American Journal of Botany 86(12): 1742–1753. 1999.

PHYLOGENETIC RELATIONSHIPS OF EURASIAN PINES (*PINUS*, PINACEAE) BASED ON CHLOROPLAST *RBCL*, *MATK*, *RPL20-RPS18* SPACER, AND *TRNV* INTRON SEQUENCES[1]

XIAO-RU WANG,[2] YOSHIHIKO TSUMURA,[3] HIROSHI YOSHIMARU,[3] KAZUTOSHI NAGASAKA,[3] AND ALFRED E. SZMIDT[2,4]

[2]Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, S-901 83 Umeå, Sweden; and [3]Bio-resource Technology Division, Forestry and Forest Products Research Institute, Ibaraki 305, Japan

The sequence divergence of chloroplast *rbcL, matK, trnV* intron, and *rpl20-rps18* spacer regions was analyzed among 32 *Pinus* species and representatives of six other genera in Pinaceae. The total aligned sequence length is 3570 bp. Of the four sequences examined, *matK* evolved much faster than *rbcL* in *Pinus* and in other Pinaceae genera. The two noncoding regions did not show more divergence than the two coding regions, especially within each *Pinus* subgenus. Phylogenetic analyses based on these four sequences gave consistent results and strongly supported the monophyly hypothesis for the genus *Pinus* and its two recognized subgenera. *Pinus krempfii,* the two-flat-needle pine endemic to Vietnam, was placed in subgen. *Strobus* and showed closer affinity to subsect. *Gerardianae.* The ancient character of sect. *Parrya* is further confirmed. However, monophyly of the sect. *Parrya* is not supported by our data. Among the Eurasian pines of subgen. *Pinus,* Mediterranean pines formed one clade and the Asian members of subsect. *Sylvestres* formed another. The Himalayan *P. roxburghii* showed considerable divergence from all the other hard pines from both regions. *Pinus merkusii* was distinctly separated from all the Asian members of subsect. *Sylvestres.* The implications of our results for *Pinus* classification are discussed.

Key words: *matK*; phylogeny; Pinaceae; *Pinus*; *rbcL*; *rpl20-rps18*; sequence divergence; *trnV* intron.

The genus *Pinus* is one of the most widely distributed genera of conifer trees in the Northern Hemisphere. The genus is usually divided into two subgenera *Strobus* (= *Haploxylon,* soft pines) and *Pinus* (= *Diploxylon,* hard pines), which are further divided into sections and subsections (Little and Critchfield, 1969). Classification of the genus differs among authors. In this paper, the classification scheme of Little and Critchfield (1969) is followed. Recently, research has become very active, in an attempt to achieve a better understanding of the evolution of the genus by various approaches (e.g., Strauss and Doerksen, 1990; Govindaraju, Lewis, and Cullis, 1992; Wang and Szmidt, 1993; Perez de la Rosa, Harris, and Farjon, 1995; Farjon, 1996; Krupkin, Liston, and Strauss, 1996; Wang, Szmidt, and Nguyen, 1999). The difficulties in genetic delineation are evident in the case of several species occurring in Asia and the Mediterranean part of Europe. The positions of several rare endemic species such as *P. krempfii, P. merkusii, P. heldreichii,* and *P. roxburghii,* as well as the relationships among and between Asian and Mediterranean pines are still not well settled (Schirone et al., 1991; Frankis, 1993; Krupkin, Liston, and Strauss, 1996; Liston et al., 1999). In most phylogenetic investigations, these species are seldom included. The study by Liston et al. (1999), based on nuclear ribosomal DNA internal transcribed spacer (ITS) sequences, involved a broad sampling of the *Pinus* subsections and covered a wide range of geographic regions. However, the topologies of the recovered phylogenetic trees gave weak support for many of the clades, possibly because of the rapidly evolving nature of the ITS sequence.

Mediterranean pines present an interesting group in the evolution of the genus, linking different geographic regions as well as different evolutionary lineages (Mirov, 1967; Klaus, 1989). According to Klaus (1989), Mediterranean pines represent an extremely heterogeneous assembly and consist mainly of relic pines from the Cretaceous–Tertiary period. Morphological, biochemical, and molecular data all indicate that Mediterranean hard pines are less uniform than the Asian taxa (Klaus, 1989; Schirone et al., 1991; Krupkin, Liston, and Strauss, 1996). Some Asian pines have been suggested to have close relationships with Mediterranean pines. The Himalayan *P. roxburghii* and *P. wallichiana* have been considered as close relatives of *P. canariensis* of the Canary Islands and *P. peuce* of the Balkan Peninsula, respectively, for instance (Mirov, 1967; Klaus, 1989). However, recent analyses of chloroplast (cp) DNA restriction site data and ITS sequences have suggested high levels of divergence among them (Wang and Szmidt, 1993; Liston

[1] Manuscript received 8 December 1998; revision accepted 25 March 1999.

The authors thank Dr. G. G. Vendramin for the DNA samples of Mediterranean pines, M. M. Ribeiro for supplying samples of *P. pinaster* and *P. canariensis,* Dr. J. W. Leverenz for samples of *P. aristata* and *P. balfouriana,* Dr. L. M. Xiong for the sample of *Cathaya,* Dr. N. Tani for the collection of *P. wallichiana* and *P. roxburghii* from Nepal, Forest Tree Breeding Center, Japan, for the access to their collection, and Dr. N. Minaka for valuable discussions on phylogeny reconstruction. This study was conducted during XRW's stay in Japan as an STA fellow, sponsored by JISTEC/JRDC, Japan. Support was also provided by the Grant-in-Aid Baikal Drilling Project from the Science and Technology Agency of Japan. AES acknowledges support from SJFR and from the Carl Tryggers Foundation, Sweden.

[4] Author for correspondence.

et al., 1999). Therefore, more evidence is needed to clarify the relationships among this group of pines.

CpDNA sequences, especially the *rbcL* gene, have been used extensively to infer plant phylogenies, including those of a number of gymnosperms (e.g., Bousquet et al., 1992; Chase et al., 1993; Gadek and Quinn, 1993; Brunsfeld et al., 1994). However, some studies have shown that this coding sequence alone is sometimes too conserved to clarify relationships between closely related taxa (Doebley et al., 1990; Plunkett, Soltis, and Soltis, 1997; Xiang, Soltis, and Soltis, 1998). Following the use of *rbcL*, the *matK* gene has become another sequence candidate for phylogenetic analysis. Recent studies have demonstrated the utility of *matK* for resolving lower level relationships in angiosperms (Johnson and Soltis, 1994, 1995; Steele and Vilgalys, 1994; Liang and Hilu, 1996; Xiang, Soltis, and Soltis, 1998). However, *matK* sequence divergence and its phylogenetic application in *Pinus* and other conifers have not been previously investigated.

In this study we selected four cpDNA regions for sequencing: *rbcL*, *matK*, the *trnV* intron, and the spacer between the *rpl20* and *rps18* genes. Considering the close relationships among pines within each subgenus, we selected *matK* to complement the *rbcL* information. Noncoding sequences tend to evolve faster than coding sequences and, thus, may provide more informative characters for phylogeny reconstruction. The *trnV* intron and the *rpl20-rps18* spacer were selected for this reason, in the expectation that they might provide more variable characters for better phylogenetic tree resolution at the tips. We included all the Mediterranean pines, most of the Asian, and four American pines in this study. In addition, six taxa representing six different genera of Pinaceae were selected as outgroups to *Pinus*. Our main objectives in the study presented here were: (1) to compare sequence divergence of coding and noncoding regions in *Pinus* and Pinaceae; (2) to evaluate the relative utility of the different sequences for phylogenetic inferences in *Pinus*; (3) to provide additional information for the assessment of relationships among and between the Asian and Mediterranean pines; and (4) to reexamine the classification of several uncertain taxa in the light of our new sequence data.

MATERIALS AND METHODS

Species sampled—Species sampled in this analysis are listed in Table 1. Our sampling mainly focused on Eurasian pines, including 17 species of subgen. *Pinus* and 15 species of subgen. *Strobus* (including *P. krempfii*), four of which are American pines. Six species, *Picea abies, Cathaya argyrophylla, Larix decidua, Pseudolarix amabilis, Keteleeria davidiana,* and *Abies numidica,* representing six other genera in *Pinaceae* were selected as outgroup species (Table 1). In total, 38 taxa were included in the present study. All the samples for each species were collected either from documented individuals grown by different institutions or from natural stands (Table 1).

DNA isolation, PCR amplification, and sequencing—Genomic DNA was isolated from needles of individual trees using the cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle, 1987). Four regions (*rbcL, matK, trnV* intron, and *rpl20-rps18*) on the cp genome were selected for polymerase chain reaction (PCR) amplification. The primers used to amplify these regions were designed on appropriate sequences from the whole sequenced cp genome of *P. thunbergii* (Wakasugi et al., 1994). The primer sequences and their positions on the *P. thunbergii* cp genome are presented in Table 2. The PCR reaction mix contained 50–100 ng DNA template, 200 μmol/L of each deoxyribonucleotide (dNTP, GibcoBRL, Life Technologies, USA), 0.5 μmol/L of each of the primer pair, and 1.5 units of Taq DNA polymerase (GibcoBRL) in a total volume of 50 μL. PCR amplification was carried out at 94°C, 3 min for initial denaturation, followed by 30 cycles of denaturation at 94°C for 45 sec, primer annealing at 58°C for 50 sec, extension at 72°C for 80 sec, and termination by 5 min at 72°C.

