Inference of higher-order conifer relationships from a multi-locus plastid data set[1]

Hardeep S. Rai, Patrick A. Reeves, Rod Peakall, Richard G. Olmstead, and Sean W. Graham

Abstract: We reconstructed the broad backbone of conifer phylogeny from a survey of 15–17 plastid loci and associated noncoding regions from exemplar conifer species. Parsimony and likelihood analyses recover the same higher-order relationships, and we find strong support for most of the deep splits in conifer phylogeny, including those within our two most heavily sampled families, Araucariaceae and Cupressaceae. Our findings are broadly congruent with other recent studies, and are inferred with comparable or improved bootstrap support. The deepest phylogenetic split in conifers is inferred to be between Pinaceae and all other conifers (Cupressophyta). Our current gene and taxon sampling does not support a relationship between Pinaceae and Gnetales, observed in some published studies. Within the Cupressophyta clade, we infer well-supported relationships among Cephalotaxaceae, Cupressaceae, Sciadopityaceae, and Taxaceae. Our data support recent moves to recognize *Cephalotaxus* under Taxaceae, and we find strong support for a sister-group relationship between the two predominantly southern hemisphere conifer families, Araucariaceae and Podocarpaceae. A local hotspot of indel evolution shared by the latter two conifer families is identified in the coding portion of one of the plastid ribosomal protein genes. The removal of the most rapidly evolving plastid characters, as defined using a likelihood-based classification of substitution rates for the taxa considered here, is shown to have little to no effect on our inferences of higher-order conifer relationships.

Key words: chloroplast genome, gymnosperms, microstructural mutations, *rps*7, seed-plant phylogeny, *Wollemia*.

Résumé : Les auteurs ont mis en place une large structure centrale pour la phylogénie des conifères à partir d'une étude sur 15–17 lieux de plastes et en associant des régions non-codantes pour des espèces de conifères de référence. Les analyses de parcimonie et de probabilité ramènent la même relation d'ordre supérieur, et on constate un fort support pour les divisions marquées dans la phylogénie des conifères, incluant celles qu'on trouve dans les familles des Araucariaceae et des Cupressaceae, fortement échantillonnées par les auteurs. Les résultats montrent une forte correspondance avec ceux d'autres études récentes et en infèrent avec un support bootstrap comparable ou amélioré. On déduit que la division phylogénétique la plus marquée chez les conifères se retrouve entre les Pinacea et tous les autres conifères. Les échantillons de gènes et de taxons réunis par les auteurs ne supportent aucune relation entre les Pinaceae et les Gnétales, comme mentionnée dans certaines publications. Pour le clade des Cupressophyte, les auteurs déduisent de fortes relations au sein des Cephalotaxaceae, Cupressaceae, Sciadopityaceae et Taxaceae. Les données supportent la récente proposition pour reconnaître le *Cephalotaxus* comme Taxaceae, et on trouve un fort support pour une relation de groupe sœur entre deux familles prédominantes dans l'hémisphère sud, les Araucariaceae et les Podocarpaceae. On a identifié un point chaud d'évolution d'indels, commun aux deux familles de conifères précitées, dans la portion codante d'un des gènes des protéines ribosomiques plastidiques. L'élimination des caractères plastidiques évoluant le plus rapidement, comme défini en utilisant une classification basée sur la probabilité des taux de substitution chez les taxons considérés, montre qu'elle a peu ou pas d'effet sur les déductions des auteurs concernant les relations d'ordre supérieur, chez les conifères.

Mots-clés : génome chloroplastique, gymnospermes, mutations microstructurales, *rps*7, phylogénie des plantes à graines, *Wollemia*.

[Traduit par la Rédaction]

Received 30 July 2007. Published on the NRC Research Press Web site at botany.nrc.ca on 12 June 2008.

H.S. Rai[2] and S.W. Graham.[4] UBC Botanical Garden & Centre for Plant Research (Faculty of Land and Food Systems), and Department of Botany, 2357 Main Mall, University of British Columbia, Vancouver, BC V6T 1Z4, Canada.
P.A. Reeves[3] and R.G. Olmstead. Department of Biology, University of Washington, Seattle, WA 98195, USA.
R. Peakall. School of Botany and Zoology, The Australian National University, Canberra, ACT 0200, Australia.

[1]This paper is one of a selection of papers published in the Special Issue on Systematics Research.
[2]Present address: The Arnold Arboretum of Harvard University, 22 Divinity Avenue, Cambridge, MA 02138, USA.
[3]Present address: Agricultural Research Service, National Center for Genetic Resources Preservation, United States Department of Agriculture, 1111 South Mason Street, Fort Collins, CO 80521-4500, USA.
[4]Corresponding author (e-mail: swgraham@interchange.ubc.ca).

 doi:10.1139/B08-062

Introduction

Conifers have a rich and deep fossil record, with taxa assignable to extant families dating back to the Triassic (Yao et al. 1997; Stockey et al. 2005). Although angiosperms are now the dominant group of seed plants in most terrestrial ecosystems, conifers are still ecologically significant in all continental floras (Enright and Hill 1995), and they dominate the northern boreal forests. Approximately 670 extant species are recognized in some 70 genera. Most conifer systematists recognize seven families to accommodate their diversity: Araucariaceae, Cephalotaxaceae, Cupressaceae, Pinaceae, Podocarpaceae, Sciadopityaceae, and Taxaceae (Phyllocladaceae are sometimes separated from Podocarpaceae, while Taxodiaceae are now usually included in Cupressaceae). Considerable progress in our understanding of conifer classification and phylogenetics has been made using multiple lines of evidence (e.g., Quinn et al. 2002). For example, Taxaceae have sometimes been treated as an order distinct from other conifers (e.g., Florin 1951), but morphological and molecular data clearly support a nested phylogenetic position for the family among the rest of the conifers (e.g., Hart 1987; Raubeson and Jansen 1992; Quinn et al. 2002).

However, there are still points of weakness in our understanding of conifer higher-order phylogenetic relationships. For example, it is still not clear whether extant conifers are monophyletic (e.g., Burleigh and Mathews 2004, 2007*a*, 2007*b*). Some studies find the enigmatic Gnetales (*Ephedra* L., *Gnetum* L., and *Welwitschia* Hook. f.) to be the sister group of the pines and their relatives (Pinaceae) with moderate to strong support (the "gnepine" hypothesis, see Bowe et al. 2000; Chaw et al. 2000; Gugerli et al. 2001), whereas others support the monophyly of extant conifers, with Gnetales being placed elsewhere in seed-plant phylogeny (e.g., Chaw et al. 1997; Rydin et al. 2002; Rai et al. 2003). A placement of Gnetales as sister to Pinaceae is difficult to justify on morphological grounds (e.g., Donoghue and Doyle 2000), and may be a strong analytical artifact, perhaps due to long-branch attraction (e.g., Burleigh and Mathews 2004). On the other hand, a sister-group relationship between Gnetales and conifers among extant seed plants, a result seen in a subset of molecular studies (e.g., Chaw et al. 1997), is less problematic from a morphological perspective (e.g., Mundry and Stützel 2004; Doyle 2005).

Although there is considerable disparity among molecular studies concerning the placement of Gnetales, there is broad agreement on a sister-group relationship between Pinaceae (or gnepines) and the remaining conifer families. Phylogenetic studies have also clarified the circumscription and interrelationships of the other conifer families. For example, they have led to the recognition of a sister-group relationship between the two predominantly southern hemisphere taxa, Araucariaceae and Podocarpaceae, and have provided support for Araucariaceae–Podocarpaceae as the sister group of a clade consisting of Cephalotaxaceae, Cupressaceae, Sciadopityaceae, and Taxaceae (Chaw et al. 1997; Stefanovic et al. 1998; Gugerli et al. 2001; Quinn et al. 2002; Rydin et al. 2002). The resulting large clade — comprising all extant conifers except Pinaceae — has been referred to informally as "conifers II" (e.g., Rydin et al. 2002), and more recently as Cupressophyta by Cantino et al. (2007).

Within Cupressophyta, various molecular and morphological phylogenetic studies support the existence of a clade consisting of members of Cephalotaxaceae and Taxaceae (e.g., Hart 1987; Cheng et al. 2000; Quinn et al. 2002), although there is some uncertainty about the limits and monophyly of Taxaceae (e.g., Page 1990*d*). For example, Quinn et al. (2002) proposed that Taxaceae should be circumscribed to include *Cephalotaxus* Siebold & Zucc. ex Endl., although this recommendation is not yet generally followed. Most taxa formerly included in Taxodiaceae are now recognized under a more broadly defined Cupressaceae, a circumscription proposed by Eckenwalder (1976) on morphological grounds, which has since been supported by numerous phylogenetic studies (e.g., Hart 1987; Brunsfeld et al. 1994; Stefanovic et al. 1998; Gadek et al. 2000; Kusumi et al. 2000; Quinn et al. 2002; Rydin et al. 2002). The distinctiveness of *Sciadopitys verticillata* (Thunb.) Siebold & Zucc., traditionally considered to belong to Cupressaceae, supports its recognition as a separate family, Sciadopityaceae (e.g., Page 1990*c*). Morphological and molecular phylogenetic data confirm this view (e.g., Hart 1987; Brunsfeld et al. 1994). Both families are now recognized as part of the larger clade that includes Cephalotaxaceae and Taxaceae. Finally, the relative arrangement of Cephalotaxaceae, Cupressaceae, Taxaceae, and Sciadopityaceae to each other is incompletely understood (e.g., Stefanovic et al. 1998; Quinn et al. 2002; Rydin et al. 2002).

