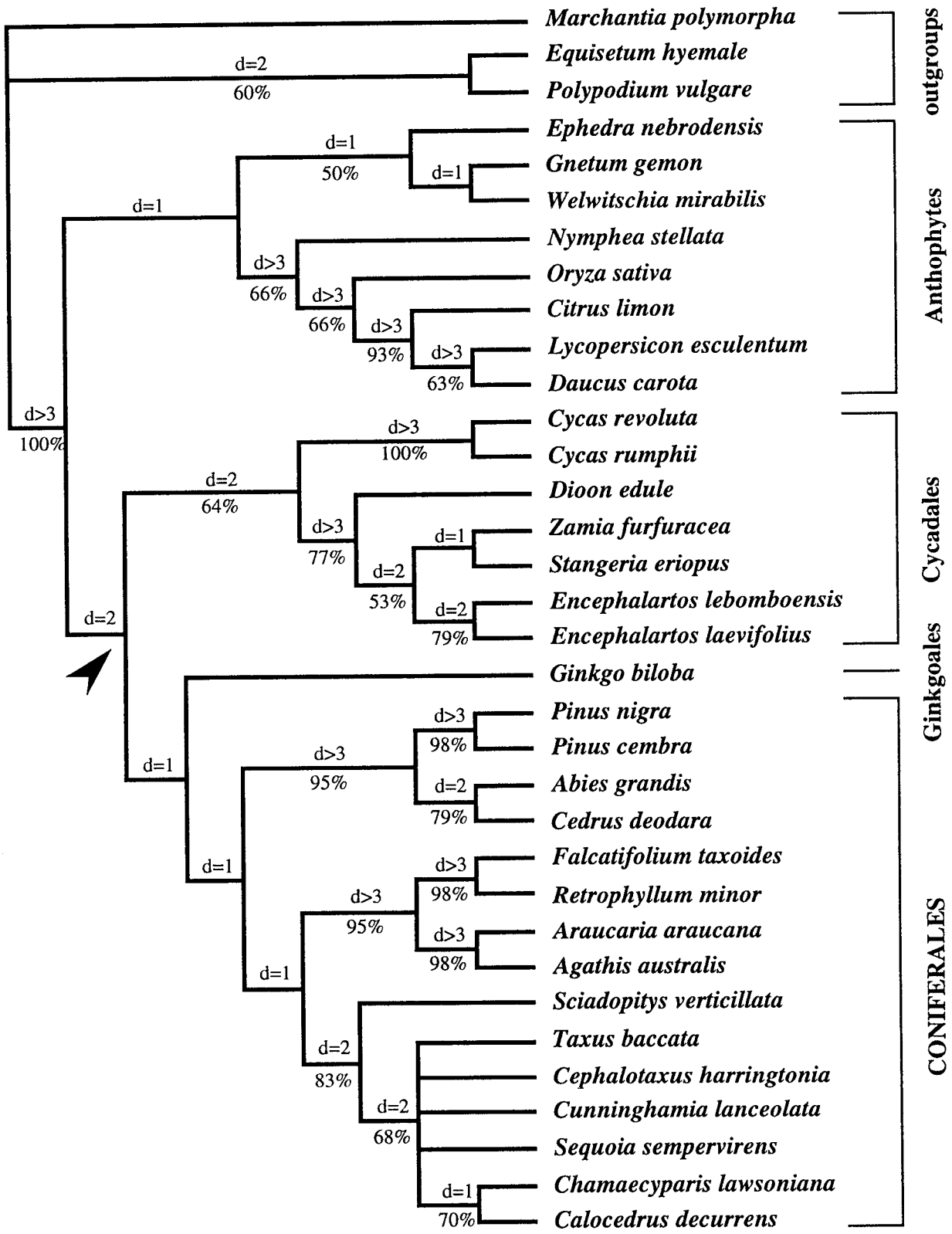

Fig. 4. Strict consensus of the eight most parsimonious trees from the ordinal-level analysis comprising 34 28S rRNA sequences of all major spermatophytes groups (L = 323; CI = 0.38; RI = 0.65). The tree is rooted using *Marchantia*, *Equisetum*, and *Polypodium* as outgroups. Only transversions are used (124 informative sites). Decay values (d) are indicated above each branch. Bootstrap percentages of nodes that occurred in ≥50% of 250 bootstrap replicates are indicated below branches. The arrow indicates the node of particular interest to this study (see text).

