African Journal of Biotechnology Vol. 11(18), pp. 4158-4162, 1 March, 2012
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB11.559
ISSN 1684–5315 © 2012 Academic Journals

Full Length Research Paper

Molecular identification of Cupressaceae (Coniferales) in Vietnam based on 18S-rRNA sequence

Nguyen Minh Tam[1,2] and Nguyen Thi Phuong Trang[1]

[1]Laboratory of Molecular Biology, Institute of Ecology and Biological Resources, Vietnam Academy of Science and Technology (VAST), 18 Hoang Quoc Viet, Cau Giay, Vietnam.
[2]Department of Experimental Taxonomy and Genetic Diversity, Vietnam National Museum of Nature, VAST, 18 Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam.

Accepted 16 February, 2012

We investigated phylogenetic relationships of nine Cupressaceae species from four subfamilies: Cupressoidae, Taxodioidae, Cunninghamioidae and Taiwanioidae using 18S ribosomal DNA sequences. These data indicated that Cupressaceae was divided into two lineages: Taiwanioidae/Cunninghamioidae containing the species in genera *Cunninghamia* and *Taiwania*, and Cupressoidae/Taxodioidae containing the species in the six remaining genera *Cupressus, Calocedrus, Fokienia, Xanthocyparis, Juniperus* and *Crytomeria*. Within the Cupressoidae, *Xanthocyparis* is strongly associated with *Cupressus*, sharing a sister relationship with *Juniperus*.

Keywords: Conifers, Cupressaceae, molecular identification, 18S-rRNA.

INTRODUCTION

Cupressaceae comprises 29 genera and about 161 species grouped within seven subfamilies. The Cupressaceae species occur in various habitats including desert, tundra, tropical forests and temperate forests. According to Luu and Thomas (2004) and Hiep et al. (2004), eight Cupressaceae species were found in Vietnam where the species are restricted to highly fragmented habitats. Many species such as *Cunninghamia lanceolata* var. *konishii* are restricted to adjoining area of Laos, *Xanthocyparis vietnamensis* in Ha Giang (Quan Ba), *Taiwania crytomerioides* in Lao Cai (Van Ban), *Glyptostrobus pensilis* in Dac Lac (Krong Nang, Ea H'leo), *Cupressus tonkinensis* in Lang Son (Huu Lung), but there are a few widespread species such as *Calocedrus macrolepis* and *Fokienia hodginsii* in central and north Vietnam

Previous molecular phylogenetic studies based on nuclear genome (18S-rRNA and 28S-rRNA gene sequences) indicated the relationships between the seven families in Coniferales (Chaw et al., 1995; 1997; Stefanovic et al., 1998). The family of Pinaceae diverges in species and the sister group of the six remaining families of Coniferales. The Taxaceae and Cupressaceae form a monophyletic group. Cheng et al. (2000) determined phylogeny of six genera (*Taxus*, *Pseudotaxus*, *Austrotaxus*, *Amentotaxus*, *Torreya* and *Cephalotaxus*) belonging two families, Taxaceae and Cephalotaxaceae using chloroplast genome and a nuclear ITS-rDNA region. Taxad genera and Cephalotaxaceae are monophyletic with the Taxodiaceae/Cupressaceae clade as their sister group. *Cephalotacus* is basal to the taxad genera, among which are the two clades, *Torreya/Amentotaxus and Taxus/Pseudotaxus/Austrotaxus*. Gadek et al. (2000) combined parisimony analysis of *matK* and *rbcL* genes, together with morphological characters within Cupressaceae and proposed a new infrafamilial classification with seven subfamilies: Cunninghamioideae, Taiwanioideae, Athrotaxidoiseae, Sequoioideae, Taxodioideae, Callitroideae and Cupressoideae. Phylogenetic analysis based on cpDNA, nrITS gene sequences and morphology (Little et al., 2004) and nrITS gene sequences (Xiang and Li, 2005) indicated that *Callitropsis* is monophyletic. Little (2006) combined data from anatomy, biochemistry, micromorphology, reproductive development, reproductive morphology, vegetative morphology together with mol-

*Corresponding author. E-mail: ngtam@hn.vnn.vn.

Table 1. List of analyzed species and GenBank accession numbers.

