RESEARCH ARTICLE

Loss of all plastid *ndh* genes in Gnetales and conifers: extent and evolutionary significance for the seed plant phylogeny

**Thomas Werner Anthony Braukmann ·
Maria Kuzmina · Saša Stefanović**

Abstract The exact phylogenetic position of Gnetales, a small, highly modified group of gymnosperms with an accelerated rate of molecular evolution, is one of the most challenging issues for seed plant systematics. Recent results from entire plastid genome (ptDNA) sequencing revealed the absence of the entire suite of plastid *ndh* genes in several species of Gnetales and the pine family (Pinaceae) potentially highlighting a major structural feature linking these two groups—concerted loss of all plastid genes for the NADH dehydrogenase complex. However, the precise extent of *ndh* gene loss in gymnosperms has not been surveyed. Using a slot-blot hybridization method, we probed all 11 *ndh* genes in 162 species from 70 of 85 gymnosperm genera. We find that all *ndh* genes are absent across Gnetales and Pinaceae, but not in any other group of gymnosperms. This feature represents either a major synapomorphy for a clade consisting of these two lineages or, less likely, a convergent loss. Our survey substantially extends previous inferences based on more limited sampling and, if the former evolutionary interpretation is correct, it provides additional support for the contentious "gnepine" hypothesis, which places Gnetales as sister to Pinaceae.

Communicated by R. Bock.

T. W. A. Braukmann (✉) · M. Kuzmina · S. Stefanović
Department of Biology, University of Toronto Mississauga,
3359 Mississauga Rd. N, Mississauga, ON L5L 1C6, Canada
e-mail: thomas.braukmann@utoronto.ca

Present Address:
M. Kuzmina
Biodiversity Institute of Ontario, University of Guelph,
579 Gordon St., Guelph, ON N1G 2W1, Canada

Introduction

Arguably the most vexing outstanding question for the seed plant phylogeny remains the placement of Gnetales, a small group of gymnosperms currently circumscribed in three genera (*Ephedra*, *Gnetum*, and *Welwitschia*), despite intensive cladistic investigations over past two decades. Multiple permutations for the relationship of Gnetales to other extant lineages of spermatophytes (otherwise comprising *Ginkgo*, cycads, Pinaceae, cupressophyte conifers and flowering plants) have been proposed by different studies, using different types of data and/or different analytical methods (reviewed by Magallón and Sanderson 2002; Soltis et al. 2002; Burleigh and Mathews 2004, 2007a; Mathews 2009). However, because of their repeated recovery across many studies, six particular hypotheses on the placement of Gnetales stand out as the most frequently invoked ones (Fig. 1). A series of early cladistic analyses based on morphological data and including relevant fossil taxa suggested the placement of Gnetales close to angiosperms (Crane 1985; Doyle and Donoghue 1986; Loconte and Stevenson 1990; Nixon et al. 1994), in accordance with the "anthophyte" hypothesis (Wettstein 1907). Later, some molecular studies recovered the same results, albeit only with weak support (Stefanović et al. 1998; Rydin et al. 2002). However, most early analyses of molecular data resulted in so-called "gnetales-sister" hypotheses, featuring Gnetales either sister to the rest of gymnosperms ("gnetales-sister I"; Hasebe et al.
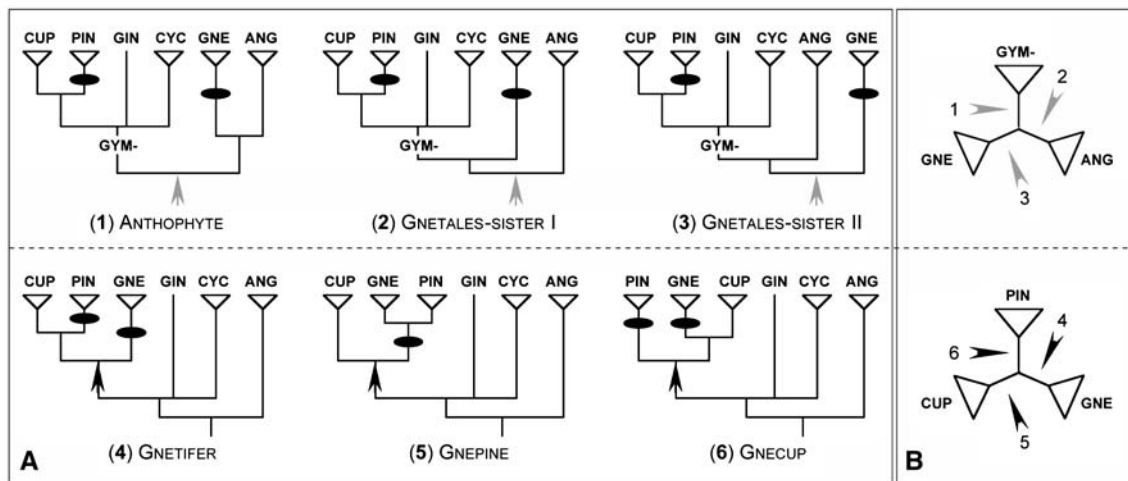
Fig. 1 Correspondence between competing rooted and unrooted seed plant phylogenetic hypotheses. **a** Six most prominent rooted hypotheses of relationships among major lineages of seed plants. *Arrows* indicate alternative root placements as in **b**, *1–6* Parsimony reconstructions of gene losses for the *ndh* gene suite are mapped (oval) under the assumption of irreversibility. **b** Unrooted trees showing that the underlying topologies for each of the two major sets of hypotheses are identical when their respective outgroups, the other land plants (*1–3*) or other seed plants (*4–6*), are excluded. Note that these two unrooted trees are incompatible. Taxon abbreviations: *ANG* angiosperms, *CUP* cupressophytes, *CYC* cycads, *GIN Ginkgo*, *GNE* Gnetales, *GYM–* gymnosperms minus Gnetales, *PIN* Pinaceae

1992; Goremykin et al. 1996; Samigullin et al. 1999; Frolich and Parker 2000; Mathews and Donoghue 2000) or sister to the rest of seed plants, including angiosperms ("gnetales-sister II"; Hamby and Zimmer 1992; Albert et al. 1994; Rai et al. 2003, 2008). Subsequent analyses, based mainly on sequences from multiple genes and genomes, suggested a "gnetifer" hypothesis, in which Gnetales are found as sister to all conifers (Chaw et al. 1997, 2000). But one of the most surprising results of molecular plant systematics placed Gnetales *within* conifers, as sister to Pinaceae, in what became known as the "gnepine" hypothesis (Bowe et al. 2000; Chaw et al. 2000; Gugerli et al. 2001; Qiu et al. 2007). More recently, another hypothesis rendering conifers paraphyletic has gained some momentum. According to the "gnecup" hypothesis, Gnetales are also placed within conifers, but as sister to cupressophytes, not Pinaceae (see Nickrent et al. 2000; Rydin and Källersjö 2002; Doyle 2006; Chumley et al. 2008).

One major cause of ambiguity for spermatophyte phylogeny inference can be attributed to the ambiguous rooting. While these six competing hypotheses seem significantly different from one another when polarized with outgroups (i.e., rooted; Fig. 1a), they fall into only two categories, each with an identical underlying unrooted tree (Fig. 1b). For example, it becomes apparent that there are no topological differences among the first three major phylogenetic hypotheses (anthophyte and gnetales-sister, I and II; Fig. 1a, 1–3) when the other land plant outgroups (represented by grey arrows in Fig. 1) are pruned off and the remainder is taken as an unrooted tree. Similarly, the only

difference between the remaining three alternative scenarios within seed plants (gnetifer, gnepine, and gnecup hypotheses; Fig. 1a, 4–6), is the placement of the root for the clade comprising conifers and Gnetales. Besides rooting issues, another important source of ambiguity is due to the substantial sequence divergence among living lineages of seed plants and their nearest outgroups (ferns and lycopods). In seed plants, this is particularly evident in long branches leading to angiosperms, Gnetales, and Pinaceae as seen in most molecular trees, regardless of whether derived from plastid, nuclear, or mitochondrial sequences (Chaw et al. 1997, 2000; Bowe et al. 2000; Rydin et al. 2002; Graham and Iles 2009). In conjunction, these two phenomena frequently lead to the strongly supported yet spurious tree rooting due to long-branch attraction or related artifacts (Felsenstein 1983; Hendy and Penny 1989). Further complicating the issue, different methods of analyses and different optimality criteria frequently support alternative topologies, even though based on the same sequences (Bowe et al. 2000; Rydin et al. 2002; Burleigh and Mathews 2004, 2007b). Surprisingly, the removal of most rapidly evolving sequences or sites was shown to have little to no effect on the inferences on higher-order seed plant relationships (Burleigh and Mathews 2004; Rai et al. 2008), raising the question of the limits of nucleotide data (Mathews 2009).