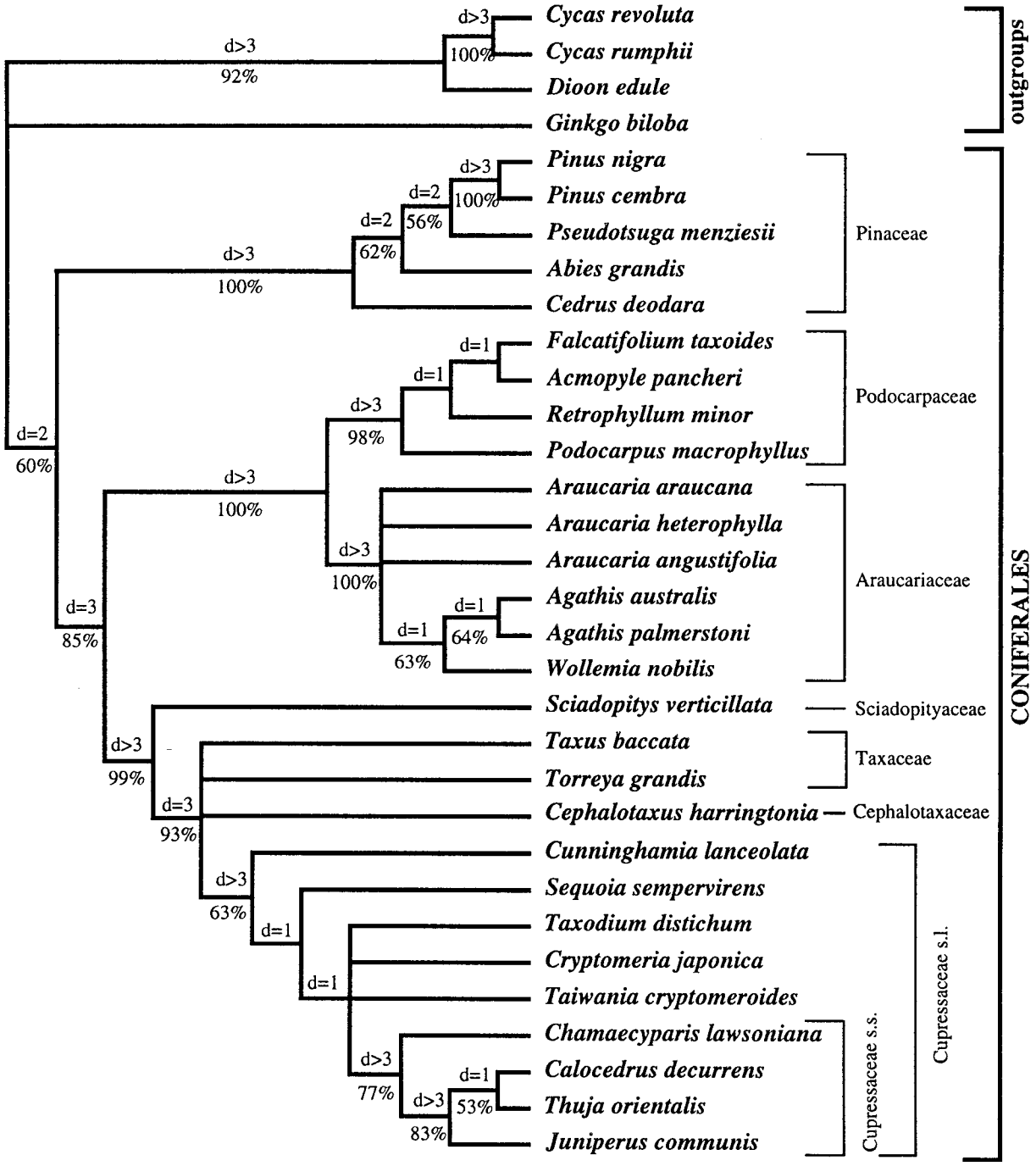

Fig. 5. Strict consensus of the six most parsimonious trees from the family-level analysis comprising 32 28S rRNA sequences belonging to all traditionally recognized conifer families (L = 442; CI = 0.53; RI = 0.78). The tree is rooted using *Ginkgo* and cycads as outgroups. All differences are used (149 informative sites). Decay values (d) are indicated above each branch. Bootstrap percentages of nodes that occurred in ≥50% of 250 bootstrap replicates are indicated below branches.

of conifers. Within this large group two distinct subclades are supported by more than three steps of decay and high bootstrap percentages: (1) Podocarpaceae plus Araucariaceae; and (2) a subclade comprising *Sciadopitys,* Taxaceae, Cephalotaxaceae, Taxodiaceae, and Cupressaceae. The Taxaceae are nested within the latter subclade, having the closest relationships with Cephalotaxaceae. The same topology was obtained in analyses using only *Ginkgo* or only cycads as outgroups.

Alternative topologies—Four specific alternative hypotheses proposed by different authors were tested. The first three correspond to those based on precladistic views, and the remaining one to a morphological cladistic