The main goal of our study is to obtain well-supported relationships for the deep branches of conifer phylogeny by surveying a large multigene plastid data set (15–17 plastid genes and associated noncoding regions) for a broad range of exemplar conifers. Increasing the amount of nucleotide data sampled per taxon has been shown to be an effective way to clarify our understanding of deep phylogenetic relationships in various groups of plants, and to generally increase support for phylogenetic inferences (for empirical examples using the current gene set see Graham and Olmstead 2000*a*; Rai et al. 2003; Graham et al. 2006; Saarela et al. 2007; Zgurski et al. 2008). Our study focuses on relationships among the families, but we also sampled Araucariaceae and Cupressaceae at sufficient taxonomic depth to address basic features of their internal phylogenetic structure. Rapidly evolving characters can have a substantial impact on the inference of overall seed-plant relationships (Burleigh and Mathews 2004, 2007*b*; H.S. Rai and S.W. Graham, unpublished data, 2008), and so we assess whether this affects phylogenetic inference within the conifers by including or excluding the most rapidly evolving characters from consideration. We also characterize a curious structural mutation in one of the plastid ribosomal protein genes from two families of conifers, Araucariaceae and Podocarpaceae.

Materials and methods

Plant material and genomic sampling

We surveyed 17 genes, which together with their associated noncoding regions represent between one-eighth and one-ninth of the entire plastid genome (~120 kb in

Table 1. Source information and GenBank numbers.

Taxon (voucher, herbarium)	Gene or region	
	*atp*B	*ndh*F
Araucariaceae		
Agathis australis (D. Don) Loudon (H.S. Rai 1002, ALTA)	AY664829	AY902169
Agathis robusta (C. Moore ex F. Muell.) F.M. Bailey (037944–037947, GAU)	EF490502	EF494250
Araucaria bidwillii Hook. (H.S. Rai 1006, ALTA)	AY664830	AY902170
Araucaria cunninghamii Aiton ex D. Don (037942 & 037943, GAU)	EF490503	EF494251
Wollemia nobilis W.G. Jones, K.D. Hill & J.M. Allen (no voucher†)	EF490504	EF494249
Cephalotaxaceae		
Cephalotaxus harringtonii (Knight ex J. Forbes) K. Koch (R.G. Olmstead 2000-55, WTU	AY664831	AY902171
Cupressaceae s.l.		
Cunninghamia lanceolata (Lamb.) Hook. (P.A. Reeves & J. Metropulos 18, WTU)	AY664833	AY902174
Juniperus communis L. (H.S. Rai 1011, ALTA)	AY664834	AY902175
Taxodium distichum (L.) Rich. (K. Ikegama 2002-1, WTU	AY664835	AY902176
Thuja plicata Donn ex D. Don (P.A. Reeves & J. Metropulos 19, WTU)	AY664836	AY902177
Widdringtonia cedarbergensis J.A. Marsh (H.S. Rai 1001, ALTA)	n/a	AY902178
Pinaceae		
Abies lasiocarpa (Hook.) Nutt. (R.G. Olmstead 2001-82, WTU)	AY664825	n/a
Pseudotsuga menziesii (Mirb.) Franco (H.S. Rai 1022, ALTA)	AY664826	n/a
Podocarpaceae		
Phyllocladus alpinus Hook. f. (R.G. Olmstead 2000-54, WTU)	AY664827	AY902167
Saxegothaea conspicua Lindl. (D.M. Cherniawsky ZB-VI-VII, ALTA)	AY664828	AY902168
Taxaceae		
Taxus brevifolia Nutt. (A. Colwell 2000-32, WTU)	AF528864	AY902172
Torreya californica Torr. (H.S. Rai 1008, ALTA)	AY664832	AY902173

*Previously published sequences; see Graham and Olmstead (2000*a*,2000*b*) and Rai et al. (2003) for a complete list of taxa and accession numbers for *Metasequoia glyptostroboides* Hu & W.C. Cheng (Cupressaceae), *Podocarpus chinensis* Wall. ex J. Forbes (Podocarpaceae) and *Sciadopitys verticillata*

†This voucherless sample is from the same population as vouchered specimens described in Jones et al. (1995).

Pinus L.; Wakasugi et al. 1994). The coding regions include photosynthetic genes (*atp*B, *rbc*L, and 10 photosystem II, *psb*, genes), translation apparatus genes (the plastid ribosomal protein genes *rpl*2, *rps*7, and 3′-*rps*12), and two chlororespiratory genes (*ndh*B and *ndh*F, which code for two of the subunits of plastid NADH dehydrogenase). The noncoding regions consist of three introns (in *rpl*2, 3′-*rps*12, and *ndh*B) and eight intergenic spacer regions (Table 1). We used exemplar-based taxon sampling to represent the major branches of conifer phylogeny; in choosing representatives for each nonmonotypic family we attempted to represent their internal systematic diversity as broadly as possible, at least as understood from prior studies. In total, we included 22 exemplar conifer species and multiple outgroups (11 other seed plants, 2 monilophytes, and 3 bryophytes). Source and GenBank information is provided in Table 1.

Recovery of plastid sequences, DNA alignment, and characterization of an indel hotspot

We extracted DNA from fresh and silica-dried specimens following Doyle and Doyle (1987) and Rai et al. (2003). DNA samples of *Wollemia* W.G. Jones, K.D. Hill, & J.M. Allen, *Agathis robusta* (C. Moore ex F. Muell.) F.M. Bailey and *Araucaria cunninghamii* Aiton ex D. Don were extracted as described in Peakall et al. (2003). DNA amplification and sequencing methods follow Graham and Olmstead (2000*a*). We sequenced all regions at least twice for each taxon, and with a few exceptions completely sequenced all regions in both directions. Several regions confirmed as lost from the plastid genome or that we could not amplify were coded as missing data in the final matrix. Two genes that are missing (or not retrievable) for Pinaceae are *ndh*B and *ndh*F (see Wakasugi et al. 1994); these two genes and *rpl*2 were also not retrievable from the Gnetales exemplars examined here. We were unable to recover *atp*B from *Widdringtonia cedarbergensis* J.A. Marsh and *rpl*2 from *Thuja plicata* Donn ex D. Don. We excluded noncoding regions for three of the outgroup taxa (*Anthoceros* L., *Marchantia* L. and *Physcomitrella* Bruch & Schimp.), because these were difficult to align across land plants.

We compiled contiguous sequences, performed basecalling using Sequencher 4.1 (Gene Codes Corporation; Ann Arbor, Mich.), and added new sequences to an alignment (Graham et al. 2006) that includes sequences generated for previous studies of seed-plant phylogeny (Graham and Olmstead, 2000*a*, 2000*b*; Graham et al. 2000; Rai et al. 2003). We adjusted alignments manually for each contiguous region using Se-Al version 1.0 (Rambaut 1998), following alignment criteria in Graham et al. (2000), and used tobacco (*Nicotiana tabacum* L.), *Ginkgo* L., and *Pinus* sequences to define gene and exon boundaries, following Graham and Olmstead (2000*a*). We offset several regions that were too difficult to align in the noncoding regions [the intergenic spacers (IGS) of two of the photosystem II clusters (*psb*E–*psb*F–*psb*L–*psb*J and *psb*B–*psb*T–*psb*N–*psb*H), the IGS between *rps*7–*ndh*B, and the introns],

*psb*B, T, N, & H	*psb*D & C	*psb*E, F, L, & J	*rbc*L	*rpl*2	3′-*rps*12, *rps*7 & *ndh*B
AF528892	AF528919	AF528865	AF362993*	AY664864	AY164586
EF490512	EF490506	EF490515	EF490509	EF490521	EF490518
AY664852	AY664840	AY664846	U96472*	AY664865	AY664816
EF490513	EF490507	EF490516	EF490510	EF490522	EF490519
EF490511	EF490505	EF490514	EF490508	EF490520	EF490517
AF528896	AF528923	AF528869	AF227461*	AY664866	AY664817
AF528898	AF528925	AF528871	L25757*	AY664869	AY664820
AY664854	AY664842	AY664848	AY664859	AY664870	AY664821
AF528915	AF525949	AF528888	AF119185*	AY664871	AY664822
AF528917	AF528942	AF528890	AF127428*	n/a	AY664823
AF528918	AF528943	AF528891	AY140261	AY664872	AY664824
AY664849	AY664837	AY664843	AY664855	AY664860	AY664813
AY664850	AY664838	AY664844	AY664856	AY664861	AY664814
AF528905	AF528933	AF528879	AF249650*	AY664862	AY237142
AY664851	AY664839	AY664845	AY664857	AY664863	AY664815
AF528916	AF525948	AF528889	AF249666*	AY664867	AY664818
AY664853	AY664841	AY664847	AY664858	AY664868	AY664819

other taxa considered here, including the following conifers: *Cedrus deodora* (Roxb.) G. Don and *Pinus thunbergii* Parl. (Pinaceae), (Sciadopityaceae).