Molecular evidence is not limited only to primary sequence data. Additional sources of molecular data relevant for the number of open questions of seed plant phylogeny can be sought from so-called "rare genomic markers" (e.g., Raubeson and Jansen 1992; Rokas and Holland 2000;

Moreira and Philippe 2000; Gugerli et al. 2001). In this regard, the plastid (pt) genome seems to be particularly promising. In seed plants, this genome is highly conserved in size, structure, content, and synteny (Palmer 1991; Downie et al. 1991; Clegg et al. 1994). Because of their relatively infrequent evolutionary occurrence, any major structural mutation in the pt genome, such as inversions, gene/intron losses, and contractions/expansions of the inverted repeat (IR), are often considered to be more reliable phylogenetic markers compared to sequences (e.g., Downie et al. 1991; Downie and Palmer 1992; Raubeson and Jansen 1992; Doyle et al. 1995; Doyle et al. 1996; Bailey et al. 1997; Graham and Olmstead 2000a; Plunkett and Downie 2000; Jansen et al. 2007; but see McPherson et al. 2004; Palmer et al. 2004 for cautionary views).

The loss of *ndh* genes from the pt genome of some seed plant represents one such potentially informative structural change. The *ndh* genes encode subunits of the plastid NAD(P)H-dehydrogenase (Ndh) complex, a homologue of mitochondrial complex I (Shinozaki et al. 1986). In plastids, the Ndh complex seems to be primarily involved with transfer of electrons from stromal reductants to a plastoquinone pool, a process commonly known as "chlororespiration" (Bennoun 2002; Peltier and Cournac 2002). In addition, involvement of this complex in photooxidative stress reduction in high light intensity, regulation of photosynthesis by modulating the activity of cyclic electron flow around photosystem I, and/or leaf senescence regulation have also been suggested (Casano et al. 2001; Bukhov and Carpentier 2004; Zapata et al. 2005; Diaz et al. 2007; Tallon and Quiles 2007; Romeau et al. 2007; Endo et al. 2008).

While the precise role of the Ndh complex is still uncertain, the *ndh* genes are known to be widespread in the autotrophic seed plants and remain highly conserved over large evolutionary distances, indicating the presence of strong selection pressure for their retention. Results of entire (or extensive) pt genome sequencing from a number of individual studies indicate that the complete suite of *ndh* genes is present in *Ginkgo* (Leebens-Mack et al. 2005; Jansen et al. 2007) and selected representatives of cycads (Wu et al. 2007) and cupressophytes (Hirao et al. 2008) as well as in ~60 species of diverse angiosperms (summarized most recently by Jansen et al. 2007). As suggested by Bungard (2004), the loss of *ndh* genes in flowering plants seems to be confined only to parasitic plant lineages (dePamphilis and Palmer 1990; Olmstead et al. 2001; Stefanović and Olmstead 2005; Funk et al. 2007; McNeal et al. 2007). This link with heterotrophy is further supported by the loss of *ndh* genes in a non-photosynthetic liverwort (Wickett et al. 2008) and a green but potentially mycotrophic orchid (Chang et al. 2006). The absence of functional *ndh* genes from the pt genomes of fully autotrophic seed plants is

presently reported only from Gnetales (Wu et al. 2007, 2009; McCoy et al. 2008) and several genera of Pinaceae (Wakasugi et al. 1994; Cronn et al. 2008; Rai et al. 2008; Wu et al. 2009). If inferred to have happened concurrently, this loss could represent a strong synapomorphy for Gnetales and Pinaceae (Chaw et al. 2000; Burleigh and Mathews 2004; Wu et al. 2007). Despite the potential of this rare structural genomic character to bear significantly on the seed plant relationships and help choose among alternative phylogenetic hypotheses, the full extent of presence or absence of *ndh* genes among living gymnosperms is unknown.

In the present study, we gathered data using a comprehensive slot-blot hybridization survey of the complete suite of plastid *ndh* genes with a dense sampling of gymnosperms, the most extensive data matrix applied to this issue to date, in order to: (1) ascertain the extent and distribution of *ndh* gene losses across gymnosperms; and (2) assess the utility of these losses as phylogenetic markers for seed plant phylogeny.

Materials and methods

Taxon sampling

In total, 70 of the 85 genera and 162 of the ~1,070 species of extant gymnosperms were sampled in this study (Table 1), corresponding to 82% of their generic- and 15% of their species-richness, respectively. Our sampling encompasses all four major lineages of living gymnosperms; however, the percentage of diversity coverage differs among these groups (compare with Table 1). *Ginkgo biloba* is the sole living representative of Ginkgoales. Conifers, including both Pinaceae and cupressophytes, are represented by a total of 131 species (out of 680; 20%), grouped into 59 genera (out of 70; 85%), from all seven currently recognized families. The species-richness of cycads is represented to a significantly lesser degree, by 14 out of 305 species (5%), but our sampling covers 64% of genera (7 out of 11). Finally, 16 species of Gnetales are sampled in total (out of the 92 species; 17%), including all three recognized genera/families. Representatives of four genera of autotrophic angiosperms (Table 1), the last remaining lineage of extant seed plants, were also included in our surveys as positive controls.

DNA extraction and hybridization

Total genomic DNA was isolated using the modified 2× CTAB method (Doyle and Doyle 1987). Fresh leaf material was used where available, but approximately two-thirds of the samples were either from silica-gel dried tissue or

Table 1 Seed plant taxa surveyed for the presence/absence of plastid *ndh(A-K)* genes

Taxon	Voucher	ndhA		ndhB		ndhC	ndhD		ndhE	ndhF	ndhG	ndhH		ndhI	ndhJ	ndhK
		exon 1	exon 2	exon 1	exon 2		5′ end	3′ end				5′ end	3′ end			
Ginkgoales [1/1; 100%]																
Ginkgo biloba L.	*QIU-94015* (IND)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++
Cycadales [7/11; ~ 64%]																
Cycas media R. Br.	No voucher	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++
Cycas revoluta Thunb.	*QIU-94051* (IND)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++
Cycas taiwaniana Carruth.	*SS-05-53* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	+	+	++
Dioon spinulosum Dyer ex Eichl.	*SS-03-72* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	+	+	+	++
Encephalartos altensteinii Lehm.	*SS-05-57* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	+	+	+	++
Encephalartos barteri Carruth. ex Miq.	*SS-05-118* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	+	+	+	++
Encephalartos ferox Bertol. f.	*SS-05-116* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++
Encephalartos longifolius Lehm.	*QIU-95127* (IND)	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++
Macrozamia communis L.A.S. Johnson	*SS-05-119* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	+	+	+	++
Ceratozamia mexicana Brongn.	*QIU-94009* (IND)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++
Stangeria eriopus (Kunze) Baill.	*QIU-94078* (IND)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++
Zamia floridana A.DC.	*QIU-95035* (IND)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++
Zamia furfuracea Ait.	*SS-05-56* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++
Zamia pumila L.	*SS-05-58* (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++
Gnetales [3/3; 100%]																
Ephedra antisyphilitica Berland ex C.A. Mey.	*QIU-94004* (IND)	+	+	–	–	+	–	–	+	–	–	+	–	–	–	–
Ephedra viridis Coville	*SS-05-21* (TRTE)	–	+	–	–	+	–	–	+	–	–	+	–	–	+	–
Ephedra triandra Tul.	*SS-05-100* (TRTE)	–	+	–	–	+	–	–	+	–	–	+	–	+	+	–
Ephedra saxatilis (Stapf) Royle ex Florin	*SS-05-188* (TRTE)	–	+	–	–	+	–	–	+	–	–	+	–	–	–	–
Ephedra monosperma J.G. Gmel. ex C.A. Mey.	*SS-05-35* (TRTE)	+	+	–	–	+	+	–	+	–	–	+	–	+	+	–
Ephedra sinica Stapf	*SS-05-112* (TRTE)	+	+	–	–	+	–	–	+	–	–	+	–	+	–	–
Ephedra distachya L.	*SS-05-25* (TRTE)	–	+	–	–	+	–	–	+	–	–	+	–	+	–	–
Ephedra campylopoda C.A. Mey.	*SS-05-60* (TRTE)	–	+	–	–	+	–	–	+	–	–	+	–	–	–	–
Ephedra fragilis Fresen.	*SS-05-61* (TRTE)	+	+	–	–	–	–	–	+	–	–	+	–	+	–	–
Ephedra minuta Florin	*SS-05-181* (TRTE)	–	–	–	–	+	–	–	+	–	–	+	–	–	–	–
Ephedra procera C.A. Mey.	*SS-05-34* (TRTE)	–	+	–	–	+	–	–	+	–	–	+	–	–	–	–
Gnetum gnemon L.	*QIU-94005* (IND)	+	+	–	+	+	–	–	+	–	–	+	–	–	–	–
Gnetum leyboldii Tul.	*QIU-94017* (IND)	+	+	–	–	+	–	–	+	–	+	+	–	+	–	–
Gnetum sp.	*SS-04-119* (TRTE)	+	+	–	–	+	–	–	+	–	–	+	–	+	–	–
Gnetum ula Brongn.	*SS-06-214* (TRTE)	+	+	–	–	+	–	–	+	–	–	+	–	–	+	–