The PCR products were purified by passage through SUPREC™-02 filter columns (TaKaRa, Japan) to remove the nonincorporated primers and nucleotides. Sequencing reactions were carried out using the BigDye Terminator Cycle Sequencing Kit (Perkin Elmer) according to the manufacturer's instruction on GeneAmp PCR System 9600 (Perkin Elmer). The sequencing reaction products were purified through CENTRI-SEP columns (Princeton Separations Inc., USA) and then applied to ABI 377 automatic sequencer (Perkin Elmer). All the four selected regions were sequenced in both directions for all the operational taxonomic units (OTUs), except for *P. thunbergii,* the sequence for which was retrieved from the EMBL database. Each sequencing run from each of the primer pairs allowed complete overlap of forward and reverse sequences, to ensure high accuracy of data scoring. Automated sequencing output was further checked visually for correction of the automated base calling. The primers used in sequencing are listed in Table 2.

Sequence alignment—The sequences of each species were aligned using Clustal V software, as implemented in Sequence Navigator (ABI, Perkin Elmer, USA) and further modified manually. In most cases the placement of gaps was straightforward. Insertion/deletions (indels) in the aligned sequences were coded as 1/0 binary characters in the data matrix. Gaps of more than 1 bp in length and shared by two or more taxa were treated as a single event. Overlapping gaps were treated as multiple-event length mutations and positioned to minimize the number of required mutational events for creation of the indel. All gaps were weighted equally. Separate alignments were carried out for the subgen. *Pinus* and *Strobus,* for all the 32 *Pinus* taxa, and for all the 38 taxa, including the outgroups.

Phylogenetic analysis—Four kinds of phylogenetic analyses differing in the treatment of gaps were carried out. In the first analysis, gaps were treated as missing data, sequences across the gaps were included, and indels were coded as binary characters. In the second analysis, sequences across the alignment gaps were excluded, but each indel was coded as a binary character. In the third analysis, both sequences across the alignment gaps and the coded indels were excluded. In the fourth analysis, indels were excluded and only the point substitutions were included. Parsimonious analysis of the four data sets produced nearly identical topology. Thus, only the results from scheme 1 are presented in this paper. Maximum parsimony analysis was performed using the PAUP v. 3.1.1 program (Swofford, 1993). Heuristic searches were performed with random sequence addition with 100 replicates, MULPARS, tree-bisection-reconnection (TBR) branch swapping, and ACCTRAN branch length optimization. All character states, including indels, were specified as unordered and equally weighted. To evaluate relative robustness of the clades found in the most parsimonious trees, bootstrap (Felsenstein, 1985), consistency index (CI) (Kluge and Farris, 1969), retention index (RI) (Farris, 1989), and decay index (Bremer, 1988; Donoghue et al., 1992) were calculated. Decay indices were calculated using the AutoDecay program v. 4.0 (T. Eriksson, Department of Botany, Stockholm University, Sweden). The bootstrap analysis was conducted with simple sequence addition, 1000 replicates, and nearest-neighbor interchanges (NNI) branch swapping. Sequence divergence in different regions was computed as the average number of nucleotide differences per site between two sequences according to Nei (1987; Eqs. 10.5 or 10.6, uncorrected *p* distance), and Jukes and Cantor (1969), using the DnaSP 3.0 program (Rozas and Rozas, 1999). The distance

TABLE 1. List of species sampled. *Pinus* classification according to Little and Critchfield (1969). Plant material origin: 1 Lac Duong, Lamdong, Vietnam; 2 Umeå, Sweden; 3 Forestry and Forest Product Research Institute, Japan; 4 Tree Breeding Center, Japan; 5 Hørsholm Aboretum, Denmark; 6 Chiang Mai, Thailand; 7 Royal Botanical Garden Godawari, Nepal; 8 Nanjing Forest University, China; 9 Institute of Forest Genetics at Placerville, USA; 10 DNA samples from G.G. Vendramin, 11 Escola Superior Agraria de Castelo Branco, Portugal; 12 Sichuan, China; 13 Wakasugi et al. (1994) published sequence.

Subgenus	Section and subsection	Species	Source	DDBJ/EMBL/GenBank accession no.			
				rbcL	*matK*	*trnV* intron	*rpl20-rps18*
Ducampopinus	Sect. *Ducampopinus*						
	Subsect. *Krempfianae*	*P. krempfii* Lecomte	1	GBAN-AB019794[a]	GBAN-AB019831	GBAN-AB019868	GBAN-AB019905
Strobus	Sect. *Strobus*						
	Subsect. *Cembrae*	*P. cembra* L.	2	GBAN-AB019795	GBAN-AB019832	GBAN-AB019869	GBAN-AB019906
		P. pumila (Pallas) Regel	3	GBAN-AB019796	GBAN-AB019833	GBAN-AB019870	GBAN-AB019907
		P. koraiensis Sieb. et Zucc.	3	GBAN-AB019797	GBAN-AB019834	GBAN-AB019871	GBAN-AB019908
	Subsect. *Strobi*	*P. strobus* L.	3	GBAN-AB019798	GBAN-AB019835	GBAN-AB019872	GBAN-AB019909
		P. monticola Dougl.	4	GBAN-AB019799	GBAN-AB019836	GBAN-AB019873	GBAN-AB019910
		P. parviflora Sieb. et Zucc.	3	GBAN-AB019800	GBAN-AB019837	GBAN-AB019874	GBAN-AB019911
		P. wallichiana Jackson	3, 7	GBAN-AB019801	GBAN-AB019838	GBAN-AB019875	GBAN-AB019912
		P. kwangtungensis Chun	8	GBAN-AB019802	GBAN-AB019839	GBAN-AB019876	GBAN-AB019913
		P. peuce Grisebach	4, 5	GBAN-AB019803	GBAN-AB019840	GBAN-AB019877	GBAN-AB019914
		P. armandii Franchet	4	GBAN-AB019804	GBAN-AB019841	GBAN-AB019878	GBAN-AB019915
	Sect. *Parrya*						
	Subsect. *Balfourianae*	*P. aristata* Engelm.	5	GBAN-AB019805	GBAN-AB019842	GBAN-AB019879	GBAN-AB019916
		P. balfouriana Grev. et Balf.	5	GBAN-AB019806	GBAN-AB019843	GBAN-AB019880	GBAN-AB019917
	Subsect. *Gerardianae*	*P. gerardiana* Wall.	9	GBAN-AB019807	GBAN-AB019844	GBAN-AB019881	GBAN-AB019918
		P. bungeana Zucc.	3	GBAN-AB019808	GBAN-AB019845	GBAN-AB019882	GBAN-AB019919
Pinus	Sect. *Pinus*						
	Subsect. *Sylvestres*	*P. sylvestris* L.	2	GBAN-AB019809	GBAN-AB019846	GBAN-AB019883	GBAN-AB019920
		P. tabuliformis Carr.	4	GBAN-AB019810	GBAN-AB019847	GBAN-AB019884	GBAN-AB019921
		P. thunbergii Parl.	13	GBAN-D17510	GBAN-D17510	GBAN-D17510	GBAN-D17510
		P. merkusii Jungh. et De Vriese	6	GBAN-AB019811	GBAN-AB019848	GBAN-AB019885	GBAN-AB019922
		P. hwangshanensis Hsia	8	GBAN-AB019812	GBAN-AB019849	GBAN-AB019886	GBAN-AB019923
		P. kesiya Royle	6	GBAN-AB019813	GBAN-AB019850	GBAN-AB019887	GBAN-AB019924
		P. densiflora Sieb. et Zucc.	3	GBAN-AB019814	GBAN-AB019851	GBAN-AB019888	GBAN-AB019925
		P. massoniana Lamb.	8	GBAN-AB019815	GBAN-AB019852	GBAN-AB019889	GBAN-AB019926
		P. yunnanensis Franchet	8	GBAN-AB019816	GBAN-AB019853	GBAN-AB019890	GBAN-AB019927
		P. nigra Arnold	5	GBAN-AB019817	GBAN-AB019854	GBAN-AB019891	GBAN-AB019928
		P. pinaster Aiton	11	GBAN-AB019818	GBAN-AB019855	GBAN-AB019892	GBAN-AB019929
		P. halepensis Mill.	10	GBAN-AB019819	GBAN-AB019856	GBAN-AB019893	GBAN-AB019930
		P. brutia Ten	10	GBAN-AB019820	GBAN-AB019857	GBAN-AB019894	GBAN-AB019931
		P. heldreichii Christ	10	GBAN-AB019821	GBAN-AB019858	GBAN-AB019895	GBAN-AB019932
	Sect. *Pinea*						
	Subsect. *Pineae*	*P. pinea* L.	10	GBAN-AB019822	GBAN-AB019859	GBAN-AB019896	GBAN-AB019933
	Subsect. *Canarienses*	*P. canariensis* C. Smith	11	GBAN-AB019823	GBAN-AB019860	GBAN-AB019897	GBAN-AB019934
		P. roxburghii Sargent	7	GBAN-AB019824	GBAN-AB019861	GBAN-AB019898	GBAN-AB019935
Outgroups		*Picea abies* (L.) Karst.	2	GBAN-AB019825	GBAN-AB019862	GBAN-AB019899	GBAN-AB019936
		Larix decidua Mill.	4	GBAN-AB019826	GBAN-AB019863	GBAN-AB019900	GBAN-AB019937
		Abies numidica De Lannoy ex Carr.	5	GBAN-AB019827	GBAN-AB019864	GBAN-AB019901	GBAN-AB019938
		Keteleeria davidiana (Bertr.) Beissn.	3	GBAN-AB019828	GBAN-AB019865	GBAN-AB019902	GBAN-AB019939
		Pseudolarix amabilis (Nelson) Rehder	3	GBAN-AB019829	GBAN-AB019866	GBAN-AB019903	GBAN-AB019940
		Cathaya argyrophylla Chung et Kuang	12	GBAN-AB019830	GBAN-AB019867	GBAN-AB019904	GBAN-AB019941