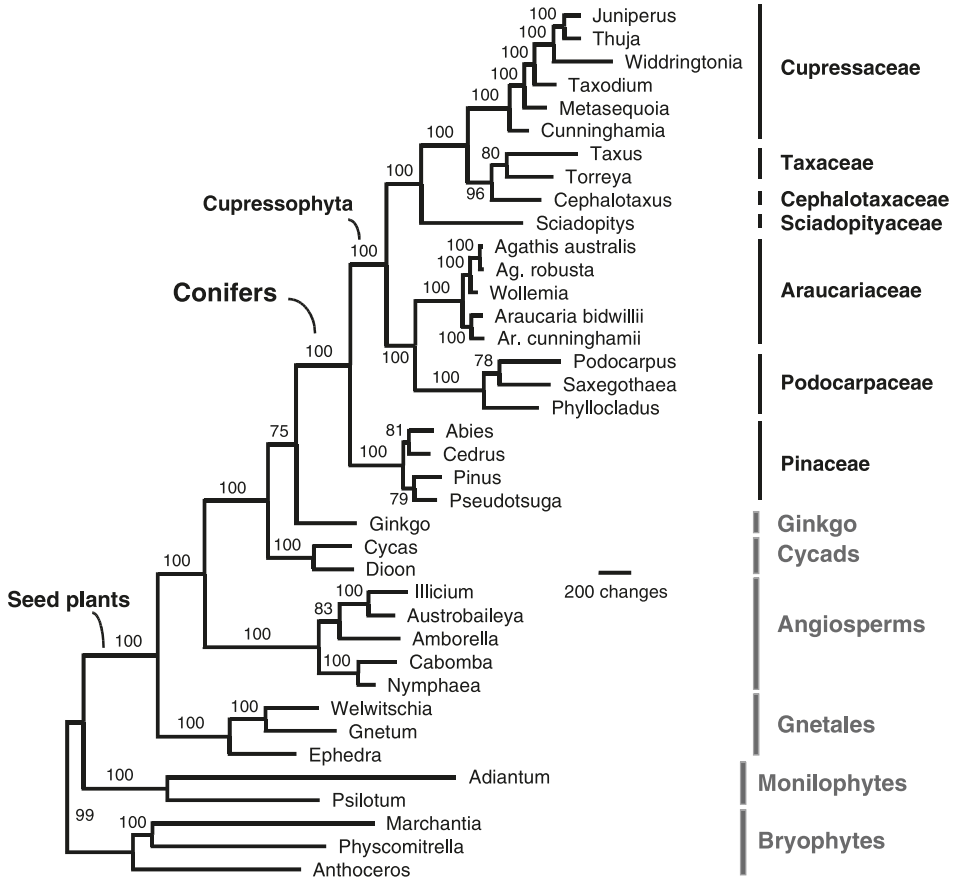
following Graham et al. (2006). The resulting staggered regions were frequently limited to single taxa, which are effectively ignored for parsimony-based tree searches and scores (Graham et al. 2006) and should have only minimal effect for model-based methods (e.g., on estimation of base frequency parameter values). Subsets of the offset regions include aligned blocks involving two or more taxa. The final alignment is 25 687 bp in length, derived from ~14 kb of unaligned data per taxon (e.g., 14.1 kb in *Agathis australis* (D. Don) Loudon). Of the total, 5384 aligned sites are potentially parsimony informative, and 2575 variable but parsimony uninformative. We also characterized a structural mutation in the ribosomal protein gene *rps*7 of Araucariaceae and Podocarpaceae, using the Dotlet browser-based application (version 1.5; Junier and Pagni 2000) to make pairwise amino-acid comparisons under the PAM-30 matrix of amino-acid substitution.

Phylogenetic analyses

We performed heuristic maximum parsimony (MP) and maximum likelihood (ML) searches using PAUP* (version 4.0b10; Swofford 2002) and PhyML (version 2.4.4; Guindon and Gascuel 2003). For the MP analysis (using PAUP*), we treated all characters and character-state changes as equally weighted, and used TBR (tree bisection–reconnection) branch swapping with 100 random addition replicates. PAUP* defaults were used for all other settings. For the ML search (using PhyML), we first chose a model of DNA sequence evolution with the hierarchical likelihood ratio test (hLRT) and the Akaike Information Criterion (AIC), using Modeltest (version 3.7; Posada and Crandall 1998), estimating model parameters from the data in each case. Both assessment methods recovered the same optimal DNA substitution model, GTR + Γ + I [i.e., the general-time-reversible (GTR) model, with among-site rate variation accounted for by considering the proportion of invariable sites (I), and the gamma (Γ) distribution, with four substitution-rate categories for the shape parameter alpha (α)]. We estimated substitution model parameters (base frequencies, the proportion of invariable sites, and the gamma distribution parameter) during the ML search. We assessed branch support using the nonparametric bootstrap (Felsenstein 1985) with 100 bootstrap replicates (in the MP search using one random addition replicate per bootstrap replicate). We use "weak," "moderate," and "strong" in reference to clades that have bootstrap support values <70%, 70%–89%, and ≥90%, respectively (e.g., Graham et al. 1998).

We re-analyzed the main matrix after removal of the most rapidly evolving characters to assess whether they distort the inference of conifer higher-order relationships. We used HyPhy (Kosakovsky Pond et al. 2005; and see Burleigh and Mathews 2004) to classify each alignment site into one of nine rate change classes (referred to as RC0-RC8, with RC0 being the zero-rate category and RC8 the fastest). The single most parsimonious tree (see below) was used as a reference tree for estimating GTR model parameters and the site rate

Fig. 1. Plastid-based phylogeny of the conifers and relatives inferred from MP for 15–17 chloroplast genes and associated noncoding regions (three introns and eight intergenic spacer regions). This single most-parsimonious tree (21 532 steps, CI = 0.541, RI = 0.659) is depicted as a phylogram, with ACCTRAN optimization of branch lengths. MP bootstrap values are indicated beside branches.



classifications in HyPhy. We reran the MP and ML analyses after excluding the two fastest rate categories from consideration (i.e., RC7 and RC8, see Burleigh and Mathews 2004).

Results

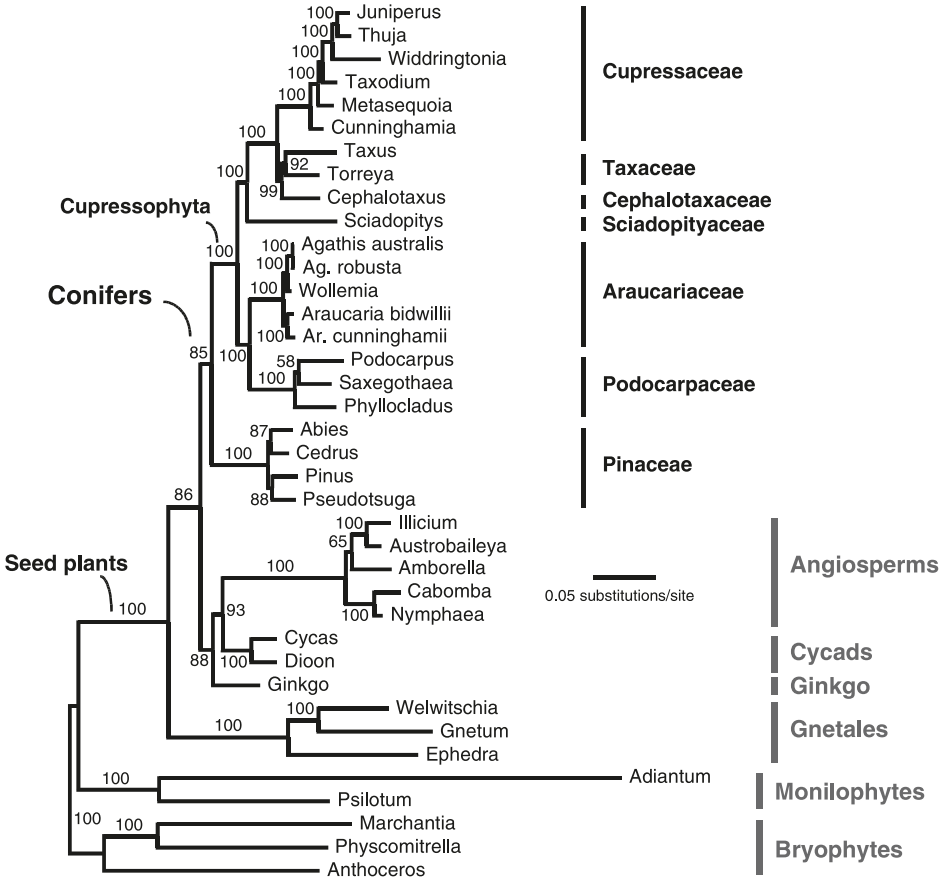
The relationships inferred among the major groups of seed plants differ in the MP and ML analysis of the full plastid data set, with *Ginkgo* (MP) or a clade consisting of *Ginkgo*, cycads, and angiosperms (ML) inferred to be the sister group of conifers, in both cases with moderately strong support (Figs. 1 and 2). Neither method supports a placement of Gnetales in or near the conifers, and both support conifer monophyly (100% from MP, 85% bootstrap support from ML). Both methods infer identical relationships within conifers (Figs. 1 and 2), with four of the five nonmonogeneric conifer families strongly supported as monophyletic, at least at the current level of sampling of taxa.