Table 1 continued

Taxon	Voucher	ndhA		ndhB		ndhC	ndhD		ndhE	ndhF	ndhG	ndhH		ndhI	ndhJ	ndhK
		exon 1	exon 2	exon 1	exon 2		5' end	3' end				5' end	3' end			
Welwitschia mirabilis Hook. f.	SS-05-20 (TRTE)	+	+	−	−	+	+	−	+	−	−	+	−	+	−	−
Pinaceae [10/11; ~91%]																
Abies balsamea (L.) Mill.	SS-05-208 (TRTE)	+	+	−	−	+	−	−	−	−	+	+	−	−	−	+
Abies concolor (Gordon & Glend.) Hildebr.	SS-05-157 (TRTE)	+	+	+	−	+	−	−	+	−	+	+	−	−	−	+
Abies firma Siebold & Zucc.	QIU-96269 (IND)	+	+	+	−	+	−	−	+	−	+	+	−	−	−	−
Abies homolepis Siebold & Zucc.	QIU-96224 (IND)	+	+	+	−	?	−	−	−	−	−	+	−	−	−	+
Abies koreana E.H. Wilson	SS-05-39 (TRTE)	+	+	+	−	+	−	−	+	−	+	+	−	−	−	+
Abies lasiocarpa (Hook.) Nutt.	SS-05-158 (TRTE)	+	+	+	−	+	−	−	−	−	+	+	−	−	−	+
Keteleeria davidiana Beissn.	QIU-96248 (IND)	+	+	+	−	+	−	−	−	−	−	+	−	−	−	−
Keteleeria sp.	SS-06-224 (TRTE)	+	+	−	−	+	−	−	+	−	−	+	−	−	−	−
Tsuga canadensis Carrière	QIU-94077 (IND)	+	+	+	−	+	−	−	+	−	+	+	−	−	−	‡
Pseudolarix amabilis Rehder	SS-06-206 (TRTE)	+	+	+	−	+	−	−	+	−	−	+	−	−	−	+
Pseudotsuga menziesii (Mirb.) Franco	SS-03-108 (TRTE)	+	+	−	−	+	−	−	−	−	−	+	−	−	−	+
Larix decidua Mull.	QIU-96242 (IND)	+	+	+	−	+	−	−	+	−	−	+	−	−	−	‡
Larix larcina (Du Roi) K. Koch	SS-05-37 (TRTE)	+	+	+	−	+	+	−	+	−	+	+	−	−	−	‡
Picea abies (L.) H. Karst.	SS-03-99 (TRTE)	+	+	−	−	+	−	−	+	−	+	+	−	−	−	‡
Picea excelsa Link.	SS-03-105 (TRTE)	+	+	+	−	+	−	−	+	−	+	+	−	−	−	‡
Picea glehnii Mast.	SS-05-155 (TRTE)	+	+	−	−	+	−	−	+	−	+	+	−	−	−	‡
Picea engelmannii Parry ex Engelm.	SS-05-210 (TRTE)	+	+	+	−	+	+	−	+	−	+	+	−	−	−	‡
Picea glauca (Moench) Voss	SS-05-209 (TRTE)	+	+	+	+	+	+	−	+	−	+	+	−	−	−	‡
Picea pungens Engelm.	SS-03-106 (TRTE)	+	+	+	−	+	+	−	+	−	+	+	−	−	−	‡
Picea mariana Britton Sterns & Poggenb.	SS-05-156 (TRTE)	+	+	+	−	+	−	−	+	−	+	+	−	−	−	‡
Picea omorika (Pančić) Purk.	SS-03-96 (TRTE)	+	+	+	−	+	−	−	+	−	+	+	−	−	−	‡
Picea orientalis (L.) Link	SS-03-100 (TRTE)	+	+	−	−	+	+	−	+	−	+	+	−	−	−	‡
Cathaya argyrophylla Chun & Kuang	SS-06-233 (TRTE)	+	+	−	−	+	−	−	+	−	+	+	−	−	−	−
Pinus aristata Engelm.	SS-05-40 (TRTE)	+	+	−	−	+	−	−	−	−	+	+	−	−	−	+
Pinus quadrifolia Parry ex Parl.	SS-06-222 (TRTE)	+	+	−	−	+	−	−	+	−	−	+	−	−	−	‡
Pinus cembra L.	SS-05-104 (TRTE)	+	+	−	−	+	−	−	+	−	+	+	−	−	−	‡
Pinus parviflora Sieb. & Zucc.	SS-05-42 (TRTE)	+	+	−	−	+	−	−	+	−	+	+	−	−	−	‡
Pinus strobus L.	SS-05-36 (TRTE)	+	+	−	−	+	−	−	−	−	+	+	−	−	−	‡
Pinus wallichiana A.B. Jacks.	SS-03-97 (TRTE)	+	+	−	−	+	−	−	−	−	+	+	−	−	−	‡
Pinus canariensis C. Sm.	SS-05-64 (TRTE)	+	+	−	−	+	−	−	+	−	+	+	−	−	−	‡
Pinus heldreichii Christ	SS-03-98 (TRTE)	+	+	−	−	+	−	−	+	−	+	+	−	−	−	+
Pinus kesiya Royle ex Gordon	SS-06-223 (TRTE)	+	+	−	−	+	−	−	−	−	+	+	−	−	−	−

Table 1 continued

Taxon	Voucher	ndhA		ndhB		ndhC	ndhD		ndhE	ndhF	ndhG	ndhH		ndhI	ndhJ	ndhK
		exon 1	exon 2	exon 1	exon 2		5' end	3' end				5' end	3' end			
Pinus nigra J.F. Arnold	*SS-03-107* (TRTE)	+	+	–	–	+	–	–	+	–	+	+	–	–	–	++
Pinus pinaster Aiton	*SS-05-65* (TRTE)	+	+	–	–	+	–	–	+	+	+	+	–	–	–	++
Pinus pinea L.	*SS-05-63* (TRTE)	+	+	–	–	+	–	–	+	+	+	+	–	–	–	++
Pinus resinosa Aiton	*SS-05-164* (TRTE)	+	+	–	–	+	–	–	+	–	+	+	–	–	–	++
Pinus contorta Douglas ex Loudon	*SS-05-205* (TRTE)	+	+	–	–	+	–	–	+	+	+	+	–	–	–	++
Pinus coulteri D. Don	*SS-05-28* (TRTE)	+	+	–	–	+	–	–	+	+	+	+	–	–	–	++
Pinus radiata D. Don	*SS-05-99* (TRTE)	+	+	–	–	+	–	–	+	+	+	+	–	–	–	+
Cedrus atlantica Manetti	*SS-03-101* (TRTE)	–	+	+	–	+	–	–	–	–	+	+	–	–	–	–
Cedrus deodara (Roxb. ex Lambert) G. Don	*QIU-96265* (IND)	–	+	+	–	+	–	–	+	–	+	+	–	–	–	–
Araucariaceae [3/3; 100%]																	
Agathis sp.	*SS-05-66* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++
Agathis robusta (F. Muell.) F.M. Bailey	*SS-06-180* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++
Araucaria araucana K. Koch	*SS-05-117* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++
Araucaria bidwillii Hook.	*SS-05-67* (TRTE)	++	++	–	+	++	++	++	+	++	++	++	++	+	+	++
Araucaria cunninghamii Aiton ex D. Don	*SS-05-106* (TRTE)	++	++	–	+	++	++	++	+	++	++	++	+	?	+	++
Araucaria heterophylla (Salisb.) Franco	*QIU-96220* (IND)	++	++	+	+	++	++	++	+	++	++	++	++	+	++	++
Araucaria rulei F. Muell.	*SS-05-101* (TRTE)	++	++	+	+	++	++	++	+	++	++	++	++	?	++	++
Araucaria sp.	*SS-05-21* (TRTE)	++	++	–	+	++	++	++	+	++	++	++	++	?	++	++
Wollemia nobilis W.G. Jones K.D. Hill & J.M.Allen	*SS-05-176* (TRTE)	++	++	–	+	++	++	++	+	++	++	++	++	+	++	++
Podocarpaceae [14/19; ~ 74%]																	
Phyllocladus alpinus Hook. f.	*SS-06-231* (TRTE)	++	++	?	+	++	++	++	++	++	++	++	++	+	++	++
Phyllocladus asplenifolius (Labill.) Hook.f.	*SS-06-199* (TRTE)	++	++	?	?	++	++	++	++	++	++	++	++	+	++	++
Halocarpus bidwillii (Hook. f. ex T. Kirk) Quinn	*SS-06-201* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	+	++	++
Lagarostrobos colensoi (Hook.) Quinn	*SS-06-187* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	+	++	++
Prumnopitys andina (Poepp. ex Endl.) de Laub.	*SS-06-188* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	+	++	++
Sundacarpus amara (Blume) C-N. Page	*SS-06-177* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	+	++	++
Saxegothea conspicua Lindl.	*QIU-96243* (IND)	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++
Microstrobos fitzgeraldii (F.Muell.) J.Garden & L.A.S. Johnson	*SS-06-232* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	+	++	++
Acmopyle pancheri Pilg.	*SS-06-182* (TRTE)	++	++	?	+	++	++	++	++	++	++	++	?	?	?	++
Dacrycarpus dacrydioides (A.Rich.) de Laub.	*SS-06-186* (TRTE)	++	++	–	+	++	++	++	++	?	?	++	++	+	++	++
Dacrydium cupressinum Sol. ex Lamb.	*SS-05-120* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	+	++	++
Falcatifolium sp.	*SS-06-179* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	+	++	++
Podocarpus coriaceus Rich. & A. Rich.	*SS-05-111* (TRTE)	++	+	+	+	++	++	++	++	++	++	++	++	+	++	++