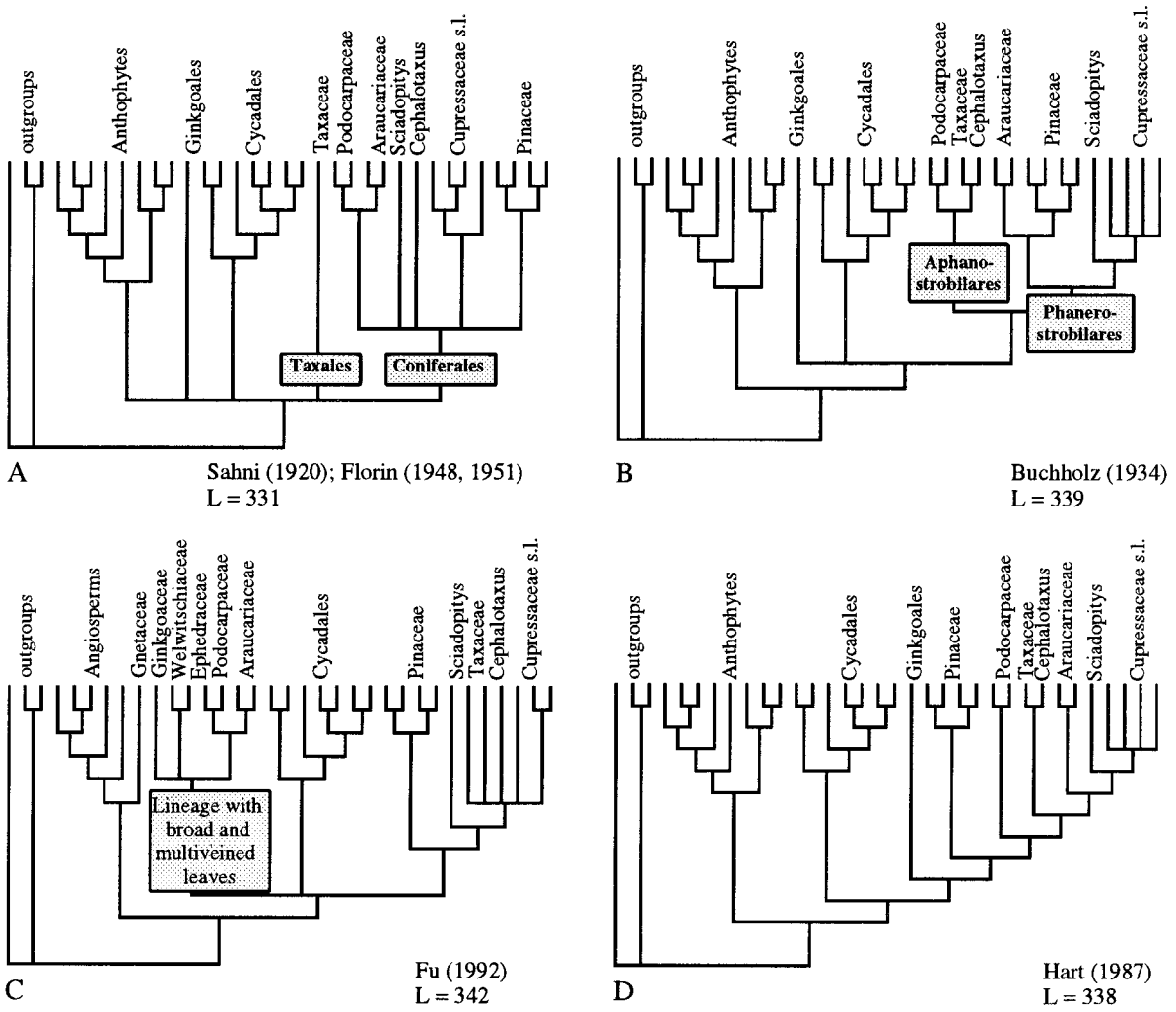

Fig. 6. Lengths (L) of alternative trees, with the following topological constraint implemented: (A) Taxales as a distinct order separate from the Coniferales; (B) conifers monophyletic but divided into two suborders: Phanerostrobilares and Aphanostrobilares; (C) monophyly of lineage with broad and multiveined leaves; (D) cladistic analysis based on morphological data. Authors of specific alternative hypotheses of gymnosperms relationships are indicated.

analysis (Fig. 6). All four alternative topologies resulted in trees 8–19 steps longer than the most parsimonious trees (Table 3). The recognition of Taxales as an order distinct from the Coniferales based on the absence of a clearly defined seed cone (Sahni, 1920; Florin, 1948, 1951) is relatively the least penalized hypothesis. Nevertheless, it requires eight additional steps. The hypothesis of Buchholz (1934), which considers conifers to be monophyletic but segregates them into two suborders, Phanerostrobilares and Aphanostrobilares, needs 16 additional steps. The greatest length penalty, 19 extra steps, is required to force the monophyly of a lineage with broad and multiveined leaves without midrib (Fu, 1992). When the topology based on a morphological cladistic analysis (Hart, 1987), which regards conifers as a monophyletic group, but with different relative positions of the families was imposed, it yielded trees 15 steps longer than our shortest trees (323 steps).

TABLE 3. Comparison of unconstrained, most parsimonious trees with specific alternative topologies proposed by different authors.

Topology	Length	Length penalty
Unconstrained	323	—
Taxales as a distinct order separate from the Coniferales (Sahni, 1920; Florin, 1948, 1951)	331	+8
Coniferales divided into two suborders: Phanerostrobilares and Aphanostrobilares (Buchholz, 1934)	339	+16
Monophyly of lineage with broad and multiveined leaves (Fu, 1992)	342	+19
Cladistic analysis based on morphological evidence (Hart, 1987)	338	+15

DISCUSSION

Our results based on 28S rRNA are not consistent with the hypothesis of Fu (1992) according to which the multiveined leaves of *Nageia* s.s. and those of Araucariaceae, Ephedraceae, Welwitschiaceae, and Ginkgoaceae are related and linked with extinct Cordaitales in a multinerve

lineage. Although in our data *Nageia* s.s. itself was not represented, it can be replaced by *Retrophyllum,* which, according to morphologic cladistic analysis of Podocarpaceae (Kelch, in press), belongs to the same clade—*Nageia* s.l. The preponderance of characters overwhelmingly supports the retention of Podocarpaceae and Araucariaceae within the conifers.