[a] The prefix GBAN- has been added for linking the online version of *American Journal of Botany* to GenBank and is not part of the actual DDBJ/EMBL/GenBank accession number.

TABLE 2. Primer sequences for template PCR amplification and sequencing. The primer positions are relative to the *Pinus thunbergii* chloroplast genome.

Region	Sequence 5′–3′ (F: forward; R: reverse)	Position
rbcL	1F: CAGCAGCTAGTTCAGGACTC	43102
	1R: ACAATGGCCTACTTCTTCAC	43598
	2F: GGACATACGCAATGCTTTAG	43511
	2R: CCCTGCTTATTCCAAAACTT	44032
	3F: ACCCAATTTTGGTTTGATAG	43937
	3R: ATGTCACCAAAAACAGAGACT	44453
matK	1F: GAACTCGTCGGATGGAGTG	1530
	1R: GAGAAATCTTTTTCATTACTACAGTG	2017
	2F: CGTACTTTTATGTTTACAGGCTAA	1928
	2R: TAAACGATCCTCTCATTCACGA	2567
trnV intron	F: GTAGAGCACCTCGTTTACAC	47451
	R: CTCGAACCGTAGACCTTCTC	48015
rpl20-rps18	F: CTTCGTCGTTTGTGGATTAC	31377
	R: AGTCGATTTATTAGTGAGCA	31946

matrices for all pairwise sequence combinations were analyzed with the neighbor-joining (NJ) method of phylogenetic tree construction (Saitou and Nei, 1987) with 1000 bootstrap replications, using the program Clustal X (Thompson, Higgins, and Gibson, 1994)

RESULTS

Sequence characterization—Our *rbcL* sequence included 1331 nucleotides for all the 38 OTUs. The *rbcL* gene in *P. thunbergii* is 1428 bp long and is located between positions 43046 and 44473 (Wakasugi et al., 1994). Our 1331-bp sequence starts at position 43122 and ends at 44452, covering 93.2% of the gene. There is no insertion/deletion in this region, and all the 38 aligned sequences have the same length (Table 3).

The *matK* gene in *P. thunbergii* is 1548 bp long and is located in the intron of *trnK* between positions 1715 and 3262 (Wakasugi et al., 1994). Our primers for the *matK* region cover about half (863 bp) of the *matK* gene and 176 bp of the 3′-flanking region within the *trnK* intron. Relative to the *P. thunbergii* cp genome, our *matK* sequence lies between positions 1539 and 2577. Length variation was found in this region among the 38 OTUs. Within subgen. *Pinus*, *P. canariensis* has the longest sequence (1052 bp) and *P. nigra* the shortest (1033 bp). All the Asian members of subgen. *Pinus*, except for *P. merkusii* and *P. roxburghii*, have the same length as *P. thunbergii* (1039 bp). *Pinus merkusii*, *P. roxburghii*, and the other members of Mediterranean pines, as well as all the species of subgen. *Strobus* have a length of 1046 bp. Among the six outgroups, *P. abies*, *L. decidua*, and *P. amabilis*, have sequences 1046 bp long. The *K. davidiana* and *C. argyrophylla* sequences are 1058 and 1051 bp long, respectively. *Abies numidica* has the longest sequence (1060 bp).

The aligned sequence length for the *matK* region is 1076 bp, and it contains nine indels of different lengths (1–12 bp). Most of the indels (five out of nine) were introduced by the outgroups, and they are mainly located in the *matK* 3′-end and the 3′-flanking region. When only the 32 *Pinus* species are included in the alignment, the aligned sequence length is 1052 bp (Table 3). Two deletions in the *matK* 3′-flanking region, one of 6 bp and

TABLE 3. Summary of sequence variation among the analyzed species and phylogenetic tree statistics, excluding uninformative characters.

Region	Taxa group	Alignment length (bp)	Variable sites (Informative)	Indels (Informative)	Most parsimonious trees			
					No.	Length	CI	RI
rbcL	Subgen. *Pinus*	1331	24 (10)	0	5	13	0.7692	0.9143
	Subgen. *Strobus*	1331	30 (18)	0	1	22	0.8636	0.9423
	Pinus 32 taxa	1331	57 (38)	0	26	55	0.7091	0.9475
	All 38 taxa	1331	128 (65)	0	1	123	0.5935	0.8845
matK	Subgen. *Pinus*	1052	58 (19)	4 (2)	9	30	0.7000	0.8767
	Subgen. *Strobus*	1046	23 (14)	0	1	16	0.8750	0.9512
	Pinus 32 taxa	1052	91 (51)	4 (2)	9	74	0.7568	0.9624
	All 38 taxa	1076	208 (102)	9 (3)	9	168	0.6964	0.9213
trnV intron	Subgen. *Pinus*	548	9 (5)	0	5	6	0.8333	0.9444
	Subgen. *Strobus*	547	8 (5)	0	1	5	1.0000	1.0000
	Pinus 32 taxa	548	16 (13)	1 (1)	5	20	0.7500	0.9677
	All 38 taxa	555	44 (20)	5 (2)	10	38	0.6579	0.9347
rpl20–rps18	Subgen. *Pinus*	582	12 (4)	5 (1)	1	5	1.0000	1.0000
	Subgen. *Strobus*	555	8 (5)	0	9	7	0.7143	0.8750
	Pinus 32 taxa	582	32 (25)	7 (3)	21	34	0.8529	0.9838
	All 38 taxa	608	102 (56)	19 (6)	18	107	0.6729	0.9182
All regions combined	Subgen. *Pinus*	3513	103 (38)	9 (3)	6	58	0.7069	0.8768
	Subgen. *Strobus*	3479	69 (42)	0	2	54	0.7963	0.9035
	Pinus 32 taxa	3513	196 (127)	12 (6)	6	190	0.7316	0.9591
	All 38 taxa	3570	482 (243)	33 (11)	12	450	0.6378	0.9046

TABLE 4. Placement of phylogenetic informative indels in the aligned sequences from representative taxa. Positions correspond to the 38 OTUs alignment matrix for each sequence region. Dashes represent alignment gaps; dots indicate the same nucleotide as the reference *P. thunbergii.*

	matK						*trnV* intron			
	8	8	9	9	1	1	1	1	2	2
	5	8	4	6	0	6	2	3	8	9
Species	5	3	9	6	5	0	8	8	4	1
P. thunbergii	GAAAACTTA-	----------TTTGACC	ACTAA-	-----TAGGAAA	TGAAG-AGCAG		CCGGT-AAGAT		TGACACTT	
P. sylvestris	.....	.....	.....	.....	.....		.....	.....	.....	
P. merkusii	.....	.....	.....AATGAT.G	.....	.....G	.....	.....	.....	.....	
P. roxburghii	.....	.....	.....AATGAT	.....	.....G	.....	.....	.....	.....	
P. brutia	.....	.....	.....AATGAT	.....	.....G	.....	.....	.....	.....	
P. pinea	.....	.....	.....AATGAT	.....	.....G	.....	.....	.....	.....	
P. canariensis	.....T	.....	.....AATGAT	.....	.....G	.....	.....	.....	.....	
P. krempfii	.....	.....	.....AATGAT	.....	.....G	.....	..C.G-	.....	...-	....
P. aristata	.....	.....	.....AATGAT	.....	.....G	.....	..C.G-	.....	...-	....
P. strobus	.....	.....	.....AATGAT	.....	.....G	.....	..C.G-	.....	...-	....
P. abies	.....	.....	.....AATGAT	.....	.....G	.....	..A.GA	.....	.....	
C. argyrophylla	.....	.....	.....AATGAT	.....	.....G	.....	.TA.GA	....G	....G	...
L. decidua	.....	.....	.....AATGAT	.....	.....G	.....	..A.G-	.....	.....	
P. amabilis	.....	.....	....GAATGAT	....G..	.....G	.....	..A.G-	.....	.....	
K. davidiana	.....TGG.AAACTAAACTTA	.....	.....AATGAT	....G..	.....G	.....	..A.G-	.....	.....	
A. numidica	T....TAG.-	----AAACTTA...T...	.....AATGAT	....G..	.....G	.....	..A.G-	.....	.....	

another of 1 bp, were found in the Asian members of subgen. *Pinus,* but not in *P. merkusii* and *P. roxburghii* (Table 4). In addition, an insertion of 6 bp and a deletion of 6 bp were found in the coding region of *matK* in *P. canariensis* and *P. nigra,* respectively. No alignment gaps were found among the taxa of subgen. *Strobus.* Among the nine indel characters, only three are phylogenetically informative (Table 4).