Table 1 continued

Taxon	Voucher	ndhA exon 1	ndhA exon 2	ndhB exon 1	ndhB exon 2	ndhC	ndhD 5′ end	ndhD 3′ end	ndhE	ndhF	ndhG	ndhH 5′ end	ndhH 3′ end	ndhI	ndhJ	ndhK
Podocarpus costalis C. Presl	*QIU-96149* (IND)	++	++	+	+	++	++	++	++	++	++	++	++	+	++	++
Podocarpus elatus Endl.	*SS-05-107* (TRTE)	++	++	–	+	++	++	++	++	++	++	++	++	+	+	++
Podocarpus falcatus (Thunb.) Mirb.	*SS-05-114* (TRTE)	++	++	–	+	++	++	++	++	++	++	++	++	+	+	++
Podocarpus fasciculus de Laub.	*QIU-96148* (IND)	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++
Podocarpus grayae de Laub.	*SS-06-204* (TRTE)	++	++	–	+	++	++	++	++	++	++	++	++	+	+	++
Podocarpus latifolius R. Br.	*SS-05-179* (TRTE)	++	++	–	+	++	++	++	++	++	++	++	++	+	+	++
Podocarpus macrophyllus (Thunb.) D. Don	*QIU-95006* (IND)	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++
Podocarpus nivalis Hook.	*SS-05-103* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	+	++	++
Podocarpus sp.	*SS-03-39* (TRTE)	++	++	–	–	++	++	++	++	++	++	++	++	–	+	++
Podocarpus totara G. Benn. ex D. Don	*SS-05-178* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++
Nageia nagi Britton & P. Wilson	*QIU-96146* (IND)	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++
Afrocarpus gracilior (Pilg.) C.N. Page	*SS-03-55* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++
Retrophyllum comptonii (Buchh.) C.N. Page	*SS-06-181* (TRTE)	++	++	?	?	++	++	++	?	?	?	?	?	++	?	++
Retrophyllum minor (Carrière) C.N. Page	*SS-06-178* (TRTE)	++	++	?	?	++	++	++	?	?	?	?	?	+	?	++
Taxaceae [3/5; 60%]																
Amentotaxus formosana H.L. Li	*SS-06-211* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	+	++	++
Taxus baccata L.	*SS-05-109* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	+	++	++
Taxus canadensis Marshall	*SS-05-41* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	+	++	++
Taxus cuspidata Sieb. & Zucc.	*SS-05-38* (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++
Taxus × media Rahder	*QIU-96218* (IND)	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++
Torreya californica Torr.	*SS-05-121* (TRTE)	++	++	+	+	++	++	++	+	++	++	++	++	+	++	++
Torreya nucifera Sieb. & Zucc.	*SS-05-180* (TRTE)	++	++	–	–	++	++	++	+	++	++	++	++	+	–	++
Torreya taxifolia Arn.	*QIU-96251* (IND)	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++
Cephalotaxaceae [1/1; 100%]																
Cephalotaxus harringtonia K. Koch	*SS-03-102* (TRTE)	++	++	+	+	++	++	++	++	++	+	++	++	+	++	++
Cephalotaxus fortunei Hook.	*SS-06-221* (TRTE)	++	++	+	+	++	++	++	++	++	+	++	++	+	++	++
Sciadopityaceae [1/1; 100%]																
Sciadopitys verticillata Sieb. & Zucc.	*QIU-96087* (IND)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++
Cupressaceae s.l. [27/30; 90%]																
Cunninghamia lanceolata Hook.	*QIU-96254* (IND)	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++
Taiwania flousiana Gaussen	*QIU-96266* (IND)	++	++	+	++	++	++	++	++	++	+	++	++	++	++	++
Athrotaxis sp.	*SS-06-230* (TRTE)	++	++	?	+	++	++	++	?	++	++	++	++	++	++	++
Sequoia sempervirens Endl.	*QIU-96091* (IND)	++	++	+	++	++	++	++	++	++	+	++	++	++	++	++
Sequoiadendron giganteum (Lindl.) J. Buchholz	*QIU-96089* (IND)	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++

Table 1 continued

Taxon	Voucher	ndhA		ndhB		ndhC	ndhD		ndhE	ndhF	ndhG	ndhH		ndhI	ndhJ	ndhK
		exon 1	exon 2	exon 1	exon 2		5′ end	3′ end				5′ end	3′ end			
Cryptomeria japonica D. Don	QIU-96267 (IND)	++	++	+	++	++	++	++	++	++	+	++	++	++	++	++
Glyptostrobus pensilis (*lineatus*) K. Koch	QIU-96264 (IND)	++	++	+	++	?	++	++	++	++	++	++	++	?	++	++
Taxodium distichum H-B. & K.	SS-05-113 (TRTE)	++	++	+	++	++	++	++	++	++	+	++	++	++	++	++
Austrocedrus chilensis (D.Don) Florin & Boutelje	SS-06-190 (TRTE)	++	++	+	+	++	++	++	++	++	+	++	++	++	++	++
Libocedrus plumosa Druce	SS-06-196 (TRTE)	++	++	+	+	++	++	++	++	++	+	++	++	++	++	++
Pilgerodendron uviferum Florin	SS-06-202 (TRTE)	++	++	+	+	?	?	?	?	?	++	++	++	?	++	++
Actinostrobus arenarius C.A. Gardner	SS-06-226 (TRTE)	++	++	+	+	++	++	++	++	?	?	++	++	+	++	++
Callitris drummondii (Parl.) F. Muell.	QIU-96250 (IND)	++	++	+	+	++	++	++	++	++	+	++	++	++	++	++
Diselma archeri Hook. f.	SS-06-203 (TRTE)	++	++	−	+	++	++	++	++	++	++	++	++	++	++	++
Fitzroya cupressoides I.M. Johnst.	SS-06-200 (TRTE)	++	++	+	+	++	++	++	++	?	?	++	++	++	++	++
Thuja occidentalis L.	QIU-96223 (IND)	++	++	+	+	++	++	++	++	++	+	++	++	++	++	++
Thuja orientalis L.	SS-05-69 (TRTE)	++	++	+	+	++	++	++	++	++	+	++	++	++	++	++
Thuja plicata Donn ex D. Don	SS-05-211 (TRTE)	++	++	+	+	++	++	++	++	++	+	++	++	++	++	++
Thujopsis dolabrata Siebold & Zucc.	QIU-96246 (IND)	++	++	−	+	++	++	++	++	++	+	++	++	++	++	++
Fokienia hodginsii A. Henry & H·H. Thomas	SS-06-198 (TRTE)	++	++	−	−	++	++	++	++	+	+	+	++	+	+	++
Chamaecyparis obtusa Siebold & Zucc.	SS-05-187 (TRTE)	++	++	+	+	++	++	++	++	++	?	++	++	++	++	++
Chamaecyparis pisifera (Siebold & Zucc.) Endl.	SS-05-161 (TRTE)	++	++	?	+	?	++	++	++	++	?	++	++	++	++	++
Chamaecyparis thyoides (L.) Britton Sterns & Poggenb.	SS-05-74 (TRTE)	++	++	+	+	++	++	++	++	+	+	++	++	++	++	++
Tetraclinis articulata Mast.	SS-05-72 (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	++	++	++
Calocedrus decurrens (Torr.) Florin	QIU-96245 (IND)	++	++	+	+	++	++	++	++	++	+	++	++	++	++	++
Microbiota decussata Kom.	QIU-96252 (IND)	++	++	−	++	++	++	++	++	++	+	++	++	++	++	++
Platycladus orientalis (L.) Franco	SS-06-192 (TRTE)	++	++	+	+	++	++	++	++	++	+	++	++	++	++	++
Cupressus sempervirens L.	SS-05-122 (TRTE)	++	++	+	+	++	++	++	++	++	+	++	++	++	++	++
Cupressus cashmeriana Royle ex Carriere	SS-06-01 (TRTE)	++	++	+	+	++	++	++	++	?	?	++	++	?	++	++
Cupressus dupreziana A. Camus	SS-05-70 (TRTE)	++	++	+	+	++	++	++	++	?	?	++	++	?	++	++
Juniperus chinensis L.	SS-05-183 (TRTE)	++	++	+	+	++	++	++	++	++	++	++	++	?	++	++
Juniperus communis L.	SS-05-83 (TRTE)	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++
Juniperus horizontalis Moench	SS-05-160 (TRTE)	++	++	+	++	++	++	++	++	++	+	++	++	++	++	++
Juniperus indica Bertol.	QIU-96270 (IND)	++	++	++	++	++	++	++	++	++	+	++	++	++	++	++
Juniperus procumbens Siebold	SS-05-159 (TRTE)	++	++	+	++	++	++	++	++	++	+	++	++	++	++	++
Juniperus sabina L.	SS-05-163 (TRTE)	++	++	+	+	++	++	++	++	++	+	++	++	++	++	++
Juniperus scopulorum Sarg.	SS-05-162 (TRTE)	++	++	+	++	++	++	++	++	++	++	++	++	++	++	++
Juniperus virginiana L.	SS-05-84 (TRTE)	++	++	+	++	++	++	++	++	++	+	++	++	++	++	++