Species	Collected location	GenBank accession no
Fokienia hodginsii	Thai An secondary forest, Quan Ba, Ha Giang TA06122404	EU273295
Cunninghamia lanceolata var. *konishii*	Tay Son secondary forest, Ky Son, Nghe An TS07042301	EU273292
Cupressus tonkinensis	Plantation, Huu Lien, Huu Lung, Lang Son HL07051405	EU273296
Taiwania cryptomerioides		D38250 (Chaw et al.,1995)
Xanthocyparis vietnamensis	Bat Dai Son Nature Reserve, Quan Ba, Ha Giang BDS06121902	EU273293
Calocedrus rupestris	Bat Dai Son Nature Reserve, Quan Ba, Ha Giang BDS06121803	EU273294
C.formesana		D85298 Chaw et al., 1997
Juniperus chinensis		D38243 Chaw et al., 1995
Cryptomeria japonica		D85304 Chaw et al., 1997

ecular sequence data (*matK, NEEDLY intron2*, *nrITS, rbcL* and *trnl*) to establish evolutionary relationships within *Cupressoidae. Callitropsis, Cupressus* and *Juniperus* form a monophyletic group. Old World species of *Cupressus* are sister to *Juniperus*. *Callitropsis* and the 16 New World species of *Cupressus* are determined as the sister group to the Old World species of *Cupressus* and *Juniperus* clade.

In this study, we used 18S-rRNA sequences to study phylogenetic relationships of the five Cupressaceae species in Vietnam.

MATERIALS AND METHODS

Taxon sampling

The conifer species used in this study are presented in Table 1, with information on collected location and GenBank accession numbers for previously published DNA sequences. Young leaves or inner barks were collected, wrapped with marked aluminum paper and put in liquid nitrogen. They were transferred to the Laboratory of Molecular Biology, Institute of Ecology and Biological Resources and subsequently, stored at -76°C until ready for use in DNA extraction. The samples collected were *Fokienia hodginsii* (TA06122404) from Thai An secondary forest; *Cunninghamia lanceolata* var. *konishii* (TS07042301) from Tay Son secondary forest; *Cupressus tonkinensis* (HL07051405) from Huu Lien plantation; *Xanthocyparis vietnamensis* (BDS06121902) and *Calocedrus rupestris* (BDS06121803) from Bat Dai Son forest. The samples were identified on basis of previous taxonomic treatments of collected specimens from these localities.

DNA isolation and sequencing

Genomic DNA was extracted from young leaves or inner barks using the modified cetyl trimethylammonium bromide (CTAB) method by Xavier and Karine (2000). About 100 mg of leaves or inner barks were ground in liquid nitrogen. Subsequently, the extraction buffer consisting of 640 µl of CTAB extraction buffer (100 mM Tris-HCl pH 8.0, 20 mM ethylenediaminetetraacetic acid (EDTA) pH 8.0, 1.4 M NaCl and 0.2% β-mercaptoethanol) and 160 µl of 10% CTAB was added, and the mixture was incubated at 60°C for 1 (leaves) or 3 h (inner barks). Then, 500 µl phenol : chloroform : isoamylalcohol (25:24:1) was added to the mixture gently for 5 min to form an emulsion and centrifuged at 10,000 g for 6 min. DNA was precipitated by adding 2/3 volume of cold isopropanol solution and refrigerated for 20 min to get the supernatant. The DNA pellet was washed with 200 µl of 5 M ammonium acetate and 600 µl of absolute ethanol, dried by air pump and dissolved in Tris-EDTA (TE) buffer (10 mM Tris-HCl pH 8.0 and 1 mM EDTA pH 8.0) with 1 µl RNase (1 µg/ml) per 100 µl DNA. The concentration of total DNA was determined using a fluorometer. 18S rRNA genes were amplified through the following PCR cycling profile: an initial heating step at 95°C for 4 min, followed by incubating for 35 cycles of 95°C for 1 min, 65, 58 and 55°C for 45 s, with pairs of primers 18S1, 18S2 and 18S3, respectively and 72°C for 45 s, and completed by incubating at 72°C for 10 min. All PCR reactions were performed in 25 µl volumes using Gene Amp PCR System 2400. Three pairs of primers, 18SF1: 5'-GTG CCA GCA GCC GCG GTA ATT-3' (forward) and 18SR1: 5'-G TTT AAG TTT CAG CCT TGC GAC-3' (reverse), 18SF2: 5'-CCT TCT GAG AAA TCA GAG TGT TTG-3'; 18SR2: 5'-CTT CTC CTT CCT CTA AAT GAT AAG-3'; 18SF3: 5'-TCA AAG ATT AAG CCA TGC ATG TCT-3' and 18SR3: 5'-TAC GAG CTT TTT AAC TGC AAC AAC were used to amplify the 18S rRNA genes with about 1700 nucleotides. We designed these primers on the basis of the 18S-rRNA sequence of *Calocedrus formesana*, with GenBank accession number D85298.