The *trnV* intron in *P. thunbergii* is 543 bp long, lying between positions 47471 and 48013 (Wakasugi et al., 1994). Our sequence for this region includes the whole intron and six nucleotides from the 3′-end of the *trnV* exon1. The length variation in this region is very minor (Table 3). Five gaps were found in the aligned matrix among the 38 OTUs, four of 1 bp and one of 5 bp in length. Four of the five gaps were introduced by the outgroup taxa, and only two of the five indel characters are phylogenetically informative (Table 4). The aligned sequence length for this region is 555 bp. Subgenus *Pinus* and subgen. *Strobus* differed by 1 bp in length.

The spacer between *rpl20* and *rps18* proved to be the most length-variable region among the four analyzed in this study. Our sequence for this region covers half (186 bp) of the *rpl20* gene (360 bp), the spacer in between (256 bp), and half (134 bp) of the *rps18* gene (303 bp), between positions 31383 and 31958 on the *P. thunbergii* cp genome. The sequence length varied between 555 bp in subgen. *Strobus* to 590 bp in *A. numidica.* The aligned sequence length is 608 bp (Table 3). Nineteen gaps of 1–20 bp were found in the aligned sequences, but only six are informative (Table 4). All the gaps, except for a 4-bp insertion in *A. numidica,* a 6-bp insertion in *C. argyrophylla,* and a 20-bp deletion in subgen. *Strobus* within *rpl20,* were found in the spacer region between the *rpl20* and *rps18* genes. Subgenus *Strobus* differed from subgen. *Pinus* by having a 20-bp deletion at the 3′-end of *rpl20* and a 1-bp deletion in the spacer region (Table 4). No length variation was found within subgen. *Strobus.* Within subgen. *Pinus,* five gaps of 1–5 bp were introduced into the spacer by including *P. merkusii, P. pinea,* and *P. canariensis.* The other gaps were introduced by the addition of the outgroup species.

Sequence divergence—For *P. krempfii, P. peuce,* and *P. wallichiana* each region was sequenced for two individuals. The two samples of each species gave identical sequences on all the four regions analyzed. The other taxa were each sequenced using one individual. All the 148 sequences (37 OTUs and four sequences each) reported in this paper will appear in the DDBJ/EMBL/GenBank nucleotide sequence databases with accession numbers from AB019794 to AB019941 (Table 1).

Sequence variation in each region is summarized in Table 3. When all the 32 *Pinus* species were compared, 57 variable sites were found in the *rbcL* region, 91 in the *matK* region, 16 in the *trnV* intron, and 32 in the *rpl20-rps18* region. The inclusion of the six outgroups introduced much additional variation to all the four regions. The variable sequence characters among all the 38 OTUs numbered 128 for *rbcL,* 208 for *matK,* 44 for *trnV* intron, and 102 for *rpl20-rps18* (Table 3). When the four regions were combined, the total data matrix for the 38 OTUs consisted of 3570 sequence characters and 33 binary 1/0 indel characters. There were 482 variable sequence sites, of which 243 were phylogenetically informative. The 33 coded indels contributed an additional 11 informative characters (Table 3). The positions of these informative indels are presented in Table 4.

The average number of nucleotide substitutions for the four sequences analyzed in this study is presented in Table 5. In general, the sequence divergence is low across DNA regions and clades. The uncorrected distance and Jukes and Cantor (1969) distance gave very similar results, thus only the Jukes and Cantor (1969) measures are cited below. Comparison of nucleotide substitution rates among the four sequences between the two subgenera revealed similarly low divergence within each subgenus, except for *matK,* the sequence divergence for which in subgen. *Pinus* (0.0109) was 1.8 times higher than in subgen. *Strobus* (0.0061). The two noncoding regions did

TABLE 4. Continued.

rpl20-rps18

1 7 9	2 1 6	3 0 6	3 3 4	4 5 3	4 7 2
GAAATA	------GGTTAGGTATAGATGAAATAAATAGG	TTGGATTGGAAATGAA	-----CGAAAAGA	TTTCCC	------ATGGATCA
.....	------.....	.....	-----.....	.....	------.....
.....	------.....	.....	------.....------.....	.....	------.....
.....	------.....	.....	-----.....	.....	------.....
.....	------.....	.....	-----.....	.....	------.....
.....	------.....	.....	-----.....	.....	------.....
.....	------.....	.....	-----.....	.....	------.....
.....	------.....	.....	-----.....	.....	-------.C....TC
.....	------.....	.....	-----.....	.....	-------.C....TC
.....	------.....	.....	-----.....	.....	------.C....TC
.....	------.....G.....	.....	------.....TGAAA.....	.....	-ATTCC-.....G.
.....	GGTATA.....G.....	.G.....	C...TGAAA.A....T.	.....	------.....
.....	------.....G.....	..TT.....	.....CGAAA.....	....T	----------..G.
.....	------.....A.....G.....	.....	.....CGAAA.....	....G	.ATTTCT.....
.....	------.....A.....G.....C	.....	.....CGAAA.....	.....	.ATTTCT.....
.....	------.....A....A.....G.....A	.....	A.....A..CAAAATA.....	.....	.ATTTCT.....

not show higher divergence than the two coding regions within each subgenus. Within subgen. *Strobus,* sequence divergences for *rbcL* (0.0063) and *matK* (0.0061) were similar, and both were higher than that for *trnV* intron and *rpl20-rps18.* However, within subgen. *Pinus,* sequence divergence for *matK* (0.0109) was much higher (2.7 times) than for *rbcL* (0.0041). When the two subgenera were combined, sequence divergence increased noticeably for all the four regions. The nucleotide divergence in genus *Pinus* for the *trnV* intron (0.0109) was lower than the divergence observed for the coding *rbcL* (0.0116) and *matK* (0.0220) sequences (Table 5). Sequence divergence in *rpl20-rps18* (0.0197) was higher than in *rbcL* but lower than in *matK.* The *matK* sequence appears to have evolved 1.9 times faster than the *rbcL* sequence in *Pinus.* A similar pattern of sequence divergence was found among the six outgroups. Although the substitution rate was higher for *rpl20-rps18* (0.0607) than for *matK* (0.0530), the *matK* region, as in *Pinus,* evolved two times faster than *rbcL* (0.0264) among the six outgroup genera (Table 5).

Phylogenetic reconstruction—The overall topology of phylogenetic trees based on individual sequences all strongly supported the monophyly of the genus *Pinus* and the two subgenera, *Strobus* and *Pinus* (data not shown). However, as suggested by sequence variation analysis, the trees based on *trnV* intron and *rpl20-rps18* were poorly resolved at the section level. The *rbcL* gene tree gave weak support for most of the branches. The *matK* tree is nearly identical to the topology of the combined data sets presented in Fig. 1, but with weaker bootstrap values for many branches. Although the overall topologies of individual sequence trees were similar, they differed in the placement of a few unstable taxa. For example, on the *rbcL* tree, *P. krempfii* and *P. peuce* formed one group sister to *P. bungeana* and *P. gerardiana.* However, on the *matK* tree, *P. krempfii* grouped with *P. gerardiana,* being sister to *P. bungeana.* On the *rpl20-rps18* tree, *P. merkusii* grouped together with *P. pinea* and *P. canariensis* rather than with the Asian members of subsect. *Sylvestres.* The combined *matK-rbcL* tree (data not shown) had good resolution, gave strong support for most of the clades, and agreed well with the phylogenetic tree based on the combined four sequences. When all the four regions were combined, maximum parsimonious heuristic search of this data set produced 12 equally parsimonious trees requiring 450 steps (CI = 0.6378; RI = 0.9046). Compared to the individual sequences, the combined data set gave stronger support for internal clades, and more clades received good bootstrap support (>75%) in the analysis of all four regions than in analyses of the separate data sets. The strict consensus tree of the 12 equally parsimonious trees based on the combined data set is shown in Fig. 1.

On the strict consensus tree (Fig. 1) the 32 *Pinus* species were split into two distinct groups corresponding to the subgenera *Pinus* and *Strobus.* Within subgen. *Strobus, P. aristata* and *P. balfouriana* from the sect. *Parrya,* subsection *Balfourianae,* formed a well-supported (99%) basal group. *Pinus bungeana* and *P. gerardiana* from subsect. *Gerardianae,* also in the sect. *Parrya,* formed

TABLE 5. Average number of nucleotide substitutions for different clades, based on uncorrected *p* distance (Nei, 1987) and Jukes and Cantor (1969) distance (in parentheses), averaged for all pairwise comparisons, gap sites were removed.