The two fastest rate classes (RC7 and RC8) comprise 4252 characters, corresponding to a substantial fraction of all parsimony informative characters (~79% of 5384 sites). Of these RC78 sites, only 3501 are parsimony informative, and so deleting these sites (corresponding to the RC0-6 analyses) leaves ~35% of all parsimony informative sites, not ~21% (the fraction expected if RC78 sites are all parsimony informative). The parsimony uninformative RC78 sites are all variable. We examined a subset of these and found them to be sites in alignment blocks that include only a few taxa. Presumably these are predicted to be rapidly evolving sites in the ML classification because variation is seen in them across a small taxon sampling.

When all RC78 characters are excluded from consideration, support for relationships among the five major groups of seed plants falls substantially, providing only poor to moderate support for the relevant branches (data not shown). However, relationships inferred within conifers are essentially unchanged after deletion of the fastest characters, with mostly very minor shifts in bootstrap support (cf. Figs. 1–3). Bootstrap support for conifer monophyly (moderate to strong support from ML and MP analysis, respectively) is also largely unchanged. However, the best MP and ML trees for RC0-6 do not depict Taxaceae as monophyletic (e.g., Fig. 3), and the clade consisting of these three taxa is then only moderately supported (70%–80%).

The remaining results focus on the analyses that consider all of the data. With all data included, the Cupressophyta clade is well supported as monophyletic, with 100% support from MP and ML bootstrap analysis (Figs. 1 and 2). Within this clade, Araucariaceae and Podocarpaceae are strongly supported as sister taxa (100% for MP and ML), and this two-family clade is in turn strongly supported as the sister-group of a clade consisting of Cupressaceae, Cephalotaxaceae, Sciadopityaceae, and Taxaceae (i.e., both Cupressophyta

Fig. 2. Plastid-based phylogeny of the conifers and relatives inferred from ML for 15–17 chloroplast genes and associated noncoding regions (three introns and eight intergenic spacer regions) using the GTR + Γ + I model of sequence evolution. The ML tree (–lnL = 128 297.454) is depicted as a phylogram. ML bootstrap values are indicated beside branches.

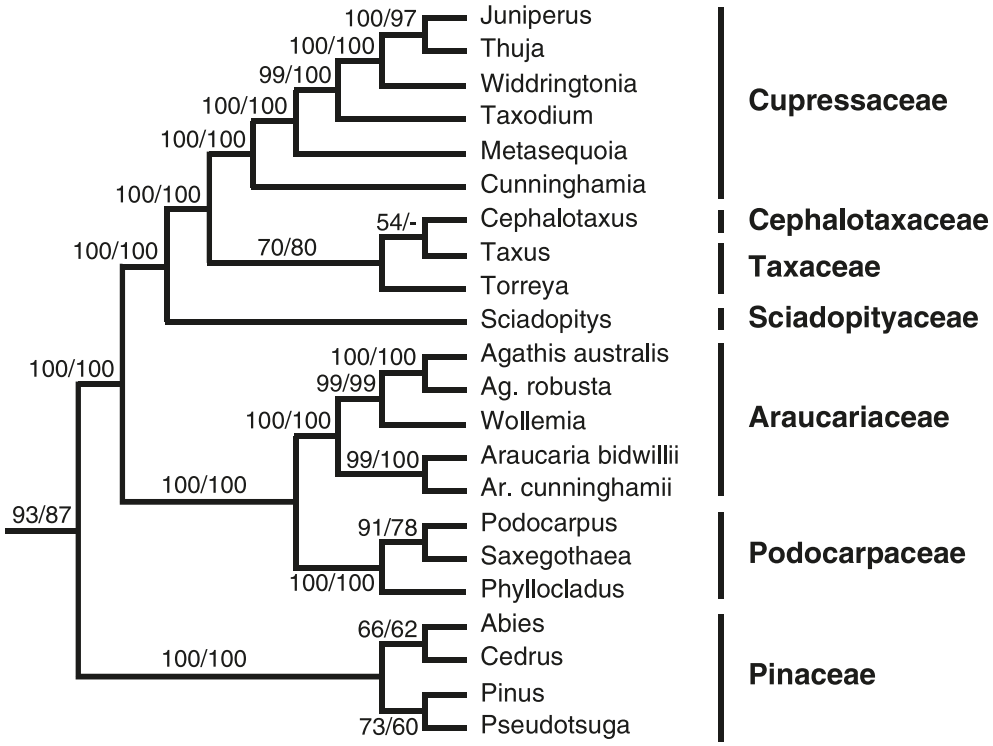


and the latter four-family clade are strongly supported). The interfamilial relationships within the latter clade are also well supported, with 96% to 100% bootstrap support from MP and ML analysis. More specifically, *Sciadopitys* Siebold & Zucc. (Sciadopityaceae) is strongly supported as the sister group of the remaining three families in the latter clade; Cephalotaxaceae and Taxaceae are sister taxa, and the Cephalotaxaceae–Taxaceae clade is then sister to Cupressaceae.

Of the four families sampled for more than two exemplar taxa (Figs. 1 and 2), Pinaceae and Podocarpaceae have only weakly to moderately supported intrafamilial backbones (58%–88% bootstrap support), although the same deep splits are inferred by the two phylogenetic methods (i.e., in Pinaceae, *Abies* Mill. is sister to *Cedrus* Trew and *Pinus* sister to *Pseudotsuga* Carrière; in Podocarpaceae, *Podocarpus* L'Hér. ex Pers. is sister to *Saxegothaea* Lindl.). In contrast, all relationships inferred within Araucariaceae and Cupressaceae have 100% MP and ML bootstrap support. In particular, the three "core Cupressaceae" taxa sampled here (*Juniperus* L., *Thuja* L., and *Widdringtonia* Endl.) are deeply nested among other members of Cupressaceae, with a basal split in the family seen between *Cunninghamia* R. Br. ex A. Rich. and other taxa. Within Araucariaceae, *Wollemia* is strongly supported as the sister group of *Agathis* Salisb.

A structural feature in the 5′-end of one of the ribosomal proteins considered here, *rps*7, is worth commenting on, as it seems to represent an otherwise quiescent region that has experienced a "recent" burst of microstructural mutations (insertions and deletions) in Araucariaceae and Podocarpaceae, including at least one tandem repeat expansion shared by these two families (Fig. 4). We refer to this hotspot of structural mutations as an "expansion region," as all taxa examined in Araucariaceae and Podocarpaceae are longer, owing to (predicted) insertions in this region relative to other land plants. However, it should be noted that the region is likely to have undergone both expansions and contractions (data not shown). The total expansion region is quite complex and includes multiple repeated motifs, a subset of which is shared among taxa in Araucariaceae and Podocarpaceae. For example, a ~15 amino-acid indel is present as six copies in *Podocarpus*, three copies in *Agathis*, *Araucaria* Juss., *Saxegothaea*, and *Wollemia*, and two copies in *Phyllocladus* Rich. ex Mirb. (e.g., Fig. 4). The tandem repeat (and broadly speaking the hotspot region itself) provides a microstructural synapomorphy for the clade consisting of these two families. The expansion region has a mean length of 149.6 bp (mean length of *rps*7 = 614.6 bp, SD = 69.7 bp) across the eight taxa included in Araucariaceae and Podocarpaceae, compared with *Pinus* (length of *rps*7 excluding stop codon = 465 bp). To provide some perspective, the mean total length of *rps*7 for the other taxa considered in this study is 466.4 bp, with a standard deviation of 3.37 bp. Although we have no experimental evidence

Fig. 3. Summary of bootstrap support after removal of sites classified as the two of nine fastest rate classes for 15–17 chloroplast genes and associated noncoding regions. Bootstrap values are indicated beside branches (left- and right-hand values are for MP and ML analysis, respectively; –, indicates <50% support). The topology shown is that of the best MP tree when the fastest sites are removed, with outgroups pruned for clarity.



that the expansion is part of the translated sequence in these taxa, this seems probable based on comparative evidence. The *rps*7 gene in the two families is both variable in length and sequence (particularly so in Podocarpaceae), and consistently in-frame in all taxa examined, including multiple taxa in both families that were not included here (data not shown). The portion of the gene containing the expansion region does not appear to be especially prone to indel events elsewhere in the land plants. A more comprehensive survey for the entire expansion region that includes all genera from these two families will be presented elsewhere.