Furthermore, our ordinal-level analysis indicates that the Gnetales are a monophyletic clade, with *Ephedra* as the sister group of *Welwitschia* and *Gnetum,* and in the position of the closest modern relatives of angiosperms. This contradicts the suggestion that *Ephedra* is more closely related to *Ginkgo* (Meyen, 1984) or that *Ephedra* and [*Gnetum* plus *Welwitschia*] are paraphyletic lineages relative to the angiosperms (Nixon et al., 1994). Our results are consistent with cladistic analyses of morphological data (Crane, 1985; Doyle and Donoghue, 1986; Loconte and Stevenson, 1990; Doyle, 1996) and phylogenetic analyses of both *rbcL* (Hasebe et al., 1992; Chase et al., 1993) and rRNA data (Hamby and Zimmer, 1992; Chaw et al., 1995).

According to 28S rRNA sequences, conifers are nested within the clade that comprises *Ginkgo* and cycads. This position is not in agreement with numerous morphologically based cladistic analyses, which place anthophytes in the position of sister group of the conifers. In our analysis, however, it is only four steps less parsimonious to associate anthophytes with conifers or with cycads. The problem of choosing outgroups for conifers is closely connected with the rooting of spermatophytes as a whole. It should also be noted that molecular data are available from living taxa only, whereas the majority of morphological analyses include fossil taxa, and it has been shown that fossil information affects the phylogenetic location of extant taxa (Donoghue et al., 1989; Doyle and Donoghue, 1992). Therefore, it is possible that the relative position of conifers would be different in our trees if fossil sequences were included. In view of the small number of steps separating the different topologies and the fact that the actual relationships of seed plants remain problematical, we used *Ginkgo* and/or cycads as conifer outgroups based on the relationships inferred by our most parsimonious trees.

Our 28S rRNA sequence analyses lend support to the theories of a single origin for the conifers (here including Taxaceae and Cephalotaxaceae), largely following Pilger's (1926) classification. This is consistent with the morphological cladistic analysis of Hart (1987), the study of structural mutations of chloroplast DNA by Raubeson and Jansen (1992) and molecular phylogenetic analyses (Chaw et al., 1993, 1995; Price et al., 1993). Hart (1987) has already pointed out that the autapomorphies that separate conifers from all other living seed plants are principally embryological characters (embryo derived from five or less free nuclei and tiered proembryo). Other features such as siphonogamic sperm transfer, absent lagenostome, or epigeal germination can also be used to support the monophyly of conifers, although they show homoplasy within seed plants.

Within the conifers, the ''type'' family, the Pinaceae, is the first to diverge. The basal position of this family, whose belonging to conifers was never questioned, is also congruent with other cladistic analyses.

Two subclades diverge after the separation from Pinaceae, the first comprising *Sciadopitys,* Taxaceae, Cephalotaxaceae, and Cupressaceae and the second Podocarpaceae and Araucariaceae. The well-supported monophyly of these subclades implies that the fleshy cone character arose more than once, i.e., independently in some genera of Podocarpaceae and Taxaceae. Moreover, the portion of the ovulate cone that becomes fleshy varies across the different taxa. This is obviously an example of analogous structures possessing the same biological function—adaptation to animal dispersal. All this invalidates Buchholz's (1934) proposition to segregate conifers into two suborders, Phanerostrobilares and Aphanostrobilares, based on cone morphology.

The most parsimonious explanation for the position of Taxaceae, nested within the former subclade, suggests that their uniaxial and simple ovule is derived from a compound, biaxial one by reduction and subsequent shifting from a lateral to the terminal position, a scenario proposed by Harris (1976). Had the theory of Sahni (1920) or that of Florin (1948, 1951) been plausible, we should have found, in the one case the Taxaceae plus Cephalotaxaceae and in the other only Taxaceae at the base of the conifers (or even in a closer relationship with *Ginkgo*) on the shortest trees. Since such a solution is penalized by eight additional steps (Fig. 6.) it is clear that the terminal ovule position is a derived character rather than a primitive one, an autapomorphy of the Taxaceae.