Region	Subgen. *Pinus*	Subgen. *Strobus*	*Pinus* 32 taxa	Six outgroups
rbcL	0.0041 (0.0041)	0.0062 (0.0063)	0.0115 (0.0116)	0.0260 (0.0264)
matK	0.0108 (0.0109)	0.0061 (0.0061)	0.0216 (0.0220)	0.0511 (0.0530)
trnV intron	0.0042 (0.0042)	0.0030 (0.0030)	0.0107 (0.0109)	0.0207 (0.0210)
rpl20-rps18	0.0038 (0.0038)	0.0043 (0.0043)	0.0193 (0.0197)	0.0583 (0.0607)

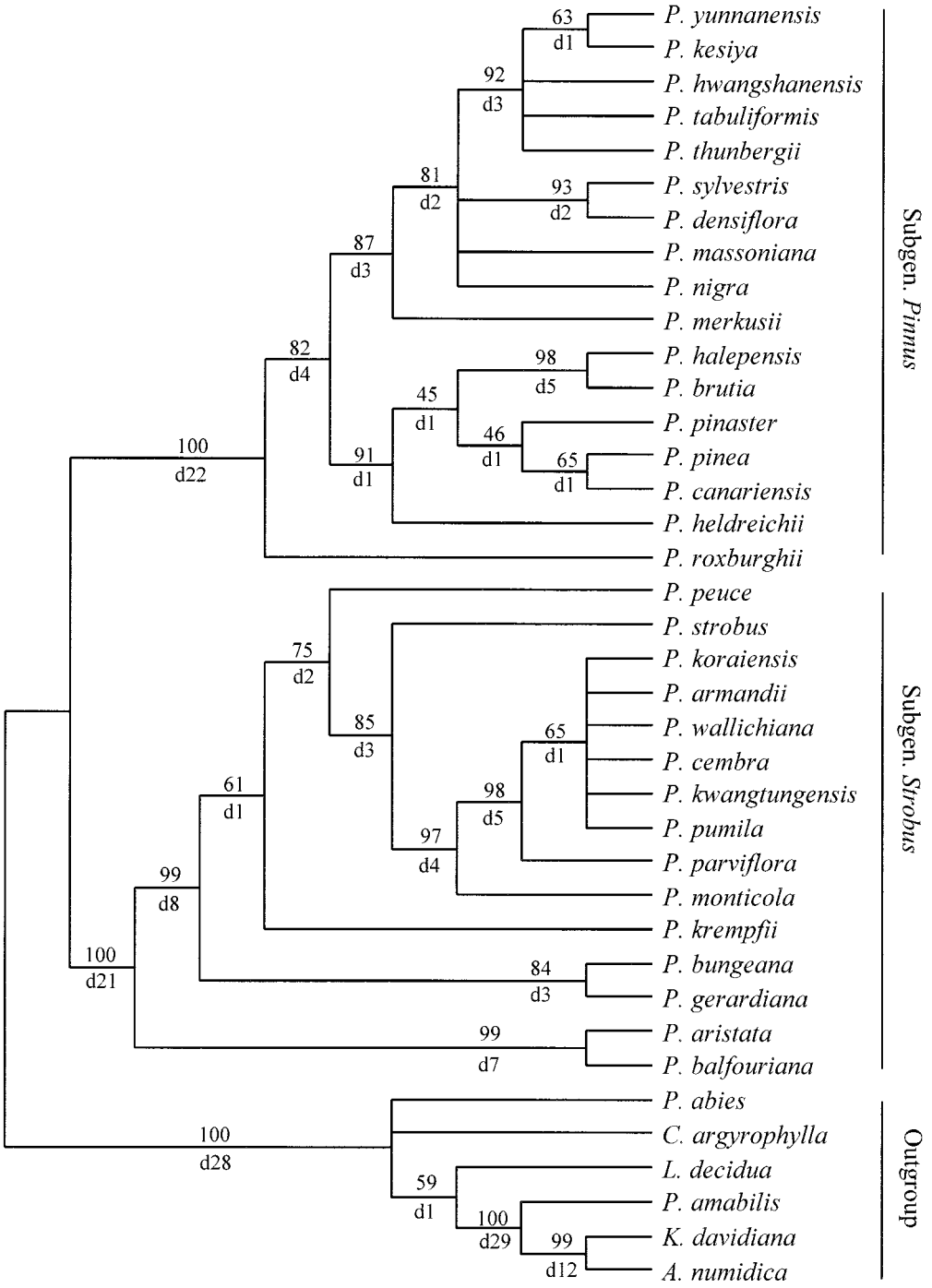

Fig. 1. Strict consensus tree of the 12 most parsimonious trees based on combined sequences from 32 *Pinus* species and six outgroups. Bootstrap percentages (above) of 1000 replicates and decay values (below, preceded by ''d'') are mapped along each branch.

another separate group. *Pinus krempfii* was sister species to the sect. *Strobus* clade. Within the sect. *Strobus* clade, *P. peuce* was sister species to the rest of the section. The two American taxa of subsect. *Strobi, P. strobus* and *P. monticola,* were separated from the Asian members of the subsection. The remaining Eurasian species of subsects. *Strobi* and *Cembrae* formed one unresolved cluster. The topology of the NJ tree (Fig. 2) is essentially the same as the strict consensus tree. However, on the NJ tree *P. krempfii* was grouped together with *P. bungeana* and *P. gerardiana,* with 69% bootstrap support.

In the subgen. *Pinus* clade, the Himalayan *P. roxburghii* was a sister species to all the remaining taxa of the subgenus. Its divergent character was shown consistently on all the individual sequence trees (data not shown). The remaining 16 species were split into two distinct clades. One of these two clades included pines occurring in the Mediterranean. In this clade, the rare and endangered *P.*

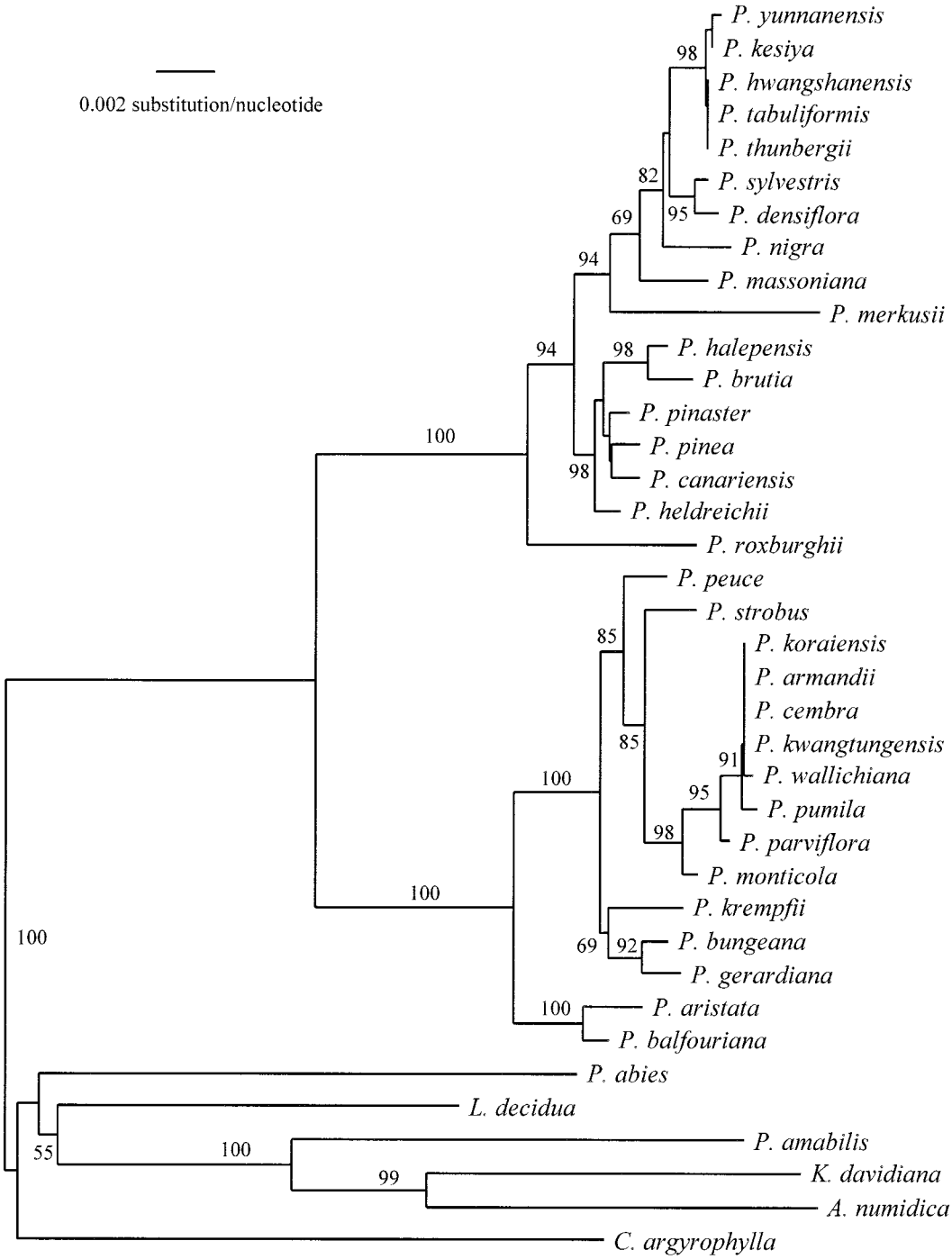

Fig. 2. The neighbor-joining tree based on overall pairwise substitution rates (*p* distance) of the combined sequences. Branch lengths are proportional to the scale of substitution rate given above. Bootstrap percentages of 1000 replicates are noted for each branch. Only bootstrap values higher than 50% are shown.