Table 2. Nucleotide base compositions (%) of the 18S-rRNA sequences of nine *Cupressaceae* species.

Codon position	Base				Length (bp)	R=Ts/Tv
	A	C	G	T		
All positions	23.9	23.2	27.9	25.0	1656.7	1.0
1st positions	21.8	22.7	30.0	25.5	550.6	1.6
2nd positions	25.5	22.9	27.4	24.2	552.1	0.7
3rd positions	24.3	23.9	26.4	25.3	554	0.8

Table 3. Number of the base substitutions per site between species for 18S-rRNA.

S/N	Parameter	1	2	3	4	5	6	7	8
1	*Calocedrus rupestris*								
2	*C. formasana*	0.010							
3	*Fokienia hodginsii*	0.021	0.018						
4	*Cunninghamia lanceolata* var. *konishii*	0.023	0.020	0.025					
5	*Taiwania cryptomerioides*	0.019	0.017	0.020	0.015				
6	*Cryptomeria japonica*	0.013	0.010	0.017	0.014	0.010			
7	*Juniperus chinensis*	0.010	0.008	0.016	0.016	0.013	0.007		
8	*Xanthocyparis vietnamensis*	0.014	0.014	0.023	0.023	0.020	0.013	0.009	
9	*Cupressus tonkinensis*	0.035	0.032	0.040	0.038	0.037	0.031	0.027	0.031

PCR products of the 18S rRNA genes were directly sequenced using the primers 18SF1, 18SF2, 18SR2 and 18SF3 on an Avant 3100 Automated DNA sequencer with the Dye Terminator Cycle sequencing kit (PE Applied Biosystems).

Phylogenetic analyses

Sequence alignments were made with ClustalX (Thompson et al., 1997) GenDoc (Nicholas and Nicholas, 1997) and adjusted manually. We used MEGA4 (Tamura et al., 2007) to analyse our data. Nucleotide sequence divergence between two species was estimated as numbers of synonymous and nonsynonymous substitutions per site, using the method of Nei and Gojobori (1986) with Juke and Cantor (1969) correction. The transition (Ts) and transversion (Tv) were also estimated separately at the first two codon positions using Kimura's (1980) two parameter method. A neighbor-joining (NJ) tree was reconstructed from Ts values and the relative support for each node was tested by the bootstrap method with 1000 replicates.

Parsimony analysis were also conducted by PAUP*, version 4.0 (Swofford, 1993) to infer phylogeny based on nucleotide substitutions. We used the heuristic search with three options: tree bisection reconnection (TBR) branch swapping, the MULPARS option and stepwise addition to look for the most parsimonious (MP) trees. To obtain an estimate of the strength of support for each node, the bootstrap method with heuristic search was performed as well.

RESULTS

The aligned 18S-rRNA sequences for nine *Cupressaceae* species were 1668 bp in length, of which 123 nucleotide sites were variable and 19 were parsimony informative. The *Cupressaceae* species in the study had GC contents ranging from 50.8 (*Xanthocyparis vietnamensis*) to 51.6% (*Taiwania cryptomerioides*), an average of 51.1% (Table 2). The mean base compositions were 23.9, 23.2, 27.9 and 25% for A, C, G and T, respectively. For all sequence pairs when comparing the average rate of transition to transversion was 0.93. The level of transition/transversion rate was high for pyrimidines (2.87) and lower for purines (1.26). The rates of different transitional substitutions were found varying from 7.42 (A to G) to 17.79% (T to C), an average of 12.63, while those of transversional substitutions were lower, an average of 6.18 (5.75 in C to A or C to G to 6.94% in G to T or G to C). Parsimony analysis generated a single most parsimonious tree with a tree length of 197, a consistency index of 0.924, a retention index of 0.722 and a homoplasy index of 0.076.