The latter subclade links the Araucariaceae and Podocarpaceae, suggesting a closer relationship between these families, both restricted today to the Southern Hemisphere. One possible explanation is that these two families originated from a common ancestor located in Gondwana. Although the interpretation of biogeography at this level directly from the cladograms is complicated, in this case an important correlation exists between the distribution pattern and evolution. Nevertheless, due to the great age of these groups and the difficulties linked to the interpretation of fossils attributed to these families, this question requires further consideration (Broutin et al., unpublished data).

A newly described genus *Wollemia* (Jones, Hill, and Allen, 1995), with its broad and multiveined leaves without any midrib, wingless pollen, and fully fused bract-scale complex of the ovulate cones, belongs to the Araucariaceae, bearing the closest relationship to *Agathis.*

Our cladistic analysis of 28S rRNA sequences reinforces the view that Cupressaceae s.s. and Taxodiaceae, without *Sciadopitys,* form a monophyletic conifer lineage. The most parsimonious trees suggest that the major lineages of Taxodiaceae diverged first, and monophyletic Cupressaceae s.s. are derived from within Taxodiaceae. Our results are consistent with other recent interpretations (Eckenwalder, 1976; Hart, 1987; Price and Lowenstein, 1989; Brunsfeld et al., 1994). *Sciadopitys,* often classified in Taxodiaceae, is not closely related to Cupressaceae s.l. and should be excluded from that family (Price and Lowenstein, 1989; Price et al., 1993). Taking into account its well-supported solitary position on the shortest trees our analysis provides evidence that *Sciadopitys* represents a distinct monogeneric family, Sciadopityaceae, as proposed by Hayata (1931). This segregation is further supported by many traits, which include morphology and chromosome number (Schlarbaum and Tsuchia, 1985).

Although our data are devoid of some of the taxa that

have been proposed as segregate lineages (e.g., *Phyllocladus*), our analyses imply that the Coniferales are a natural order. Within this order at present we can recognize seven monophyletic families: Pinaceae Lindl., Podocarpaceae Endl., Araucariaceae Henkel & W. Hochst, Sciadopityaceae Hayata, Taxaceae Gray, Cephalotaxaceae Dumort., and Cupressaceae Rich ex Bartl.

LITERATURE CITED

BREMER, K. 1988. The limits of amino acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* 42: 795–803.

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 *rbc*L sequences. *Systematic Botany* 19: 253–262.

BUCHHOLZ, J. T. 1934. The classification of Coniferales. *Transactions of the Illinois State Academy of Science* 25: 112–113.

CHAMBERLAIN, C. J. 1935. Gymnosperms, structure and function. University of Chicago Press, Chicago, IL.

CHASE, M. W., ET AL. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from plastid gene *rbc*L. *Annals of the Missouri Botanical Garden* 80: 528–580.

CHAW, S. M., H. LONG, B. S. WANG, A. ZHARKIKH, AND W. H. LI. 1993. The phylogenetic positions of Taxaceae based on 18S rRNA sequences. *Journal of Molecular Evolution* 37: 624–630.

———, H. M. SUNG, H. LONG, A. ZHARKIKH, AND W. H. LI. 1995. The phylogenetic positions of the Conifer genera *Amentotaxus, Phyllocladus,* and *Nageia* inferred from 18S rRNA sequences. *Journal of Molecular Evolution* 41: 224–230.

CRANE, P. R. 1985. Phylogenetic analysis of seed plants and the origin of angiosperms. *Annals of the Missouri Botanical Garden* 72: 716–793.

DONOGHUE, M. J., J. A. DOYLE, J. GAUTHIER, A. G. KLUGE, AND T. ROWE. 1989. The importance of fossils in phylogeny reconstruction. *Annual Review of Ecology and Systematics* 20: 431–460.