heldreichii appeared as sister species to the rest of the cluster. *Pinus halepensis* and *P. brutia* formed a strongly supported (98%) group. The remaining species, *P. pinea, P. canariensis,* and *P. pinaster,* formed one group with weak (<50%) bootstrap support on both the NJ and strict consensus trees. The second clade of the subgen. *Pinus,* consisted of species from subsect. *Sylvestres,* including all the Asian members and the European *P. nigra.* In this clade *P. merkusii* was placed on a separate long branch, sister to the rest of the Asian members (Fig. 2). Among the remaining species, *P. yunnanensis, P. kesiya, P. hwangshanensis, P. tabuliformis* (previous spelling *P. tabulaeformis*), and *P. thunbergii* formed one unresolved group, and *P. sylvestris* and *P. densiflora* formed a sep-

arate strongly supported group on both the NJ and strict consensus trees.

DISCUSSION

Sequence divergence in Pinus—Sequence divergence varied considerably among the four cpDNA regions analyzed, as well as among clades for a given sequence. Previous studies have suggested that the *rbcL* gene is conservative within lineages of seed plants (Bousquet et al., 1992; Chase et al., 1993). Our results give further confirmation of the conservative nature of the *rbcL* sequence within *Pinus* and among genera of Pinaceae (Table 5). In this region all the 38 OTUs had the same length, and only 57 variable sites were found among the 32 pines in a sequence 1331 bp long. Based on *rbcL* sequence alone, the topology of the recovered tree was not well resolved, and most branches had weak support.

On the other hand, the *matK* gene has been often found to be more variable than other coding cpDNA sequences tested (Johnson and Soltis, 1994, 1995; Soltis et al., 1996; Plunkett, Soltis, and Soltis, 1997; Xiang, Soltis, and Soltis, 1998). Previous studies in angiosperms have shown that *matK* evolves much faster (2–3 times) than *rbcL* (Johnson and Soltis, 1994; Soltis et al., 1996; Plunkett, Soltis, and Soltis, 1997; Xiang, Soltis, and Soltis, 1998). Thus far, there has been no report on *matK* variation in gymnosperms. Our results provide the first information on this subject. The *matK* sequences analyzed in our study suggested there was a distinctly higher rate of evolution in this region than in the *rbcL* sequence, both within subgen. *Pinus,* and among different genera of Pinaceae. In addition, the variation of *matK* in *Pinus* was even higher than that of the noncoding regions. Surprisingly, however, within our samples of subgen. *Strobus,* unlike in subgen. *Pinus,* the *matK* diverged at a rate very similar to *rbcL*. One scenario that could explain this observation is uneven rate of divergence over time among lineages for different sequences. Another possible explanation is homogenizing sequence evolution within different lineages caused by differing types of recurrent mutations. The unequal evolution rate of different cpDNA sequences within and among lineages found in this study and the interspecific rate heterogeneity reported by Bousquet et al. (1992) indicate that care must be taken when selecting sequence candidates for estimating branching dates.

An unexpected result from this study was the low sequence divergence of the two noncoding regions, especially within each *Pinus* subgenus. This was particularly manifest in the case of the *trnV* intron, which appeared to evolve slower than the *rbcL* sequence and contained only one 1-bp indel in the alignment matrix of the 32 *Pinus* species. The basis for the apparently slow evolution of this intron cannot be elucidated with our data.

Although the sequence divergence across the four DNA regions was generally low within each subgenus, a sharp increase was noticed when the two subgenera were combined. This can result from the differences in mutation sites between the two groups. Although both have similar mutation rates, individual mutations can occur in a different genome region in each group. As a consequence, a sharp increase in divergence would occur when we combine them, because one group contributes changes that do not occur in the other. In fact, this is the case for many of the mutations we observed in our data set, which further stresses the distinct split between the two subgenera.

Phylogenetic implications—*Subgenus Strobus*—One of the most morphologically unique species in *Pinus* is *P. krempfii,* which is endemic to Vietnam. Morphologically it differs from all the other pines by having two flat leaf-like needles rather than typical pine needles (Lecomte, 1921, 1924). Several specific morphological and wood anatomy features giving unusual combinations of characters have made classification of this taxon difficult (Chevalier, 1944; De Ferré, 1948, 1953; Buchholz, 1951; Erdtman, Kimland, and Norin, 1966). It has been suggested that the taxon may represent a link between the genus *Pinus* and other genera of the family Pinaceae such as *Keteleeria* and *Pseudolarix* (De Ferré, 1948, 1953; Mirov, 1967). However, both previous cpDNA RFLP analysis (Wang, Szmidt, and Nguyen, 1999) and the present sequence data do not support this hypothesis. The relationship between *P. krempfii* and *Keteleeria* and *Pseudolarix* is clearly remote. Chevalier (1944) elevated this taxon to an independent monospecific genus in Pinaceae and named it *Ducampopinus krempfii.* Other authors, however, created a separate subgenus, *Ducampopinus,* in the genus *Pinus* to accommodate this taxon (De Ferré, 1953; Gaussen, 1960; Little and Critchfield, 1969). In Pilger's (1926) classification, the species was placed in the same subsection, *Balfourianae,* as *P. aristata* and *P. balfouriana.* Farjon (1984) following the subdivision of Van der Burgh (1973) placed *P. krempfii* in sect. *Parrya,* monospecific subsect. *Krempfianae.* Our previous analysis of cpDNA restriction site variation could only place this taxon in subgen. *Strobus* (Wang, Szmidt, and Nguyen, 1999). In the present study, however, *P. krempfii* was found outside sect. *Strobus* and could not be placed in the subsect. *Balfourianae;* rather it seems to have a closer affinity to subsect. *Gerardianae,* as indicated by the *matK* and *rbcL* trees and the combined NJ tree. Taking into account its unique morphology, our results tend to support the placement of *P. krempfii* in the sect. *Parrya,* monotypic subsect. *Krempfianae.* Although by now the available molecular data clearly suggest the placement of *P. krempfii* in genus *Pinus,* subgen. *Strobus,* the evolution of its unique needle morphology remains to be explained.

Species representing subsects. *Balfourianae* and *Gerardianae* of the sect. *Parrya* were placed in two separate, strongly supported groups. Our results, similar to the conclusions of Perez de la Rosa, Harris, and Farjon (1995) and Liston et al. (1999), also showed that this section is not monophyletic. The distinct character of these subsections has been recognized in most other phylogenetic studies (Strauss and Doerksen, 1990; Wang and Szmidt, 1993). It has been suggested that sect. *Parrya* represents the most ancient pines (Farjon, 1984, 1996; Strauss and Doerksen, 1990). Strauss and Doerksen (1990) suggested that the ancestral species in *Parrya* are perhaps North American and gave rise to the Asian group subsect. *Gerardianae,* which then gave rise to the section *Strobus.* The position of *P. krempfii* revealed in the present study suggests the taxon might represent a link between sect. *Parrya* and sect. *Strobus.* According to Millar (1998, and

references therein) pine originated in the early-middle Mesozoic in middle latitudes. At the beginning of the Mesozoic, there was one landmass. By the early Jurassic, a northern super-continent, Laurasia, separated and began to drift from a southern continent. During the Cretaceous, the genus was already differentiated into the two subgenera, and pines were widely distributed throughout the Northern Hemisphere, indicating that wherever within middle latitudes they originated, their main route of migration was east and west. Cretaceous fossil records of sect. *Parrya,* especially *Balfourianae* and *Gerardianae,* are very poor, making it difficult to track the path of these pines. Nevertheless, our results support the ancient character of sect. *Parrya,* and the divergence between subsects. *Balfourianae* and *Gerardianae* seems very advanced.