Among the three codon positions, the base compositions were 51.7 and 50.3% in CG at the first and second position, respectively. The CG content of 50.3% was also found at the third position. For all nucleotide pair comparisons, the average of Tv (transitions)/Ts (transversions) rate at the first position was 1.6 times, higher than those at two remaining codon positions (Table 2).

To estimate evolutionary divergence between sequences for the nine Cupressaceae species, the number of base substitutions per site between sequences was calculated (Table 3) using the maximum composite likehood method. Pairwise comparison between sequences indicated that sequence divergences ranged from 0.7 (*Cryptomeria japonica*/*Juniperus chinensis*) to 4% (*F. hodginsii*/*C. tonkinensis*), with an average of 2%.

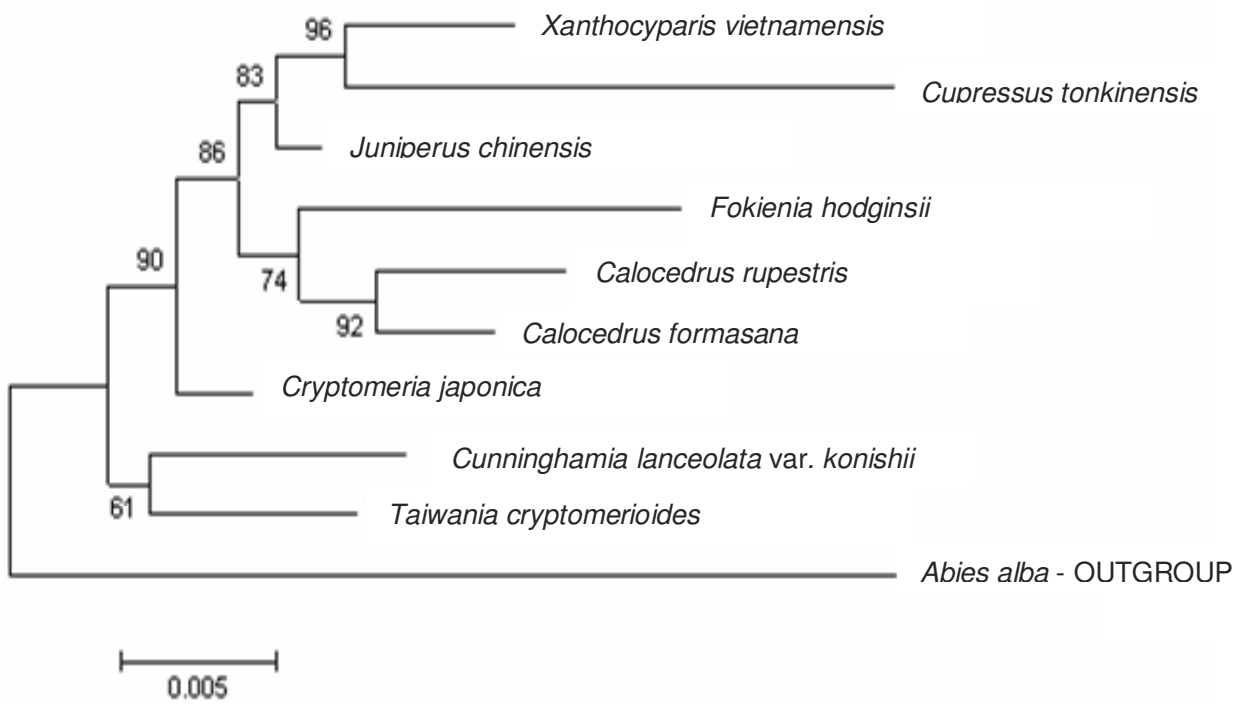

Figure 1. Neighbor-joining tree of *Cupressaceae* based on 1668 bp of the 18S-Rrna.