———, R. G. OLMSTEAD, J. F. SMITH, AND J. D. PALMER. 1992. Phylogenetic relationships of Dipsacales based on *rbc*L sequences. *Annals of the Missouri Botanical Garden* 79: 333–345.

DOYLE, J. A. 1996. Seed plant phylogeny and the relationships of Gnetales. *International Journal of Plant Sciences* 157 (6 Suppl.): S3–S39.

———, AND M. J. DONOGHUE. 1986. Seed plant phylogeny and the origin of angiosperms: an experimental cladistic approach. *Botanical Review* 52: 321–431.

———, AND ———. 1987. The importance of fossils in elucidating seed plant phylogeny and macroevolution. *Review of Paleobotany and Palynology* 50: 63–95.

———, AND ———. 1992. Fossils and seed plant phylogeny reanalyzed. *Brittonia* 44: 89–106.

DOYLE, J. J., AND J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

ECKENWALDER, J. E. 1976. Re-evaluation of Cupressaceae and Taxodiaceae: a proposed merger. *Madroño* 23: 237–256.

FARRIS, J. S. 1989. The retention index and the rescaled consistency index. *Cladistics* 5: 417–419.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791.

FLORIN, R. 1948. On the morphology and relationships of the Taxaceae. *Botanical Gazette* 110: 31–39.

———. 1951. Evolution in cordaites and conifers. *Acta Horti Bergiani* 15: 1–389.

FU, D. Z. 1992. Nageiaceae—a new gymnosperm family. *Acta Phytotaxonomica Sinica* 30: 515–528.

HAMBY, R. K., AND E. A. ZIMMER. 1992. Ribosomal RNA as a phylogenetic tool in plant systematics. *In* P. S. Soltis, D. E. Soltis, and J. J. Doyle [eds.], Molecular systematics of plants, 50–91. Chapman and Hall, London.

HARRIS, T. M. 1976. The mesozoic gymnosperms. *Review of Paleobotany and Palynology* 21: 119–134.

HART, J. A. 1987. A cladistic analysis of conifers: preliminary results. *Journal of the Arnold Arboretum* 68: 269–307.

HASEBE, M., M. ITO, R. KOFUJI, K. IWATSUKI, AND K. UEDA. 1992. Phylogenetic relationships in Gnetophyta deduced from *rbc*L gene sequences. *Botanical Magazine (Tokyo)* 105: 385–391.

HASSOUNA, N., B. MICHOT, AND J. P. BACHELLERIE. 1984. The complete nucleotide sequence of mouse 28S rRNA gene: implication for the process of size increase of the large subunit rRNA in higher eukaryotes. *Nucleic Acids Research* 12: 3563–3583.

HAYATA, B. 1931. The Sciadopityaceae represented by *Sciadopitys verticillata* Sieb. et Zucc., an endemic species of Japan. *Botanical Magazine (Tokyo)* 45: 567–569.

HILLIS, D. M. 1991. Discriminating between phylogenetic signal and random noise in DNA sequences. *In* M. M. Miyamoto and J. Cracraft [eds.], Phylogenetic analysis of DNA sequences, 278–294. Oxford University Press, Oxford.

HUELSENBECK, J. P. 1991. Tree-length distribution skewness: an indicator of phylogenetic information. *Systematic Zoology* 40: 257–270.

JONES, W. G., K. D. HILL, AND M. J. ALLEN. 1995. *Wollemia nobilis,* a new living Australian genus and species in the Araucariaceae. *Telopea* 6: 173–176.

KELCH, D. G. In press. The phylogeny of the Podocarpaceae based on morphological evidence. *Systematic Botany.*

KENG, H. 1974. The phylloclade of *Phyllocladus* and its possible bearing on the branch systems of Progymnosperms. *Annals of Botany* 38: 757–764.

———. 1978. The genus *Phyllocladus* (Phyllocladaceae). *Journal of the Arnold Arboretum* 59: 249–273.

KLUGE, A. G., AND J. S. FARRIS. 1969. Quantitative phyletics and the evolution of anurans. *Systematic Zoology* 18: 1–32.