The close relationship of subsects. *Strobi* and *Cembrae* has been revealed by several previous analyses (Strauss and Doerksen, 1990; Govindaraju, Lewis, and Cullis, 1992; Wang and Szmidt, 1993). However, as in other phylogenetic analyses, further divisions among pines from these two subsections were not resolved in the present study. The two American species, *P. strobus* and *P. monticola,* from the subsect. *Strobi* were separated from Eurasian members of subsects. *Strobi* and *Cembrae,* indicating advanced divergence between Old World and New World soft pines as well as relatively recent diversification among the Eurasian taxa. Since only four North American pines were sampled in this study, patterns of divergence between North American and Eurasian pines cannot be generalized. A wider sampling of North American pines from different subsections would be necessary for such a comparison. In general, patterns of divergence among species within and between continents would largely depend on their origin and development history. Many of the extant pines were developed from scattered refugia throughout the Tertiary (Millar, 1998). In the sect. *Strobus* clade, *P. peuce* was clearly separated from the remaining members, which suggests the species is distinctly different from others in this section. According to Klaus (1989), *P. peuce* is a small-cone relative of the Himalayan *P. wallichiana* and represents an Eurasian relic pine that has been isolated from other pines of subgen. *Strobus* since Tertiary times (Mirov, 1967; Klaus, 1989). However, from our sequence data *P. wallichiana* was clearly associated with other Asian soft pines, which together formed an unresolved group. Thus, the relationship between *P. peuce* and *P. wallichiana* requires further investigation.

Subgenus Pinus—Within subgenus *Pinus,* the 16 species excluding *P. roxburghii* were split into two strongly supported clades, one containing all Asian members of subsect. *Sylvestres* and *P. nigra,* and the other comprising all the Mediterranean pines of subsects. *Sylvestres, Pineae,* and *Canarienses.* The Himalayan *P. roxburghii* was found as a strongly divergent taxon from all the remaining hard pines. Based on analysis of the ITS region, Liston et al. (1999) found *P. roxburghii* had a sister relationship to the American and Mexican pines of subsects. *Ponderosae, Leiophyllae, Contortae, Oocarpae, Australes,* and *Attenuatae* and that it was paraphyletic to the Asian and Mediterranean hard pines. The strong morphological resemblance of *P. roxburghii* to *P. canariensis* has promoted the classification of the two taxa into the same subsection, *Canarienses* (Little and Critchfield, 1969; Farjon, 1984; Klaus, 1989). Klaus (1989) suggested that *P. roxburghii* originated from Mediterranean ancestors of *P. canariensis* that followed the Tethys coast to the east and reached the Himalayan region in the Upper Cretaceous–Lower Tertiary and led to the rise of *P. roxburghii.* On the other hand, Mirov (1967) suggested an eastern Asia origin for *P. roxburghii,* from where it purportedly migrated to the Himalayas via the mountain ranges that once extended from eastern Asia to the Caucasus and farther west. By this route, Mirov (1967) suggested, the closely related *P. canariensis* reached the Canary Islands. The highly divergent character of *P. roxburghii* revealed by our present and other (Liston et al., 1999) molecular evidence does not clearly support a Mediterranean descent for *P. roxburghii,* rather it suggests a very early split of the *P. roxburghii* lineage from the Mediterranean pines. Alternatively, *P. roxburghii* might represent an ancestral stock to the Eurasian hard pines.

In the clade comprising the Mediterranean pines, *P. heldreichii* was a sister species to the remaining members. *Pinus heldreichii* is an endemic species that grows in southern Italy and the Balkan Peninsula (Mirov, 1967). By some authors this species is called *P. leucodermis* (Farjon, 1984; Schirone et al., 1991; Boscherini et al., 1994). The taxonomic position of *P. heldreichii* remains uncertain, and it has seldom been studied at the molecular level (Schirone et al., 1991; Boscherini et al., 1994). In general, *P. heldreichii* is regarded as more closely related to *P. nigra, P. sylvestris,* and other Asian hard pines than to the true Mediterranean pines (Klaus, 1989). Shaw (1914) even considered it as a variety of *P. nigra.* However, chemical analysis revealed that *P. heldreichii* has a different terpene composition than *P. nigra* (Mirov, 1967). Seed protein analysis revealed a ''divider'' position for *P. heldreichii* between Mediterranean pines and other members of subsect. *Sylvestres,* though it is more closely related to the Mediterranean taxa (Schirone et al., 1991). Our present results clearly support a distinct taxonomic status for this rare and endangered pine and its close affinity to the ''true'' Mediterranean pines.

Within the Mediterranean pine clade, *P. halepensis* and *P. brutia* formed a highly supported (98% on both NJ tree and consensus tree) group. The clear resemblance in their seed protein profiles (Schirone et al., 1991), and allozyme patterns (Conkle, Schiller, and Grunwald, 1988) and their ability to hybridize in nature (Panetsos et al., 1997) all indicate a close relationship between the two. *Pinus brutia* is even described as a variety of *P. halepensis* by some authors (Farjon, 1984). Allozyme (Conkle, Schiller, and Grunwald, 1988) and morphology (Frankis, 1993) studies have suggested that *P. halepensis* is derived from a *P. brutia*-like ancestor and that *P. brutia* has retained greater ancestral variation, showing affinities not only to *P. halepensis* but also to other Mediterranean pines, e.g., *P. pinaster* and *P. canariensis* (Frankis, 1993). Our present results support this suggestion.

Pinus pinaster, P. pinea, and *P. canariensis* formed one group, albeit with weak (<50%) bootstrap support.

Pinus pinea is considered by many authors as an enigmatic and isolated species (Mirov, 1967; Klaus, 1989). Traditionally, *P. pinea* is placed in the monotypic subsect. *Pineae* (Little and Critchfield, 1969; Farjon, 1984). However, our present results do not reveal such a distinct separation of *P. pinea* from other Mediterranean pines. Klaus (1989) noted that *P. pinea, P. pinaster,* and *P. canariensis* share many cone and vegetative characters. Frankis (1993) combined *P. pinaster, P. canariensis, P. halepensis,* and *P. brutia* into one subsection, *Pinaster,* but both authors still placed *P. pinea* in a separate subsection. Our present results lend additional support to the grouping of these species into one subsection, *Pinaster,* suggested by Frankis (1993), but indicate that *P. pinea* may also belong to this subsection.

The Asian members of the subsect. *Sylvestres* formed a strongly (94% on the NJ tree and 87% on the consensus tree) supported monophyletic group that is clearly separated from the Mediterranean clade. In this clade, *P. merkusii* appeared as strongly diverged from all the other members (Fig. 2). Morphological, chemical, and population studies have revealed that *P. merkusii* is very different from other neighboring Asian hard pines (Cooling, 1968; Weissmann and Lange, 1987; Szmidt, Wang, and Changtragoon, 1996). Its distinct separation from the rest of the Asian members of subsect. *Sylvestres* at the molecular level was first reported in a study based on cpDNA restriction site data (Wang and Szmidt, 1993) and was further confirmed by a recent analysis of the nuclear ITS region (Liston et al., 1999). It appears that the distinct character of *P. merkusii* is a result of an early separation and prolonged isolation of this species from other Asian members. During the Jurassic and Cretaceous periods, tropical pines were present in southeastern Asia (Mirov, 1967). It is possible that *P. merkusii* has continued to develop in this region ever since, while the other extant pines migrated to and developed in southeastern Asia not earlier than the Tertiary (Mirov, 1967). Considering all these lines of evidence, it appears that *P. merkusii* could be excluded from subsect. *Sylvestres.* A similar suggestion was made by Frankis (1993). Based on the similarity of cones of *P. merkusii* and *P. brutia,* Frankis (1993) placed the former species in subsect. *Pinaster* together with other Mediterranean pines. On our *rbcL* and *trnV* intron trees the position of *P. merkusii* appeared uncertain. On the *rpl20-rps18* tree this taxon was grouped together with *P. pinea* and *P. canariensis.* However, on the *matK* tree and the combined sequence tree, *P. merkusii* was placed in the same clade as other Asian members of subsect. *Sylvestres.* Thus, the classification scheme proposed by Frankis (1993) is not fully supported by our combined sequence topology. We feel reluctant to express a strong opinion about its placement in subsect. *Sylvestres* or *Pinaster.* Taking into account inconsistent characters of the available morphological and molecular evidence, we believe that additional studies are necessary for its placement in subgen. *Pinus.*

LITERATURE CITED

BOSCHERINI, G., M. MORGANTE, P. ROSSI, AND G. VENDRAMIN. 1994. Allozyme and chloroplast DNA variation in Italian and Greek populations of *Pinus leucodermis. Heredity* 73: 284–290.

BOUSQUET, J., S. H. STRAUSS, A. H. DOERKSEN, AND R. A. PRICE. 1992. Extensive variation in evolutionary rate of *rbcL* gene sequences among seed plants. *Proceedings of the National Academy of Sciences, USA* 89: 7844–7848.

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

BRUNSFELD, S. J., P. S. SOLTIS, D. E. SOLTIS, P. A. GADEK, C. J. QUINN, D. D. STRENGE, AND T. A. RANKER. 1994. Phylogenetic relationships among the genera of Taxodiaceae and Cupressaceae: evidence from *rbcL* sequences. *Systematic Botany* 19: 253–262.

BUCHHOLZ, J. T. 1951. A flat-leaved pine from Annam, Indo-China. *American Journal of Botany* 38: 245–252.

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

CHEVALIER, A. 1944. Notes sur les conifères de l'Indochine. *Revue de Botanique Appliquée et d'Agriculture Tropicale* 24: 7–34.