Discussion

Rapidly evolving plastid DNA sites and the inference of higher-order conifer relationships

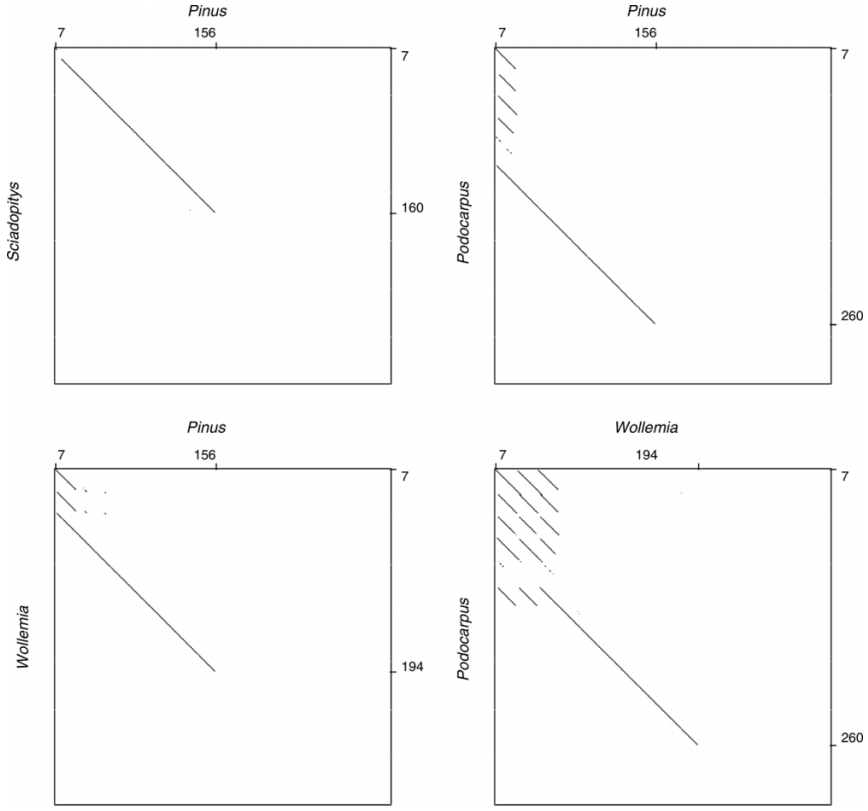
Classifying characters into different rate classes and then removing the fastest ones is a useful alternative approach to dealing with so-called saturated sites, alignment positions that may be misinformative for phylogenetic inference due to "unseen" multiple hits. In principle we might expect that ML analysis should be unaffected by removal of these rate classes, as the method should properly correct for multiple hits if the DNA substitution model is adequate (e.g., Sullivan and Swofford 2001). However, this adjustment might be expected to improve the accuracy of MP results if the amount of saturation is substantial enough to affect phylogenetic inference (e.g., Burleigh and Mathews 2004).

Removing all third-codon positions from protein-coding genes is arguably less desirable than excluding the most rapid ML rate classes, as the former approach is an overly coarse approach for correcting for multiple hits (Olmstead et al. 1998; Yang 1998; Sanderson et al. 2000), and the latter is applicable to both coding and noncoding data. We would also argue that an exclusion method based on site rates is preferable to the use of parsimony-based successive weighting methods, as in the study of conifer higher-order relationships by Quinn et al. (2002). Successive weighting has been criticized because it may lead to heuristic searches becoming trapped on local optima that depend on starting trees (Swofford et al. 1996). While the rate classification method used here may also partly depend on the starting tree, it has the potential advantage that the substitution rates used to cull the data are explicitly model-based estimates (see Olmstead et al. 1998 for a parsimony-based approach for excluding highly variable characters). However, as we find that deleting the most rapidly evolving characters has little to no effect on our major findings within the conifers (Figs. 1–3), and only a small effect on branch support, the debate could be considered moot for conifer phylogeny inference. The slight to modest reduction in bootstrap support observed within conifers after the removal of the two fastest rate classes is consistent with an expectation of increased sampling error due to fewer characters. In summary, we find no evidence here that the most rapidly evolving sites distort the inference of higher-order conifer relationships.

The ovulate cone and conifer systematics

The ovulate (seed) cone has been considered to be particularly significant in conifer systematics (e.g., Pilger 1926; Florin 1951; Miller 1999). For example, because members

Fig. 4. Dot-plot showing the pairwise similarity of complete translated sequences of the plastid *rps*7 locus from selected conifers (*Pinus*, Pinaceae; *Podocarpus*, Podocarpaceae; *Wollemia*; Araucariaceae; *Sciadopitys*, Sciadopityaceae) using a PAM-30 amino-acid substitution model, an 11-residue sliding window and a gray scale of 58%–77%.



of Taxaceae lack the "typical" compound ovulate cone of conifers, this was used to justify their recognition as Taxales, an order distinct from the remaining conifers (e.g., Florin 1951). Our data confirm the widely accepted view that Taxaceae have a nested position within the Cupressophyta clade of conifers, sister to Cephalotaxaceae (Figs. 1 and 2). If a compound ovulate cone of the sort found in Pinaceae was ancestral in extant conifers, as is usually assumed, this would require that the ovule-bearing arrangement in Taxaceae (an apparently "coneless conifer" from the perspective of its ovules) was derived by reduction from the more complex form (e.g., Chamberlain 1935; Takhtajan 1953; Hart 1987; Doyle 1998; Quinn et al. 2002). However, Tomlinson and Takaso (2002) discuss general difficulties in applying Florin's model (Florin 1951) in most families of conifers, owing to the extreme modification or apparent absence of the ovuliferous scale in these taxa (the ovuliferous scale is a condensed ovule-bearing secondary shoot axis whose underlying structure seems clearest in the ovulate cones of Pinaceae and Sciadopityaceae; Tomlinson and Takaso 2002). Hart (1987) suggested that too much weight has been placed on the compound ovulate cone in higher-order conifer systematics.

The case for recognizing *Cephalotaxus* as a member of Taxaceae

We infer *Cephalotaxus* to be the sister group of the two genera of Taxaceae sampled here, *Taxus* L. and *Torreya* Arn.; Taxaceae s. str. are monophyletic at this taxon sampling (Figs. 1 and 2). A sister-group relationship between Cephalotaxaceae and Taxaceae was first recovered by Hart (1987) using morphological data, and subsequently recovered in a morphological analysis by Doyle (1998). A *mat*K-based analysis of the two families that surveyed all five genera of Taxaceae (Cheng et al. 2000) found strong support for the monophyly of Taxaceae (96%, from MP analysis), but included only a handful of outgroups. In the more broadly based study of two plastid loci (*mat*K and *rbc*L) by Quinn et al. (2002), parsimony analysis found strong support for a clade comprising *Cephalotaxus* and Taxaceae. Their analysis did not strongly support the monophyly of Taxaceae s. str. unless the data were re-weighted using successive weighting, a method that may yield artifactual results (see above). However, it should be noted that Quinn et al. (2002) consistently recovered two strongly supported clades of Taxaceae in equally and unequally weighted analyses, one of which includes *Taxus* and the other *Torreya*. These are the two genera that we included as exemplars for the family, and which we found to make up a clade. When we removed the fastest evolving characters from phylogenetic analysis (RC7 and RC8), relationships among *Cephalotaxus*, *Taxus*, and *Torreya* are no longer strongly supported (Fig. 3).

Our main analyses support the view of Quinn et al. (2002) that it is no longer useful to recognize Cephalotaxaceae (Figs. 1 and 2); *Cephalotaxus* should be returned to its original home in Taxaceae. The rationale for a circumscription of Taxaceae that includes *Cephalotaxus* is straightforward, should the latter prove to be nested in the former, because

the more broadly defined family would then be monophyletic. However, if *Cephalotaxus* is instead shown to be the sister group of the five genera usually assumed to be in Taxaceae s. str. (i.e., *Austrotaxus* Compton, *Amentotaxus* Pilg., *Pseudotaxus* W.C. Cheng, *Taxus*, and *Torreya*), a straightforward case can also be made for reducing Cephalotaxaceae to synonymy. Backlund and Bremer (1998) have argued that higher-order classifications that recognize two families in this situation (where a small monogeneric family is the sister group of a larger one) do not optimize phylogenetic information, and ought to be considered redundant. Furthermore, the morphological distinction between *Cephalotaxus* and Taxaceae is clearly not so great that their combination would create a morphologically unrecognizable taxon (see Angiosperm Phylogeny Group (APG II) 2003). For example, *Cephalotaxus* also has a relatively simple ovule-bearing arrangement (a pair of ovules in the axil of a bract, the two separated by a narrow flange of tissue of uncertain origin; Tomlinson and Takaso 2002). A morphological connection between *Cephalotaxus* and Taxaceae s. str. is uncontroversial (Doyle 1998; Stützel and Röwekamp 1999). However, the relationship among the six genera of Taxaceae in its broadened circumscription ought to be addressed by including more taxa in phylogenetic analysis for a sampling of plastid data at least as large as that examined here.