Table 1 continued

Taxon	Voucher	ndhA		ndhB		ndhC	ndhD		ndhE	ndhF	ndhG	ndhH		ndhI	ndhJ	ndhK
		exon 1	exon 2	exon 1	exon 2		5′ end	3′ end				5′ end	3′ end			
Xanthocyparis nootkatensis (D. Don) Farjon & D.K. Harder	*QIU-94088* (IND)	++	++	+	++	++	++	++	++	++	+	++	++	++	++	++
Callitropsis arizonica (glabra) (Greene) D.P. Little	*QIU-96254* (IND)	++	++	+	++	++	++	++	++	++	+	++	++	++	++	++
Callitropsis bakeri (Jeps.) D.P. Little	SS-05-73 (TRTE)	++	++	+	+	++	++	++	++	++	+	++	++	+	++	++
Callitropsis sargentii (Jeps.) D.P. Little	SS-05-190 (TRTE)	++	++	−	+	++	++	++	++	++	+	++	++	++	++	++
Angiosperms																
Amborella trichopoda Baill.	No voucher	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++
Victoria sp.	No voucher	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++
Acorus gramineus Soland.	SS-03-70 (TRTE)	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++
Glycine max (L.) Merr.	No voucher	++	++	++	++	++	++	++	++	++	++	++	++	++	++	++

Numbers in square parentheses following taxa above generic rank indicate number of genera in that taxon and the approximate percentage of coverage. Within major groups of seed plants, species are arranged according to their presumptive phylogenetic relationships (see text for details). Based on slot-blot results, a "++" symbol indicates presence of full hybridization, "+" diminished hybridization, "−" absence of hybridization signal in comparison to positive controls, and "?" unable to score. Selected slot-blot hybridization results for representatives from all major groups (species indicated in boldface) are shown in Fig. 1. Abbreviations of herbaria follow Index Herbariorum

herbarium specimens. Quality of DNA varied significantly, from high molecular weight to considerably degraded (from some herbarium material). Because of the poor quality of a number of samples, frequently accompanied also by limited quantity, the standard restriction endonuclease digestion followed by electrophoretic separation approach to Southern hybridization (Sambrook et al. 1989) could not be used. Instead, the slot-blot hybridization method was used, as described in detail by Doyle et al. (1995).

In brief, a slot-blot apparatus (Bio-Rad) was used to make seven sets of pseudoreplicate filter-blots, following the manufacturer's protocol. Approximately 500–800 ng of total DNA (per sample and per set) was bound to Immobilon-Ny + nylon membrane (Millipore). DNAs from several species with sequenced pt genomes (e.g., *Nicotiana*, *Acorus*, *Amborella*, *Ginkgo,* or *Welwitschia*) were included on each membrane as positive or negative controls. Membranes were prehybridized and hybridized at 60°C–62°C in 5× standard saline citrate (SSC), 0.1% SDS, 50 mM Tris (pH 8.0), 10 mM EDTA, 2× Denhardt's solution, and 5% dextran sulfate. After hybridization, filters were washed twice for 30–45 min in 0.5% SDS and 2× SSC at the hybridization temperature. Probes were labeled with ^{32}P using random oligonucleotide primers (Invitrogen). Autoradiography was carried out using intensifying screens at −80°C for 18–48 h. Filters were stripped of probe between hybridizations by boiling twice for 5–10 min in 0.1% SSC. The absence of carryover signal from previous hybridizations was assured by an overexposure (3–5 days) prior to new rounds of hybridization.

Hybridization probes for all 11 plastid-encoded *ndh* genes and small plastid ribosomal subunit (16S rDNA; used as a control probe) were derived from tobacco via polymerase chain reaction (PCR). Primer names and sequences are provided in Supplementary Table 1. A total of 17 probes were constructed and their relative positions are indicated in Supplementary Fig. 1. For the two *ndh* genes, usually interrupted by introns (*ndhA* and *ndhB*), two probes were used, each covering one exon. Two additional longer *ndh* genes (*ndhD* and *ndhH*) were surveyed with two probes situated at the 5′- and 3′-ends, respectively. In addition, to estimate the unspecific background hybridization levels, an initial negative hybridization control was performed under the same stringency conditions (see above) and the same amount of ^{32}P, but without probe added.

Results

Interpretation of slot-blots

The slot-blot data ranged from no diminution to complete absence of signal and were for the most part readily

interpretable. The presence or absence of *ndh* genes was determined by eye, by comparison of hybridization signal to the corresponding 16S probe, which is used as a control to establish the presence of significant amounts of ptDNA on the blots. For each blot set and probe combination, the strength of signal was estimated by comparison with a number of positive and negative controls; namely, the species known to contain functional *ndh* genes (e.g., *Ginkgo, Amborella, Acorus*, etc.) or to lack these genes (e.g., *Welwitschia, Pinus* spp., etc.), based on previously available entire ptDNA sequence data.

Representative hybridization results, arranged phylogenetically, are shown in Fig. 2, and the scores for all surveyed species and probes are summarized in Table 1. For every probe, the relative presence or absence of signal was scored for each taxon as showing either full (++), diminished (+), or absent (−) hybridization in comparison to their respective 16S positive controls (Table 1). Full hybridization strength is assumed to indicate that the gene is present and functional. For genes that have two probes (i.e., two exons or 5′- and 3′-end) full hybridization to both probes is required to indicate that the gene is functional. Diminished signals, where hybridization is weaker than the control but there is definite signal presence, can be interpreted in two different ways. It can indicate that the gene is divergent with respect to tobacco but still present and functional or that the gene is present but pseudogenized (i.e., rendered nonfunctional). Absence was scored if no detectable hybridization to the probe was observed. Under our experimental conditions, plants in which a gene has been transferred to the nucleus would fail to produce a detectable hybridization signal when compared to a plant that retains the gene in its plastid genome, due to significant reduction in copy number and an increase in nucleotide substitution rates (Wolfe et al. 1987). Hence, the absence of signal implies either outright loss of genes or their functional transfer to the nucleus. Given the generally conservative substitution rates of ptDNA, it is less likely that the absence of signal represents a highly divergent yet functional gene. Lastly, in certain cases we were unable to determine the presence or absence of signal and we scored these taxa as unknown ("?"; see Table 1). These ambiguities are due to insufficient amounts or poor quality of ptDNA for a given pseudoreplicate.

Altogether, these assumptions on the presence or absence of genes can lead to potential underestimates or overestimates of gene losses. For example, signals that appear present could potentially represent relatively recent pseudogenized genes while significantly diminished signals might be due to divergent but functional genes. Nevertheless, despite these caveats, a hybridization approach remains a cost effective and efficient method for surveying numerous and diverse samples (Doyle et al. 1995; Adams et al. 2002).

Distribution of *ndh* gene losses

As expected, the full hybridization signal was observed for all taxa used as positive controls. Based on results from whole pt genome sequencing, the entire *ndh* gene suite is known to be present in autotrophic angiosperms, indicating that its presence is the shared ancestral condition for this group of plants (Jansen et al. 2007). The presence of all 11 *ndh* genes is confirmed here across the representatives chosen to span the basal nodes of flowering plant diversity (Fig. 2; Table 1). Also, for *Ginkgo*, the hybridization to all probes derived from the *ndh* genes was similar in strength to the control DNA (Fig. 2; Table 1). This was expected as well, given the known presence of these genes based on extensive sequencing of its pt genome (Leebens-Mack et al. 2005; Jansen et al. 2007). In addition, the relative strength of the hybridization of tobacco-derived probes to *Ginkgo* illustrates the conserved nature of the *ndh* genes across large phylogenetic distances, including the angiosperm-gymnosperm divergence (>325 Mya; Judd et al. 2002; Palmer et al. 2004).

As a group, cycads strongly hybridized to seven of 11 *ndh* genes. For a few taxa (Table 1), the signal was diminished with probes for *ndhB* (both exons), *ndhH* 3, *ndhI,* and *ndhJ*. Although some of the hybridizations are weaker than the positive controls, they do not necessarily indicate loss of function of these *ndh* genes but rather that these genes are divergent to some degree in these taxa. The presence of functional *ndh* genes in cycads is expected given that they are found as open reading frames (ORFs) in the sequenced ptDNA of *Cycas taitungensis* (Wu et al. 2007).