CONKLE, M. T., G. SCHILLER, AND C. GRUNWALD. 1988. Electrophoretic analysis of diversity and phylogeny of *Pinus brutia* and closely related taxa. *Systematic Botany* 13: 411–424.

COOLING, E. N. G. 1968. *Pinus merkusii.* Commonwealth Forestry Institute, Department of Forestry, Oxford University, Oxford.

DE FERRÉ, Y. 1948. Quelques particularités anatomiques d'un pin Indochinois: *Pinus krempfii. Bulletin de la Société d'Histoire Naturelle de Toulouse* 83: 1–6.

———. 1953. Division du genre *Pinus* en quatre sous-genres. *Academie des Sciences Compte Rendu* 236: 226–228.

DOEBLEY, J., M. DURBIN, E. M. GOLENBERG, M. T. CLEGG, AND D. P. MA. 1990. Evolutionary analysis of the large subunit of carboxylase (*rbcL*) nucleotide sequence among grasses (Gramineae). *Evolution* 44: 1097–1108.

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

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

ERDTMAN, H., B. KIMLAND, AND T. NORIN. 1966. Wood constituents of *Ducampopinus krempfii* (Lecomte) Chevalier (*Pinus krempfii* Lecomte). *Phytochemistry* 5: 927–931.

FARJON, A. 1984. Pines: drawings and descriptions of the genus. E. J. Brill/DR. W. Backhuys, Leiden.

———. 1996. Biodiversity of *Pinus* (Pinaceae) in Mexico: speciation and palaeo-endemism. *Botanical Journal of the Linnean Society* 121: 365–384.

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

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

FRANKIS, M. P. 1993. Morphology and affinities of *Pinus brutia. In* International Symposium on *Pinus brutia* Ten., 11–18. Ministry of Forestry, Ankara.

GADEK, P. A., AND C. J. QUINN. 1993. An analysis of relationships within the Cupressaceae sensu stricto based on *rbcL* sequences. *Annals of the Missouri Botanical Garden* 80: 581–586.

GAUSSEN, H. 1960. Les gymnospermes actuelles et fossiles. Fassicule VI. Les conifèrales. Chap. 11. Généralités, Genre *Pinus. Travaux du Toulous Université Laboratoire Forestier* Tome 2, Sect. 1, vol. 1, 1–272.

GOVINDARAJU, D., P. LEWIS, AND C. CULLIS. 1992. Phylogenetic analysis of pines using ribosomal DNA restriction fragment length polymorphisms. *Plant Systematics and Evolution* 179: 141–153.

JOHNSON, L. A., AND D. E. SOLTIS. 1994. *matK* DNA sequences and phylogenetic reconstruction in Saxifragaceae s. str. *Systematic Botany* 19: 143–156.

———, AND ———. 1995. Phylogenetic inference in Saxifragaceae sensu stricto and *Gilia* (Polemoniaceae) using *matK* sequences. *Annals of the Missouri Botanical Garden* 82: 149–175.

JUKES, T. H., AND C. R. CANTOR. 1969. Evolution of protein molecules. *In* H. N. Munro [ed.], Mammalian protein metabolism, 21–132. Academic Press, New York, NY.

KLAUS, W. 1989. Mediterranean pines and their history. *Plant Systematics and Evolution* 162: 133–163.

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

KRUPKIN, A. B., A. LISTON, AND S. H. STRAUSS. 1996. Phylogenetic analysis of the hard pines (*Pinus* subgenus *Pinus,* Pinaceae) from chloroplast DNA restriction site analysis. *American Journal of Botany* 83: 489–498.

LECOMTE, H. 1921. Un pin remarquable de l'Annam. *Bulletin du Museum National d'Histoire Naturelle Paris* 25: 191–192.

———. 1924. Additions au sujet de *Pinus krempfii* H. Lec. *Bulletin du Museum National d'Histoire Naturelle Paris* 30: 321–325.

LIANG, H. P., AND K. W. HILU. 1996. Application of the *matK* gene sequences to grass systematics. *Canadian Journal of Botany* 74: 125–134.

LISTON, A., W. A. ROBINSON, D. PINERO, AND E. R. ALVAREZ-BUYLLA. 1999. Phylogenetics of *Pinus* (Pinaceae) based on nuclear ribosomal DNA internal transcribed spacer region sequences. *Molecular Phylogenetics and Evolution* 11: 95–109.

LITTLE, E. L., JR., AND W. B. CRITCHFIELD. 1969. Subdivisions of the genus *Pinus* (pines). USDA Forest Service, Washington DC, Miscellaneous Publication Number 1144.

MILLAR, C. I. 1998. Early evolution of pines. *In* D. M. Richardson [ed.], Ecology and biogeography of *Pinus,* 69–91. Cambridge University Press, Cambridge.

MIROV, N. T. 1967. The genus *Pinus.* Ronald Press Company, New York, NY.

NEI, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York, NY.

PANETSOS, K., A. SCALTSOYIANNES, F. A. ARAVANOPOULOS, K. DOUNAVI, AND A. DEMETRAKOPOULOS. 1997. Identification of *Pinus brutia* Ten., *P. halepensis* Mill. and their putative hybrids. *Silvae Genetica* 46: 253–257.

PEREZ DE LA ROSA, J., S. A. HARRIS, AND A. FARJON. 1995. Noncoding chloroplast DNA variation in Mexican pines. *Theoretical and Applied Genetics* 91: 1101–1106.

PILGER, R. 1926. Coniferae. *In* A. Engler and K. Prantl [eds.], Die Naturlichen Pflanzenfamilien, vol. 13 (ed.2) Gymnospermae, 121–403. Wilhelm Engelmann, Leipzig.

PLUNKETT, G. M., D. E. SOLTIS, AND P. S. SOLTIS. 1997. Clarification of the relationship between Apiaceae and Araliaceae based on *matK* and *rbcL* sequence data. *American Journal of Botany* 84: 565–580.

ROZAS, J., AND R. ROZAS. 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* 15: 174–175.

SAITOU, N., AND M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4: 406–425.

SCHIRONE, B., G. PIOVESAN, R. BELLAROSA, AND C. PELOSI. 1991. A taxonomic analysis of seed proteins in *Pinus* spp. (Pinaceae). *Plant Systematics and Evolution* 178: 43–53.

SHAW, G. R. 1914. The genus *Pinus.* Publication of Arnold Arboretum Number 5. Riverside Press, Cambridge, MA.

SOLTIS, D. E., R. K. KUZOFF, E. CONTI, R. GORNALL, AND K. FERGUSON. 1996. *matK* and *rbcL* gene sequence data indicate that Saxifraga (Saxifragaceae) is polyphyletic. *American Journal of Botany* 83: 371–382.

STEELE, K. P., AND R. VILGALYS. 1994. Phylogenetic analyses of Polemoniaceae using nucleotide sequences of the plastid gene *matK*. *Systematic Botany* 19: 126–142.

STRAUSS, S. H., AND A. H. DOERKSEN. 1990. Restriction fragment analysis of pine phylogeny. *Evolution* 44: 1081–1096.

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

SZMIDT, A. E., X.-R. WANG, AND S. CHANGTRAGOON. 1996. Contrasting patterns of genetic diversity in two tropical pines: *Pinus kesiya* (Royle ex Gordon) and *P. merkusii* (Jungh et De Vriese). *Theoretical and Applied Genetics* 92: 436–441.

THOMPSON, J. D., D. G. HIGGINS, AND T. J. GIBSON. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22: 4673–4680.

VAN DER BURGH, J. 1973. Hölzer der niederrheinischen Braunkohlenformation 2. Hölzer der Braunkohlengruben ''Maria Theresia'' zu Herzogenrath, ''Zukunft West'' zu Eschweiler und ''Victor'' (Zülpich Mitte) zu Zülpich. Nebst einer systematisch-anatomischen Bearbeitung der Gattung *Pinus* L. *Review of Palaeobotany and Palynology* 15: 73–275.

WAKASUGI, T., S. I. TSUDZUKI, M. SHIBATA, AND M. SUGIURA. 1994. A physical map and clone bank of the black pine (*Pinus thunbergii*) chloroplast genome. *Plant Molecular Biology Reporter* 12: 227–241.

WANG, X.-R., AND A. E. SZMIDT. 1993. Chloroplast DNA-based phylogeny of Asian *Pinus* species (*Pinaceae*). *Plant Systematics and Evolution* 188: 197–211.

———, ———, AND H. N. NGUYEN. 1999. The phylogenetic position of the endemic flat-needle pine *Pinus krempfii* (Pinaceae) from Vietnam, based on PCR-RFLP analysis of chloroplast DNA. *Plant Systematics and Evolution* (in press).

WEISSMANN, V. G., AND W. LANGE. 1987. Zusammensetzung der Neutralteile des Balsamkolophoniums von *Pinus massoniana* Lamb., *Pinus merkusii* Jungh. und *Pinus luchuensis* Mayr. *Holzforschung* 41: 147–154.

XIANG, Q. Y., D. E. SOLTIS, AND P. S. SOLTIS. 1998. Phylogenetic relationships of Cornaceae and close relatives inferred from *matK* and *rbcL* sequences. *American Journal of Botany* 85: 285–297.