Relationships within Cupressaceae

Relationships for three "core Cupressaceae" sampled here are in line with other studies and are well supported: *Widdringtonia* (representing the "callitroid" clade of Gadek et al. 2000) is the sister group of *Juniperus* + *Thuja* (two exemplars that represent the "cupressoid" clade of Gadek et al. 2000). Basal relationships in the family have generally not been inferred with strong support (e.g., Brunsfeld et al. 1994; Stefanovic et al. 1998; Gadek et al. 2000; Kusumi et al. 2000, although see Quinn et al. 2002). The limited taxon sampling here is congruent with these earlier studies (Figs. 1–3). As with other recent studies (Hart 1987; Stefanovic et al. 1998; Gadek et al. 2000; Kusumi et al. 2000; Quinn et al. 2002), we find that "Taxodiaceae" (represented here by *Cunninghamia*, *Metasequoia* Hu & W.C. Cheng and *Taxodium* Rich.) comprise a grade of taxa near the base of Cupressaceae. *Metasequoia* and *Taxodium* represent the "sequoioid" and "taxodioid" clades of Gadek et al. (2000), respectively; the latter is the sister group of the core Cupressaceae (Figs. 1–3). *Cunninghamia* is inferred to be the sister group of the remainder of the family, as in other recent studies with a broad taxon sampling (Gadek et al. 2000; Quinn et al. 2002). All of these relationships are well supported here.

The higher-order position of *Phyllocladus* in conifer phylogeny

Phyllocladus has sometimes been recognized as a distinct family, Phyllocladaceae, because of its highly distinctive morphology, including details of its pollen morphology, wood anatomy, and its unique (for conifers) broad leaf-like cladodes (Keng 1973, 1978; Page 1990*a*). However, it shares a relatively simple fleshy ovulate cone with the other podocarps (note that the fleshy part of the cone is not homologous across Podocarpaceae; Kelch 1997; Tomlinson and Takaso 2002), and the need to recognize it at the family level has been contested based on evidence from embryogeny and other morphological characters (Quinn 1986, 1987). We consistently find *Phyllocladus* to be the sister group of the two other taxa that we surveyed for Podocarpaceae (*Podocarpus* and *Saxegothaea*) with weak to strong support, and we consistently find strong support for the clade consisting of all three genera (Figs. 1–3). Other phylogenetic studies find *Phyllocladus* to be nested among podocarps (Hart 1987; Kelch 1997, 1998; Conran et al. 2000), sister to the rest of the family (Kelch 1998; Quinn et al. 2002; Sinclair et al. 2002), or even a grade at the base of the family (Kelch 1998), generally with poor support. However, straightforward arguments paralleling those used above to justify the recognition of *Cephalotaxus* under Taxaceae can be used to support a circumscription of the podocarps (Podocarpaceae) that includes *Phyllocladus*. Improved taxon sampling using the current plastid data set should help clarify the backbone of relationships within this large and diverse family.

The position of *Wollemia* within Araucariaceae

The recently discovered conifer *Wollemia nobilis* W.G. Jones, K.D. Hill, & J.M. Allen (Jones et al. 1995) has attracted much attention (e.g., Hogbin et al. 2000; Peakall et al. 2003) because of its status as a "living fossil" (in other words, a previously unknown living taxon that bears considerable similarity to fossil taxa). The monotypic *Wollemia* has been placed unequivocally in Araucariaceae (Jones et al. 1995), although morphological evidence on where it fits in relation to the other genera is inconclusive, since it approaches members of *Agathis* and *Araucaria* in contrasting leaf and ovulate cone characters (Chambers et al. 1998). With the exception of Setoguchi et al. (1998), who found moderate support for *Wollemia* as the sister-group of *Agathis* and *Araucaria* using *rbc*L, other studies have recovered *Wollemia* as the sister-group of *Agathis* with weak to moderate support (Gilmore and Hill 1997; Stefanovic et al. 1998; Conran et al. 2000). The placement was also strongly supported in a combined analysis of *mat*K and *rbc*L data by Quinn et al. (2002). We confirm this result here: in all analyses *Wollemia nobilis* is strongly supported as the sister group of *Agathis* (Figs. 1–3).

Significance of an expansion hotspot in the plastid ribosomal protein gene *rps*7

The morphological evidence on where the podocarps fit in higher-order conifer phylogeny is not clear (e.g., Page 1990*b*), and a close relationship between the predominantly southern hemisphere families Araucariaceae and Podocarpaceae was not firmly supported until relatively recently. Several molecular studies have demonstrated that these two families are sister taxa (Chaw et al. 1997; Stefanovic et al. 1998; Gugerli et al. 2001; Quinn et al. 2002; Rydin et al. 2002) and the phylogenies inferred here provide further support for this relationship (Figs. 1–3). The *rps*7 expansion hotspot (and the associated tandemly repeated amino-acid motif; Fig. 4) provides a microstructural synapomorphy supporting this two-family clade. Although the functional significance of this expansion region is unknown, comparable large expansions have been found in another

plastid ribosomal gene, *rps*4, in Araucariaceae and Podocarpaceae (D. Kelch, personal communication, 2007), which suggests that a survey of other plastid ribosomal protein genes might uncover additional protein structural shifts.

Seed-plant phylogeny and the position of Gnetales

We observe moderately to strongly conflicting sets of relationships here among the major seed-plant groups (cf. MP, ML analyses of the complete data set; Figs. 1 and 2). There has been extensive debate on the potential for incorrectly inferring relationships among the major groups of extant seed plants from molecular data due to systematic bias, including the question of conifer monophyly relative to a possible relationship between Pinaceae and Gnetales (e.g., Sanderson et al. 2000; Burleigh and Mathews 2004, 2007*a*, 2007*b*). [The broader issue of the monophyly of extant *and* extinct conifers is unsettled; for example, it is not clear whether voltzialean conifers such as *Emporia* are closely related to extant conifers (Rothwell and Serbet 1994; Doyle 2005).] Extinct taxa are essentially inaccessible to molecular systematists, but Burleigh and Mathews (2004) suggested that better taxon sampling of extant taxa might help reduce the observed conflict among studies regarding broad seed-plant relationships. Our substantially improved taxon sampling within conifers (compared with Rai et al. 2003) does not yield a clearer answer for seed-plant relationships as a whole, although it is possible that this picture will change with additional conifer sampling.

Removing the faster characters (RC78), which might be expected to reduce systematic bias, results in generally poorer support for relationships among the major nonconifer seed plant clades in MP and ML analyses (data not shown). For example, in the analyses of the full data set, the bootstrap support for the clade that is the sister group of Gnetales (which consists of angiosperms, conifers, cycads, and *Ginkgo*; Figs. 1 and 2) is 100% from MP analysis and 86% from ML analysis. These values fall to 79%, and 44%, respectively, with the fastest characters removed. We reexamined our bootstrap profiles to determine the levels of support for alternative relationships involving Gnetales for the full and reduced data set. We found that bootstrap support for the hypothetical gnepine clade from our MP and ML analyses is <1% and 9% (respectively) with all data included, versus 7% and 12% with the RC78 sites removed. Bootstrap support for even a loose version of the "gnetifer" hypothesis (i.e., with Gnetales and conifers in a clade, without regard to the monophyly of either) is weak at best. There is <1% and 14% bootstrap support for this clade (from MP and ML analyses, respectively) when all data are included. The latter support values show modest improvement when RC78 sites are removed, with 22% and 50% bootstrap support for this weak version of the gnetifer hypothesis from our MP and ML analyses, respectively. We will address the broader issue of seed-plant relationships more fully elsewhere.

The inference of conifer phylogeny from plastid data

Considering either the entire data set or the reduced subset of it, our MP and ML analyses support Pinaceae as the sister group of the rest of the conifers, with or without the most rapidly evolving characters included (Figs. 1–3). We therefore find moderate to strong support for conifer monophyly (setting aside for now the possibility that the gnepine hypothesis is correct but not recovered here owing to strong systematic bias; see above). We recover strong bootstrap support for the broad backbone of conifer phylogeny, comparable with or better than that found in other studies with a broad sampling of conifers (e.g., Stefanovic et al. 1998; Quinn et al. 2002). This is in line with theoretical expectations that increasing the amount of data per taxon should reduce the effect of sampling error on phylogenetic inference. Further increasing the taxon sampling within the major clades of conifers for the plastid gene set examined here, or others of comparable size, may help address the subset of relationships that we did not infer with strong support (i.e., relationships within Pinaceae, Podocarpaceae, and Taxaceae s.l.; Figs. 1–3); see Hillis (1998) for a rationale. Adding genes may also help address some of the hard-to-resolve branches within conifers (e.g., within Pinaceae) and some of the tougher questions involving the major groups of seed plants (e.g., concerning Gnetales placement). It is becoming reasonably straightforward, for example, to obtain plastid data sets of the order of size of the whole plastid genome (e.g., Leebens-Mack et al. 2005). However, the current gene sample is clearly sufficient to recover strong bootstrap support for most of the higher-order relationships that we address in conifers (Figs. 1 and 2). Indeed, even the relatively small set of characters that we infer to be among the most slowly evolving (i.e., RC0-6) provides excellent support for almost the entire broad backbone of conifer phylogeny (Fig. 3).

Acknowledgements

We thank D.M. Cherniawsky, A. Colwell, K. Ikegama, and J. Metropulos and others for assistance in obtaining plant material, and two anonymous reviewers for their comments. This research was funded by an Natural Sciences and Engineering Research Council of Canada (NSERC) postgraduate scholarship to H.S.R., National Science Foundation grant DEB 0090313 to R.G.O., and an NSERC Discovery grant to S.W.G.