In contrast to the previous lineages, the evidence for loss of *ndh* genes from pt genomes is widespread in Gnetales and Pinaceae. In Gnetales, there was no significant hybridization signal for most of the probes (Fig. 2; Table 1). For some taxa weak signal was present for *ndhA* (both exons), *ndhC,* and *ndhH* 5′-end. This pattern indicates that the loss of the *ndh* genes is common to all Gnetales (Fig. 2), in accordance with the results of entire ptDNA sequences from a small number of select representatives of this group (McCoy et al. 2008; Wu et al. 2009). Similarly, there was a generally weak to absent hybridization signal observed for most probes across Pinaceae. In particular, *ndhI* and *ndhJ* are absent from all surveyed taxa. Substantially diminished signal was common for probes derived from *ndhA* (both exons), *ndhD, ndhE, ndhH,* and *ndhI* genes. However, *ndhB* (most notably, the 5′ end)*, ndhC,* and *ndhK* appeared present in many (but not all) taxa. Given what is known from the entire plastid genome sequences of several *Pinus spp.* (Wakasugi et al. 1994; Cronn et al. 2008), *Picea sitchensis* (Cronn et al. 2008), and *Keteleeria davidiana* (Wu et al. 2009) species, it can be deduced that the presence of weak to moderate hybridization signal observed for these
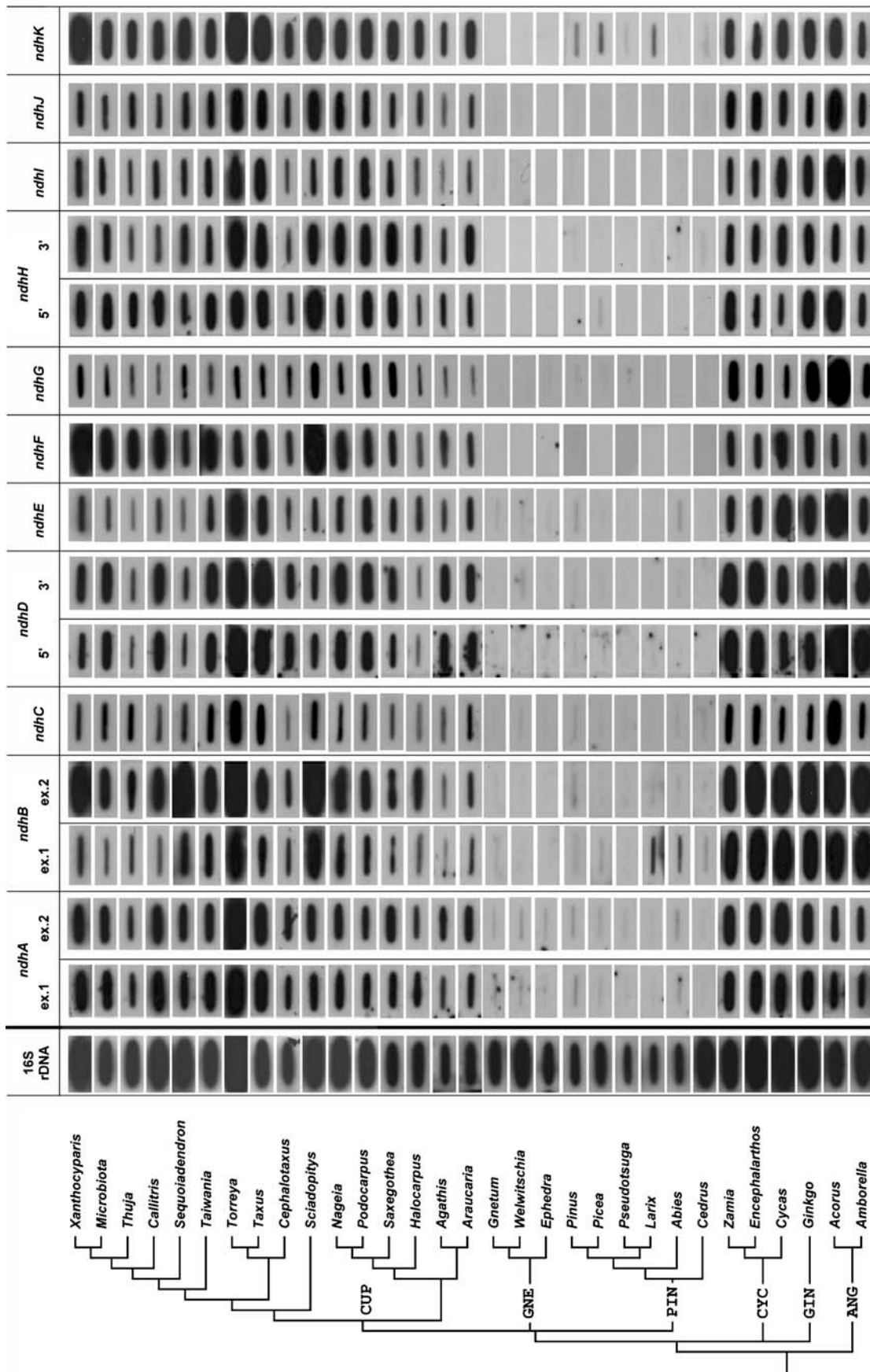
Fig. 2 Autoradiographs showing slot-blot hybridization results of probes derived from *ndh*(*A-K*) genes for 31 selected species representing seed plant (out of 166 surveyed; compare with Table 1) presented in a phylogenetic context. Small plastid ribosomal subunit (16S rDNA) was used as positive control (shown here is one representative out of seven filter-sets). The topology shown is a composite tree depicting current understanding of relationships derived from several published phylogenetic analyses (see text for references). Note that the absence or near absence of hybridization for the *ndh* probes is restricted only to the surveyed members of Gnetales and Pinaceae. Taxon abbreviations are the same as in Fig. 1

ndh genes (Fig. 2; Table 1) corresponds to pseudogenes. Compared to Pinaceae, Gnetales appear to have fewer remnants of *ndh* genes (Fig. 2; Table 1), which is expected, given the highly elevated rates of molecular evolution observed in Gnetales plastids generally (Rydin et al. 2002; Burleigh and Mathews 2007a; McCoy et al. 2008).

Unlike Gnetales and Pinaceae, the general trend across cupressophytes was a strong hybridization to almost all *ndh* probes, indicating that the entire suite of *ndh* genes is present and conserved within this group (Fig. 2; Table 1). This is fully in agreement with the only published entirely sequenced ptDNA from cupressophytes, *Cryptomeria japonica* (Cupressaceae s. lat.; Hirao et al. 2008). However, some members of Araucariaceae, Podocarpaceae, Taxaceae, and Cephalotaxaceae hybridized weakly to *ndhB* (one or both exons) and *ndhI*, while Cupressaceae s. lat. exhibit diminished hybridization signal to *ndhG* but not *ndhI* (see Table 1). In these cases, diminished signal is most likely due to the elevated sequence divergence of *ndh* genes, as evidenced by the presence of *ndhB* and *ndhF* ORFs in few representative species from these families that are currently sequenced (Rai et al. 2008).

Discussion

Implications of *ndh* losses for spermatophyte relationships

The *ndh* genes comprise about one-tenth of the ~120 genes retained in plastids of most photosynthetic seed plants. Based on numerous entirely sequenced pt genomes of angiosperms (see Jansen et al. 2007 for the most recent summary) as well as a limited number of gymnosperms (Wakasugi et al. 1994; Wu et al. 2007, 2009; Cronn et al. 2008; McCoy et al. 2008; Hirao et al. 2008), it appeared that the loss of the *ndh* genes was restricted to Gnetales (McCoy et al. 2008; Wu et al. 2009) and Pinaceae (Wakasugi et al. 1994; Cronn et al. 2008; Wu et al. 2009). The results of our survey extend the previous inferences of the *ndh* gene absence to be common to all of Gnetales and Pinaceae, but not to other gymnosperms (nor to autotrophic angiosperms). While the losses of *ndh* genes from the plastids is rare in autotrophic plants, their absence has been observed repeatedly in heterotrophic angiosperms (Olmstead et al. 2001; Stefanović and Olmstead 2005). Extrapolating from those cases, it seems that the *ndh* genes are generally lost as a suite (Bungard 2004; Krause 2008). Hence, from a phylogenetic point of view, the entire suite should be considered as a single loss (Stefanović and Olmstead 2005) and not as 11 independent losses.

Given the extent and distribution of presence and absence of the *ndh* genes among spermatophytes, the most parsimonious solution suggests that the loss of these genes

is a synapomorphy for Gnetales and Pinaceae, a shared derived character inherited from their common ancestor (Fig. 1) supporting the gnepine hypothesis. Each of the five alternatives for the Gnetales relationships with the other seed plants (i.e., anthophyte, gnetales-sister, gnetifer, and gnecup hypthotheses) would require a minimum of two independent losses of *ndh* genes, one in Gnetales and one in Pinaceae (Fig. 1). While the possibility that *ndh* genes were lost more than once in gymnosperms cannot be positively excluded, the low frequency of loss of these genes, and in particular the near-absence of loss among autotrophic seed plants, argues against such a scenario.