LI, H. L. 1953. Present distribution and habitats of the conifers and taxads. *Evolution* 7: 245–261.

LOCONTE, H., AND D. W. STEVENSON. 1990. Cladistic of the Spermatophyta. *Brittonia* 42: 197–211.

MADDISON, D. R. 1991. The discovery and importance of multiple islands of most-parsimonious trees. *Systematic Zoology* 40: 315–328.

MEYEN, S. V. 1984. Basic features of gymnosperm systematics and phylogenie as evidenced by the fossil record. *Botanical Review* 50: 1–112.

MULLIS, K. B., AND F. A. FALOONA. 1987. Specific synthesis of DNA in vitro via a polymerase catalysed chain reaction. *Methods Enzymol.* 155: 335–350.

NIXON, K. C., W. L. CREPET, D. STEVENSON, AND E. M. FRIIS. 1994. A reevaluation of phylogenetic relationships of seed plants. *Annals of the Missouri Botanical Garden* 81: 484–533.

PHILIPPE, H. 1993. MUST: a computer package of management utilities for sequences and trees. *Nucleic Acids Research* 21: 5264–5272.

———, U. SÖRHANNUS, A. BAROIN, R. PERASSO, F. GASSE, AND A. ADOUTTE. 1994. Comparison of molecular and paleontological data in diatoms suggests a major gap in the fossil record. *Journal of Evolutionary Biology* 7: 247–265.

PILGER, R. 1926. Coniferae. *In* A. Engler and K. Prantl [eds.], Die Natürlichen Pflanzenfamilien, 13: 121–407. W. Engelmann, Leipzig.

PRICE, R. A., AND J. M. LOWENSTEIN. 1989. An immunological comparision of the Sciadopityaceae, Taxodiaceae, and Cupressaceae. *Systematic Botany* 14: 141–149.

———, J. THOMAS, S. H. STRAUSS, P. A. GADEK, C. J. QUINN, AND J. D. PALMER. 1993. Familial relationships of the conifers from *rbc*L sequence data. *American Journal of Botany* 80: 172.

RAUBESON, L. A., AND R. K. JANSEN. 1992. A rare chloroplast-DNA structural mutation is shared by all conifers. *Biochemical Systematics and Ecology* 20: 17–24.

ROGERS, S. O., AND A. J. BENDICH. 1985. Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. *Plant Molecular Biology* 5: 69–76.

ROTHWELL, G. W., AND R. SERBET. 1994. Lignophyte phylogeny and the evolution of Spermatophytes: a numerical cladistic analysis. *Systematic Botany* 19: 443–482.

SAHNI, B. 1920. On certain archaic features in the seed of *Taxus baccata,* with remarks on the antiquity of Taxineae. *Annals of Botany* 34: 117–133.

SAIKI, R. K., D. H. GELFAND, S. STOFFEL, S. J. SCHARF, R. HIGUCHI, G. T. HORN, K. B. MULLIS, AND H. A. ERLICH. 1988. Primer-directed enzymatic amplificaton of DNA with thermostable DNA polymerase. *Science* 239: 487–491.

SANDERSON, M. J., AND M. J. DONOGHUE. 1989. Patterns of variation in levels of homoplasy. *Evolution* 43: 1781–1795.

SANGER, F., S. NICKLEN, AND A. R. COULSON. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences, USA* 74: 5463–5467.

SCHLARBAUM, S. E., AND T. TSUCHIA. 1985. Karyological derivation of *Sciadopitys verticillata* Sieb. et Zucc. from a protaxodiaceous ancestor. *Botanical Gazette* 146: 264–267.

SWOFFORD, D. L. 1993. PAUP: Phylogenetic analysis using parsimony, version 3.1.1. Illinois Natural History Survey, Champaign, IL.

TAKHTAJAN, A. L. 1953. Phylogenetic principles of the system of higher plants. *Botanical Review* 19: 1–45.