References

Angiosperm Phylogeny Group (APG II). 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. Bot. J. Linn. Soc. **141**: 399–436. doi:10.1046/j.1095-8339.2003.t01-1-00158.x.

Backlund, A., and Bremer, K. 1998. To be or not to be — principles of classification and monotypic plant families. Taxon, **47**: 391–400. doi:10.2307/1223768.

Bowe, L.M., Coat, G., and dePamphilis, C.W. 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. U.S.A. **97**: 4092–4097. doi:10.1073/pnas.97.8.4092. PMID:10760278.

Brunsfeld, S.J., Soltis, P.E., Soltis, D.E., Gadek, P.A., Quinn, C.J., Strenge, D.D., and Ranker, T.A. 1994. Phylogenetic relationships among the genera of Taxodiaceae and Cupressaceae: evidence from *rbc*L sequences. Syst. Bot. **19**: 253–262. doi:10.2307/2419600.

Burleigh, J.G., and Mathews, S. 2004. Phylogenetic signal in nucleotide data from seed plants: implications for resolving the

seed plant tree of life. Am. J. Bot. **91**: 1599–1613. doi:10.3732/ajb.91.10.1599.

Burleigh, J.G., and Mathews, S. 2007*a*. Assessing among-locus variation in the inference of seed plant phylogeny. Int. J. Plant Sci. **168**: 111–124. doi:10.1086/509586.

Burleigh, J.G., and Mathews, S. 2007*b*. Assessing systematic error in the inference of seed plant phylogeny. Int. J. Plant Sci. **168**: 125–135. doi:10.1086/509588.

Cantino, P.D., Doyle, J.A., Graham, S.W., Judd, W.S., Olmstead, R.G., Soltis, D.E., Soltis, P.S., and Donoghue, M.J. 2007. Towards a phylogenetic nomenclature of *Tracheophyta.* Taxon, **56**: 822–846.

Chamberlain, C.J. 1935. Gymnosperms, structure and evolution. Chicago University Press, Chicago, Ill.

Chambers, T.C., Drinnan, A.N., and McLoughlin, S. 1998. Some morphological features of Wollemi Pine (*Wollemia nobilis*: Araucariaceae) and their comparison to cretaceous plant fossils. Int. J. Plant Sci. **159**: 160–171. doi:10.1086/297534.

Chaw, S.-M., Zharkikh, A., Sung, H.-M., Lau, T.-C., and Li, W.-H. 1997. Molecular phylogeny of extant gymnosperms and seed plant evolution: Analysis of nuclear 18S rRNA sequences. Mol. Biol. Evol. **14**: 56–68. PMID:9000754.

Chaw, S.-M., Parkinson, C.L., Cheng, Y., Vincent, T.M., and Palmer, J.D. 2000. Seed plant phylogeny inferred from all three plant genomes: monophyly of extant gymnosperms and origin of Gnetales from conifers. Proc. Natl. Acad. Sci. U.S.A. **97**: 4086–4091. doi:10.1073/pnas.97.8.4086.

Cheng, Y., Nicolson, R.G., Tripp, K., and Chaw, S.-M. 2000. Phylogeny of Taxaceae and Cephalotaxaceae genera inferred from chloroplast *mat*K gene and nuclear rDNA ITS region. Mol. Phylogenet. Evol. **14**: 353–365. doi:10.1006/mpev.1999.0710. PMID:10712841.

Conran, J.G., Wood, G.M., Martin, P.G., Dowd, J.M., Quinn, C.J., Gadek, P.A., and Price, R.A. 2000. Generic relationships within and between the gymnosperm families Podocarpaceae and Phyllocladaceae based on an analysis of the chloroplast gene *rbc*L. Aust. J. Bot. **48**: 715–724. doi:10.1071/BT99062.

Donoghue, M.J., and Doyle, J.A. 2000. Seed plant phylogeny: demise of the anthophyte hypothesis? Curr. Biol. **10**: R106–R109. doi:10.1016/S0960-9822(00)00304-3. PMID:10679315.

Doyle, J.A. 1998. Phylogeny of vascular plants. Annu. Rev. Ecol. Syst. **29**: 567–599. doi:10.1146/annurev.ecolsys.29.1.567.

Doyle, J.A. 2005. Seed ferns and the origin of angiosperms. J. Torrey Bot. Soc. **133**: 169–209.

Doyle, J.J., and Doyle, J.L. 1987. A rapid isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. **19**: 11–15.

Eckenwalder, J.E. 1976. Re-evaluation of Cupressaceae and Taxodiaceae: A proposed merger. Madrono, **23**: 237–256.

Enright, N.J., and Hill, R.S. (*Editors*). 1995. Ecology of the southern conifers. Smithsonian Institution Press, Washington, D.C.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution, **39**: 783–791. doi:10.2307/2408678.

Florin, R. 1951. Evolution in cordaites and conifers. Acta Horti Bergiani, **15**: 285–388.

Gadek, P.A., Alpers, D.L., Heslewood, M.M., and Quinn, C.J. 2000. Relationships within Cupressaceae sensu lato: A combined morphological and molecular approach. Am. J. Bot. **87**: 1044–1057. doi:10.2307/2657004. PMID:10898782.

Gilmore, S., and Hill, K.D. 1997. Relationships of the Wollemi Pine (*Wollemia nobilis*) and a molecular phylogeny of the Araucariaceae. Telopea (Syd.), **7**: 275–291.

Graham, S.W., and Olmstead, R.G. 2000*a*. Utility of 17 chloroplast genes for inferring the phylogeny of the basal angiosperms. Am. J. Bot. **87**: 1712–1730. doi:10.2307/2656749. PMID:11080123.

Graham, S.W., and Olmstead, R.G. 2000*b*. Evolutionary significance of an unusual chloroplast DNA inversion found in two basal angiosperm lineages. Curr. Genet. **37**: 183–188. doi:10.1007/s002940050517. PMID:10794175.

Graham, S.W., Kohn, J.R., Morton, B.R., Eckenwalder, J.E., and Barrett, S.C.H. 1998. Phylogenetic congruence and discordance among one morphological and three molecular data sets from Pontederiaceae. Syst. Biol. **47**: 545–567. doi:10.1080/106351598260572. PMID:12066301.

Graham, S.W., Reeves, P.A., Burns, A.C.E., and Olmstead, R.G. 2000. Microstructural changes in noncoding chloroplast DNA: interpretation, evolution, and utility of indels and inversions in basal angiosperm phylogenetic inference. Int. J. Plant Sci. **161**: S83–S96. doi:10.1086/317583.

Graham, S.W., Zgurski, J.M., McPherson, M.A., Cherniawski, D.M., Saarela, J.M., Horne, E.F.C., Smith, S.Y., Wong, W.A., O'Brien, H.E., Biron, V.L., Pires, J.C., Olmstead, R.G., Chase, M.W., and Rai, H.S. 2006. Robust inference of monocot deep phylogeny using an expanded multigene plastid data set. *In* Monocots: comparative biology and evolution (excluding Poales) *Edited by* J.T. Columbus, E.A. Friar, J.M. Porter, L.M. Prince, and M.G. Simpson. Rancho Santa Ana Botanic Garden, Claremont, Calif. pp. 3–21.

Gugerli, F., Sperisen, C., Büchler, U., Brunner, I., Brodbeck, S., Palmer, J.D., and Qui, Y.-L. 2001. The evolutionary split of Pinaceae from other conifers: evidence from an intron loss and a multigene phylogeny. Mol. Phylogenet. Evol. **21**: 167–175. doi:10.1006/mpev.2001.1004. PMID:11697913.

Guindon, S., and Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. **52**: 696–704. doi:10.1080/10635150390235520. PMID:14530136.

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

Hillis, D. 1998. Taxonomic sampling, phylogenetic accuracy, and investigator bias. Syst. Biol. **42**: 182–192. doi:10.2307/2992540.

Hogbin, P.M., Peakall, R., and Sydes, M.A. 2000. Achieving practical outcomes from genetic studies of rare plants. Aust. J. Bot. **48**: 375–382. doi:10.1071/BT98080.

Jones, W.G., Hill, K.D., and Allen, J.M. 1995. *Wollemia nobilis*, a new living Australian genus and species in the Araucariaceae. Telopea (Syd.), **6**: 173–176.

Junier, T., and Pagni, M. 2000. Dotlet: diagonal plots in a web browser. Bioinformatics, **16**: 178–179. doi:10.1093/bioinformatics/16.2.178.

Kelch, D.G. 1997. The phylogeny of the Podocarpaceae based on morphological evidence. Syst. Bot. **22**: 113–131. doi:10.2307/2419680.

Kelch, D.G. 1998. Phylogeny of Podocarpaceae: comparison of evidence from morphology and 18S rDNA. Am. J. Bot. **85**: 986–996. doi:10.2307/2446365.

Keng, H. 1973. On the family Phyllocladaceae. Taiwania, **18**: 142–145.

Keng, H. 1978. The genus *Phyllocladus* (Phyllocladaceae). J. Arnold. Arbor. Harv. Univ. **59**: 249–273.