Additional lines of evidence supporting the gnepine hypothesis are provided by a couple of other plastid structural characters, each with putatively the same phylogenetic distribution among seed plants as that observed for the loss of *ndh* genes. First, an expansion of the inverted repeat (IR) that includes the duplication of *trnI*-CAU and a partial duplication of the *psbA* gene region situated at the end of the large single copy (LSC) is found in several sequenced members of Gnetales and Pinaceae but is not known from any other land plant (Wu et al. 2007, 2009; McCoy et al. 2008; Hirao et al. 2008). Second, both Gnetales and Pinaceae appear to share a common loss of *rps16*, to the exclusion of other gymnosperms and basal angiosperms lineages from which the presence of this gene is ascertained (Wakasugi et al. 1994; Wu et al. 2007, 2009; Jansen et al. 2007; McCoy et al. 2008). However, contrary to the above examples, one particular structural genomic marker does not support the gnepine phylogeny. The loss (or a significant reduction) of the IR is reported from both cupressophytes and Pinaceae, but not from Gnetales (Raubeson and Jansen 1992; Wu et al. 2007, 2009; McCoy et al. 2008; Hirao et al. 2008). This ptDNA feature favors the phylogenetic interpretation according to which the loss of the *ndh* genes occurred independently in Gnetales and Pinaceae.

Future directions

Although the preponderance of genomic structural changes currently supports the gnepine hypothesis, caution is still warranted because most of these features have only been observed in a subset of exemplar taxa. Hence, their full distribution across seed plants and their evolutionary significance remains poorly understood. To determine their relative importance, all of these underexamined markers require further in-depth surveys across a broader taxon sample, as was done here for the *ndh* genes. In particular, additional ptDNA sequences are needed from the representatives of Araucariaceae and Podocarpaceae, to help triangulate the ancestral conditions for many of these potentially phylogenetically important characters in cupressophytes. It would also be valuable to survey for *ndh* genes in *Parasit-*

axus usta (Podocarpaceae), the only known mycoheterotrophic conifer (Feild and Brodribb 2005). We predict the functional absence of all *ndh* genes in this highly derived podocarp species (Sinclair et al. 2002), which would represent an independent loss of the Ndh complex in conifers, related to its shift to a fully heterotrophic nutritional mode.

References

Adams KL, Qiu YL, Stoutemyer M, Palmer JD (2002) Punctuated evolution of mitochondrial gene content: high and variable rates of mitochondrial gene loss and transfer to the nucleus during angiosperm evolution. Proc Natl Acad Sci 99:9905–9912

Albert VA, Backlund A, Bremer K, Chase MW, Manhardt JR, Mishler BD, Nixon KC (1994) Functional constraints and *rbcL* evidence for land plant phylogeny. Ann Mo Bot Gard 81:534–567

Bailey CD, Doyle JJ, Kajita T, Nemoto T, Ohashi H (1997) The chloroplast *rpl2* intron and ORF184 as phylogenetic markers in the legume tribe Desmodieae. Syst Bot 22:133–138

Bennoun P (2002) The present model for cholororespiration. Photosynthesis Res 73:273–277

Bowe LM, Coat G, dePamphilis CW (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 97:4092–4097

Bukhov N, Carpentier R (2004) Alternative photosystem I-driver electron transport routes: mechanisms and functions. Photosynthesis Res 82:17–33

Bungard RA (2004) Photosynthetic evolution in parasitic plants: insight from the cholorplast genome. Bioassays 26:235–247

Burleigh JG, Mathews S (2004) Phylogenetic signal in nucleotide data from seed plants: implications for resolving the seed plant tree of life. Am J Bot 91:235–237

Burleigh JG, Mathews S (2007a) Assessing among-locus variation in the inference of seed plant phylogeny. Int J Plant Sci 168(2):111–124

Burleigh JG, Mathews S (2007b) Assessing systematic error in the inference of seed plant phylogeny. Int J Plant Sci 168(2):125–135

Casano LM, Martín M, Sabater B (2001) Hydrogen peroxide mediates the induction of chloroplastic Ndh complex under photooxidative stress in barley. Plant Physiol 125:1450–1458

Chang C-C, Lin H-C, Lin I-P, T-Y Chow, Chen H-H, Chen W-H, Cheng C-H, Lin C-Y, Liu S-M, Chang C-C, Chaw S-M (2006) The chloroplast genome of *Phalaenopsis aphrodite* (Orchidaceae):

comparative analysis of evolutionary rate with that of grasses and its phylogenetic implications. Mol Biol Evol 23:279–291

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

Chaw S-M, Parkinson CL, Cheng Y, Vincent TM, Palmer JD (2000) Seed plant phylogeny inferred from all three plant genomes: monophyly of extant gymnosperms and origin of Gnetales from conifers. Proc Natl Acad Sci 97:4086–4091

Chumley TW, McCoy SKR, Raubeson LA (2008) Gne-deep: exploring Gnetalean affinities in seed plant phylogeny with 83 plastid genes. Botany 2008: Joint Annual Meeting of Canadian Botanical Association, American Fern Society, American Society of Plant Taxonomists, and the Botanical Society of America, Vancouver, BC, Canada [http://2008.botanyconference.org/engine/search/index.php?func=detail&aid=770]

Clegg MT, Gaut BS, Learn GH, Morton BR (1994) Rates and patterns of chloroplast DNA evolution. Proc Natl Acad Sci 91:6795–6801

Crane PR (1985) Phylogenetic analysis of seed plants and the origin of angiosperms. Ann Mo Bot Gard 72:716–793

Cronn R, Liston A, Parks M, Gernandt DS, Shen R, Mockler T (2008) Multiplex sequencing of plant chloroplast genomes using Solexa sequencing-by-synthesis technology. Nucleic Acids Res 36(19):e122

dePamphilis CW, Palmer JD (1990) Loss of photosynthetic and chlororespiratory genes from the plastid genome of a parasitic flowering plant. Nature 348 (22):337–339

Diaz M, de Haro V, Munoz R, Quiles MJ (2007) Chlororespiration is involved in the adaptation of *Brassica* plants to heat and high light intensity. Plant Cell Environ 30:1578–1585

Downie SR, Palmer JD (1992) Use of chloroplast DNA rearrangements in reconstructing plant phylogeny. In: Soltis PS, Soltis DE, Doyle JA (eds) Molecular systematics of plants. Chapman & Hall, New York, pp 14–35

Downie SR, Olmstead RG, Zurawski G, Soltis DE, Soltis PS, Watson JC, Palmer JD (1991) Six independent losses of the chloroplast DNA *rpl2* intron in dicotyledons: molecular and phylogenetic implications. Evolution 45:1245–1259

Doyle JA (2006) Seed ferns and the origin of angiosperms. J Torrey Bot Soc 133:169–209

Doyle JA, Donoghue MJ (1986) Seed plant phylogeny and the origin of angiosperms: an experimental cladistic approach. Bot Rev 52:321–431

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

Doyle JJ, Doyle JL, Palmer JD (1995) Multiple independent losses of two genes and one intron from legume chloroplast genomes. Mol Phylogenet Evol 5:429–438

Doyle JJ, Doyle JL, Ballenger JA, Palmer JD (1996) The distribution and phylogenetic significance of a 50-kb chloroplast DNA inversion in the flowering plant family Leguminosae. Syst Bot 20:272–294

Endo T, Ishida S, Ishikawa N, Sato F (2008) Chloroplastic NAD(P)H dehydrogenase complex and cyclic electron transport around photosystem I. Mol Cells 25:158–162

Feild TS, Brodribb TJ (2005) A unique mode of parasitism in the conifer coral tree *Parasitaxus ustus* (Podocarpaceae). Plant Cell Environ 28:1316–1325

Felsenstein J (1983) Parsimony in systematics: biological and statistical issues. Annu Rev Ecol Syst 14:313–333

Frolich MW, Parker DS (2000) The mostly male theory of flower evolutionary origins: from genes to fossils. Syst Bot 25(2):155–170

Funk HT, Berg S, Krupinska K, Maier UG, Krause K (2007) Complete DNA sequences of the plastid genomes of two parasitic flowering plant species, *Cuscuta reflexa* and *Cuscuta gronovii*. BMC Plant Biol 7:45

Goremykin V, Bobrova V, Pahnke J, Troitsky A, Antonov A, Martin W (1996) Noncoding sequences from the slowly evolving chloroplast inverted repeat in addition to *rbcL* data do not support gnetalean affinities of angiosperms. Mol Biol Evol 13:383–396

Graham SW, Iles WJD (2009) Different gymnosperm outgroups have (mostly) congruent signal regarding the root of flowering plant phylogeny. Am J Bot 96:216–227

Graham SW, Olmstead RG (2000a) Evolutionary significance of an unusual chloroplast DNA inversion found in two basal angiosperm lineages. Curr Genet 37:183–188

Graham SW, Olmstead RG (2000b) Utility of 17 chloroplast genes for inferring the phylogeny of the basal angiosperms. Am J Bot 87:1712–1730

Gugerli F, Sperisen C, Buchler U, Brunner L, Brodbeck S, Palmer JD, Qiu YL (2001) The evolutionary split of Pinaceae from other conifers: evidence from an intron loss and a multigene phylogeny. Mol Phylogen Evol 21:167–175

Hamby RK, Zimmer EA (1992) Ribosomal RNA as a phylogenetic tool. In: Soltis PE, Soltis DE, Doyle JJ (eds) Molecular systematics of plants. Chapman & Hall, London, pp 50–91

Hasebe M, Kofugi R, Ito M, Kato M, Iwatsuki K, Ueda K (1992) Phylogeny of gymnosperms inferred from *rbcL* gene sequence. J Plant Res 105:673–679

Hendy MD, Penny D (1989) A framework for the quantitative study of evolutionary trees. Syst Zool 38:296–309

Hirao T, Watanabe A, Kurita M, Kondo T (2008) Complete nucleotide sequence of the *Cryptomeria japonica* D. Don. chloroplast genome and comparative chloroplast genomics: diversified genomic structure of coniferous species. BMC Plant Biol 8:70

Jansen RK, Cai Z, Raubeson LA, Daniell H, Depamphilis CW, Leebens-Mack J, Muller KF, Guisinger-Bellian M, Haberle RC, Hansen AK, Chumley TW, Lee SB, Peery R, McNeal JR, Kuehl JV, Boore JL (2007) Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. Proc Natl Acad Sci 104:19369–19374

Judd WS, Campbell CS, Kellogg EA, Stevens PF, Donoghue MJ (2002) Plant systematics: a phylogenetic approach, 2nd edn. Sinauer Associates, Sunderland, p 576

Krause K (2008) From chloroplasts to "cryptic" plastids: evolution of plastid genomes in parasitic plants. Curr Genet 54:111–121

Leebens-Mack J, Raubeson LA, Cui LY, Kuehl JV, Fourcade MH, Chumley TW, Boore JL, Jansen RK, dePamphilis CW (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

Loconte H, Stevenson DW (1990) Cladistics of the Spermatophyta. Brittonia 42:197–211

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

Mathews S (2009) Phylogenetic relationships among seed plants: persistent questions and the limits of molecular data. Am J Bot 96(1):228–236

Mathews S, Donoghue MJ (2000) Basal angiosperm phylogeny inferred from duplicate phytochromes A and C. Int J Plant Sci 161:S41–S55

McCoy SR, Kuehl JV, Boore JL, Raubeson LA (2008) The complete plastid genome sequence of *Welwitschia mirabilis*: an unusually compact plastome with accelerated divergence rates. BMC Evol Biol 8:130

McNeal JR, Kuehl JV, Boore JL, dePamphilis CW (2007) Complete plastid genome sequences suggest strong selection for retention of photosynthetic genes in the parasitic plant genus *Cuscuta*. BMC Plant Biol 7:57

McPherson MA, Fay ME, Chase MW, Graham SW (2004) Parallel loss of a slowly evolving intron from two closely related families in asparagales. Syst Bot 29:296–307

Moreira D, Philippe H (2000) Molecular phylogeny: pitfalls and progress. Int Microbiol 3:9–16

Nickrent DL, Parkinson CL, Palmer JD, Duff RJ (2000) Multigene phylogeny of land plants with special reference to bryophytes and the earliest land plants. Mol Biol Evol 17:1885–1895

Nixon KC, Crepet WL, Stevenson D, Friis EM (1994) A reevaluation of seed plant phylogeny. Ann Mo Bot Gard 81:484–533

Olmstead RG, Sweere JA (1994) Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae. Syst Biol 43:467–481

Olmstead RG, dePamphilis CW, Wolfe AD, Young ND, Elisons WJ, Reeves PA (2001) Disintegration of the Scrophulariaceae. Am J Bot 88:348–361

Palmer JD (1991) Plastid chromosomes: structure and evolution. In: Bogarad L, Vasil IK (eds) Cell culture and somatic cell genetics of plants 7A. Academic, San Diego, pp 6–53

Palmer JD, Soltis DE, Chase MW (2004) The plant tree of life: an overview and some points of view. Am J Bot 91:1437–1445

Peltier G, Cournac L (2002) Chlororespiration. Annu Rev Plant Biol 53:523–550

Plunkett GM, Downie SR (2000) Expansion and contraction of the chloroplast inverted repeat in Apiaceae subfamily Apioideae. Syst Bot 25:648–667

Qiu Y-L, Li L, Wang B, Chen Z, Dombrovska O, Lee J, Kent L, Li R, Jobson RW, Hendry TA, Taylor DW, Testa CM, Ambros M (2007) A nonflowering land plant phylogeny inferred from nucleotide sequences of seven chloroplast, mitochondrial, and nuclear genes. Int J Plant Sci 168:691–708

Rai HS, O'Brien HE, Reeves PA, Olmstead RG, Graham SW (2003) Inference of higher-order relationships in the cycads from a large chloroplast data set. Mol Phylogen Evol 29:350–359

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

Raubeson LA, Jansen RK (1992) A rare chloroplast DNA structural mutation is shared by all conifers. Biochem Syst Ecol 20:17–24

Rokas A, Holland PWH (2000) Rare genomic changes as a tool for phylogenetics. Trends Ecol Evol 15:454–459

Romeau D, Peltier G, Cournac L (2007) Chlororespiration and cyclic electron flow around PS I during photosynthesis and plant stress response. Plant Cell Environ 104:1–1051

Rydin C, Källersjö M (2002) Taxon sampling and seed plant phylogeny. Cladistics 18:485–513

Rydin C, Källersjö M, Friis EM (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

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor

Samigullin TK, Martin WF, Troitsky AV, Antonov AS (1999) Molecular data from the chlorplast *rpoC1* gene suggest a deep and distinct dichotomy of contemporary spermatophytes into two monophyla: gymnosperms (including Gnetalaes) and angiosperms. J Mol Evol 49:310–315

Shinozaki K, Ohme M, Tanaka M, Wakasugi T, Hayashida N, Matsubayashi T, Zaita N, Chunwongse J, Obokata J, Yamaguchishinozaki K, Ohto C, Torazawa K, Meng BY, Sugita M, Deno H, Kamogashira T, Yamada K, Kusuda J, Takaiwa F, Kato A, Tohdoh N, Shimada H, Sugiura M (1986) The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J 5:2043–2049

Sinclair WT, Mill RR, Gardner MF, Woltz P, Jaffré T, Preston J, Hollingsworth ML, Ponge A, Möller M (2002) Evolutionary relationships of the New Caledonian heterotrophic conifer, *Parasitaxus usta* (Podocarpaceae), inferred from chloroplast *trnL-F* intron/spacer and nuclear rDNA ITS2 sequences. Plant Syst Evol 233:79–104

Soltis PS, Soltis DE, Savolainen V, Crane PR, Barraclough TG (2002) Rate heterogeneity among lineages of tracheophytes: integration of molecular and fossil data and evidence for molecular living fossils. Proc Natl Acad Sci 99:4430–4435

Stefanović S, Olmstead RG (2005) Down the slippery slope: plastid genome evolution in Convolvulaceae. J Mol Evol 61:292–305

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

Tallon C, Quiles MJ (2007) Acclimation to heat and high light intensity during the development of oat leaves increases the NADH DH complex and PTOX levels in chloroplasts. Plant Sci 173:438–445

Wakasugi T, Tsudzuki J, Ito S, Nakashima K, Tsudzuki T, 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 91:9794–9798

Wettstein RR (1907) Handbuch der Systematischen Botanik. Franz Deuticke, Leipzig

Wickett NJZY, Hansen SK, Roper JM, Kuehl JV, Plock SA, Wolf PG, dePamphilis CW, Boore JL, Goffinet B (2008) Functional gene losses occur with minimal size reduction in the plastid genome of the parasitic liverwort *Aneura mirabilis*. Mol Biol Evol 25(2):393–411

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci USA 84:9054–9058

Wu C-S, Wang Y-N, Liu S-M, Chaw S-M (2007) Chloroplast genome (cpDNA) of *Cycas taitungensis* and 56 cp protein-coding genes of *Gnetum parvifolium*: insights into cpDNA evolution and phylogeny of extant seed plants. Mol Biol Evol 24:1366–1379

Wu C-S, Lai Y-T, Lin C-P, Wang Y-N, Chaw S-M (2009) Evolution of reduced and compact chloroplast genomes (cpDNAs) in gnetophytes: selection towards a lower cost strategy. Mol Phylogen Evol (in press)

Zapata JM, Guéra A, Esteban-Carrasco A, Martín M, Sabater JM (2005) Chloroplasts regulate leaf senescence: delayed senescence in transgenic *ndhF*-defective tobacco. Cell Death Differ 12:1277–1284