Kosakovsky Pond, S.L., Frost, S.D.W., and Muse, S.V. 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics, **21**: 676–679. doi:10.1093/bioinformatics/bti079. PMID:15509596.

Kusumi, J., Tsumura, Y., Yoshimaru, H., and Tachida, H. 2000. Phylogenetic relationships in Taxodiaceae and Cupressaceae sensu stricto based on *mat*K gene, *chl*L gene, *trn*L-*trn*F IGS region, and *trn*L intron sequences. Am. J. Bot. **87**: 1480–1488. doi:10.2307/2656874. PMID:11034923.

Leebens-Mack, J., Raubeson, L.A., Cui, L., Kuehl, J.V., Fourcade, M.H., Chumley, T.W., Boore, J.L., Jansen, R.K., and dePamphilis, C.W. 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. doi:10.1093/molbev/msi191. PMID:15944438.

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

Mundry, M., and Stützel, T. 2004. Morphogenesis of the reproductive shoots of *Welwitschia mirabilis* and *Ephedra distachya* (Gnetales), and its evolutionary implications. Org. Divers. Evol. **4**: 91–108. doi:10.1016/j.ode.2004.01.002.

Olmstead, R.G., Reeves, P.A., and Yen, A.C. 1998. Patterns of sequence evolution and implications for parsimony analysis of chloroplast DNA. *In* Molecular systematics of plants II: DNA sequencing. *Edited by* P.S. Soltis, D.E. Soltis, and J.J. Doyle. Kluwer, Boston, Mass. pp. 164–187.

Page, C.N. 1990*a*. Phyllocladaceae. *In* The families and genera of vascular plants Vol. 1. *Edited by* K.U. Kramer and P.S. Green. Springer-Verlag, Berlin, Germany. pp. 317–319.

Page, C.N. 1990*b*. Podocarpaceae. *In* The families and genera of vascular plants Vol. 1. *Edited by* K.U. Kramer and P.S. Green. Springer-Verlag, Berlin, Germany. pp. 332–346.

Page, C.N. 1990*c*. Sciadopityaceae. *In* The families and genera of vascular plants Vol. 1. *Edited by* K.U. Kramer and P.S. Green. Springer-Verlag, Berlin, Germany. pp. 346–348.

Page, C.N. 1990*d*. Taxaceae. *In* The families and genera of vascular plants Vol. 1. *Edited by* K.U. Kramer and P.S. Green. Springer-Verlag, Berlin, Germany. pp. 346–353.

Peakall, R., Ebert, D., Scott, L., Meagher, P., and Offord, C. 2003. Comparative genetic study confirms exceptionally low genetic variation in the ancient and endangered relictual conifer, *Wollemia nobilis* (Araucariaceae). Mol. Ecol. **12**: 2331–2343. doi:10.1046/j.1365-294X.2003.01926.x. PMID:12919472.

Pilger, R. 1926. Coniferae. *In* Die Natürlichen Pflanzenfamilien. 2nd ed. *Edited by* A. Engler and K. Prantl. W. Engelmann, Leipzig, Germany. pp. 121–407.

Posada, D., and Crandall, K.A. 1998. Modeltest: testing the model of DNA substitution. Bioinformatics, **14**: 817–818. doi:10.1093/bioinformatics/14.9.817. PMID:9918953.

Quinn, C.J. 1986. Embryogeny in *Phyllocladus.* N.Z. J. Bot. **24**: 575–580.

Quinn, C.J. 1987. The Phyllocladaceae Keng — a critique. Taxon, **36**: 559–565. doi:10.2307/1221846.

Quinn, C.J., Price, R.A., and Gadek, P.A. 2002. Familial concepts and relationships in the conifers based on *rbcL* and *matK* sequence comparisons. Kew Bull. **57**: 513–531. doi:10.2307/4110984.

Rai, H.S., O'Brien, H.E., Reeves, P.A., and Graham, S.W. 2003. Inference of higher-order relationships in the cycads from a large chloroplast data set. Mol. Phylogenet. Evol. **29**: 350–359. doi:10.1016/S1055-7903(03)00131-3. PMID:13678689.

Rambaut, A. 1998. "Se-Al (Sequence Alignment Editor Version 1.0)," Computer program and documentation. Department of Zoology, University of Oxford, UK.

Raubeson, L.A., and Jansen, R.K. 1992. A rare chloroplast-DNA structural mutation is shared by all conifers. Biochem. Syst. Ecol. **20**: 17–24. doi:10.1016/0305-1978(92)90067-N.

Rothwell, G.W., and Serbet, R. 1994. Lignophyte phylogeny and the evolution of spermatophytes: a numerical cladistic analysis. Syst. Bot. **19**: 443–482. doi:10.2307/2419767.

Rydin, C., Källersjö, M., and Friis, E.M. 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. doi:10.1086/338321.

Saarela, J.M., Rai, H.S., Doyle, J.A., Endress, P.K., Mathews, S., Marchant, A.D., Briggs, B.G., and Graham, S.W. 2007. Hydatellaceae identified as a new branch near the base of the angiosperm phylogenetic tree. Nature (London), **446**: 312–315. doi:10.1038/nature05612. PMID:17361182.

Sanderson, M.J., Wojciechowski, M.F., Hu, J.M., Sher Khan, T., and Brady, S.G. 2000. Error, bias, and long-branch attraction in data for two chloroplast photosystem genes in seed plants. Mol. Biol. Evol. **17**: 782–797. PMID:10779539.

Setoguchi, H., Osawa, T.A., Pintaud, J.-C., Jaffré, T., and Veillon, J.-M. 1998. Phylogenetic relationships within Araucariaceae based on *rbc*L gene sequences. Am. J. Bot. **85**: 1507–1516. doi:10.2307/2446478.

Sinclair, W.T., Mill, R.R., Gardner, M.F., Woltz, P., Jaffré, T., Preston, J., Hollingsworth, M.L., Ponge, A., and Möller, M. 2002. Evolutionary relationships of the New Caledonian heterotrophic conifer, *Parasitaxus usta* (Podocarpaceae), inferred from chloroplast *trn*L-F intron/spacer and nuclear rDNA ITS2 sequences. Plant Syst. Evol. **233**: 79–104. doi:10.1007/s00606-002-0199-8.

Stefanovic, S., Jager, M., Deutsch, J., Broutin, J., and Masselot, M. 1998. Phylogenetic relationships of conifers inferred from partial 28S rRNA gene sequences. Am. J. Bot. **85**: 688–697. doi:10.2307/2446539.

Stockey, R.A., Kvacek, J., Hill, R.S., Rothwell, G.W., and Kvacek, K. 2005. The fossil record of Cupressaceae *s. lat*. *In* A monograph of Cupressaceae and *Sciadopitys. Edited by* A. Farjon. Royal Botanic Gardens, Kew, Richmond, Surrey, UK. pp. 64–68.

Stützel, T., and Röwekamp, I. 1999. Female reproductive structures in Taxales. Flora, **194**: 145–157.

Sullivan, J., and Swofford, D.L. 2001. Should we use model-based methods for phylogenetic inference when we know that assumptions about among-site rate variation and nucleotide substitution patterns are violated? Syst. Biol. **50**: 723–729. doi:10.1080/106351501753328848. PMID:12116942.

Swofford, D.L. 2002. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Computer program and documentation. Sinauer Associates, Sunderland, Mass.

Swofford, D.L., Olsen, G.J., Waddell, P.J., and Hillis, D.M. 1996. Phylogenetic inference. *In* Molecular systematics. 2nd ed. *Edited by* D.M. Hillis, C. Moritz, and B.K. Mable. Sinauer Associates, Sunderland, Mass. pp. 407–514

Takhtajan, A.L. 1953. Phylogenetic principles of the system of higher plants. Bot. Rev. **19**: 1–45. doi:10.1007/BF02861725.

Tomlinson, P.B., and Takaso, T. 2002. Seed cone structure in conifers in relation to development and pollination: a biological approach. Can. J. Bot. **80**: 1250–1273. doi:10.1139/b02-112.

Wakasugi, T., Tsudzuki, J., Ito, S., Nakashima, K., Tsudzuki, T., and Sugiura, M. 1994. Loss of all *ndh* genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. Proc. Natl. Acad. Sci. U.S.A. **91**: 9794–9798. doi:10.1073/pnas.91.21.9794. PMID:7937893.

Yang, Z. 1998. On the best evolutionary rate for phylogenetic analysis. Syst. Biol. **47**: 125–133. doi:10.1080/106351598261067. PMID:12064232.

Yao, X., Taylor, T.N., and Taylor, E.L. 1997. A taxodiaceous seed cone from the Triassic of Antarctica. Am. J. Bot. **84**: 343–354. doi:10.2307/2446008.

Zgurski, J.M., Rai, H.S., Fai, Q.M., Bogler, D.J., Francisco-Ortega, J., and Graham, S.W. 2008. How well do we understand the overall backbone of cycad phylogeny? New insights from a large, multigene plastid data set. Mol. Phylogenet. Evol. **47**: 1232–1237. doi:10.1016/j.ympev.2008.03.002.