The two *Calocedrus* species differed from each other (1%), *F. hodginsii* by 2 to 2.3% and *C. lanceolata* var. *konishii* by 1.7 to 1.9%. The highest divergences were found between *C. tonkinensis* and the remaining species, ranging from 2.7% with *T. cryptomerioides* to 4% with *F. hodginsii*. The codon based test of neutrality for analysis between the species using the Nei-Gojobori method (Nei and Gojobori, 1986) indicated that significant deviations were found between *C. tonkiensis* with five species including *F. hodginsii, C. lanceolata* var. *konishii, T. cryptomerioides, C. japonica* and *J. chinensis* ($p < 0.05$). Similarly, significant deviations were found between *Xanthocyparis vietnamensis* with *Calocedrus* species and within *Calocedrus* species. Test of homogeneity of substitution patterns between *Cupressaceae* species using the disparity index per site (Kumar and Gadagkar, 2001) also indicated that significant deviations were found between *T. cryptomerioides* with *Calocedrus* species, *J. chinensis* and *X. vietnamensis* ($p<0.05$). Estimating evolutionary rate between species among *Cupressaceae* using the Tajima (1993) method for testing molecular clock hypothesis, the results indicated that the rate between *F. hodginsii* and *T. cryptomerioides* using *C. lanceolata* var. *konishii* as an outgroup had significance ($\chi^2 = 7.53$, $p<0.05$). Similarly, this rate between *F. hodginsii* and *X. vietnamensis* using *C. rupestris* as an outgroup was found ($\chi^2 = 4.0$, $p<0.05$). Nucleotide diversity was determined (0.0194) by Tajima (1993) method to test for nine *Cupressaceae* species with nucleotide mutation hypothesis.

The NJ analysis with nine *Cupressaceae* species using *Abies alba* as the outgroup based on the distance matrix of partial sequences of 18S-rRNA genes revealed that the *Cupressaceae* species form a monophyletic clade (Figure 1). The parsimony jacknife analysis indicated that the Cupressoidae are placed in a single clade together with Taxodioidae presented by *C. japonica*, characterized by high bootstrap value (90%). The remaining clade including two subfamilies, Cunninghamioidae with present species of *Cunninghamia lanceolata* var. *konishii* and Taiwanioidae (*Taiwania cryptomerioidae*) was poorly supported by a frequence of 61%. Within the Cupressoidae, monophyly of the two genera (*Xanthocyparis/Cupressus*) was supported by 96% bootstrap value. *Juniperus* was sister to the monophyletic *Xanthocyparis/Cupressus* subclade (83% bootstrap value).

DISCUSSION

Four subfamilies of *Cupressaceae* were recognized in our phylogenetic analysis of the nuclear 18S ribosomal DNA sequences in *Cupressaceae* showed that two monophyletic groups were distinctly separated. One group included subfamilies Cupressoidae and Taxodioidae and the second one included the two remaining subfamilies, Cunninghamioidae and Taiwanioidae. Our data shows that Taxodoideae and Cupressoideae were sister groups. This relationship was recovered by preview analyses of the 18S-rRNA sequence data (Chaw et al., 1997), cp DNA (Kusumi et

al., 2002) and chloroplast genes bcL and atpB; and 18S and 26S-rRNA sequences (Catarina et al., 2002). Here, the phylogenetic analyses support the relationship of two subfamilies Cupressoideae and Taxodioideae in family Cupressaceae.

Within Cupressoideae, our data shows that there were two sister clades *X. vietnamensis*/*C. tonkinensis*/*J. chinensis* and *F. hodginsii/C. rupestris/C. formasana.* Our results agree with those of Gadek et al. (2000) and Little et al. (2004) who found that *X. nootkatensis* has close affinity with *Cupressu*s. *Juniperus chinensis* was seen as as a distinctive group. *F. hodginsii* is within the genus, *Calosedus*.

REFERENCES

Catarina R, Mari KM, Else MF (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(2): 197-214.

Chaw SM, Sung HM, Long H, Zharkikh A, Li WH (1995). The phylogenetic positions of the conifer genera *Amentotaxus, Phyllocladus* and *Nageia* inferred from 18S-rRNA sequences. J. Mol. Evol. 41: 224-230.

Chaw SM, Zharkikh A, Sung HM, Leu TC, Li WH (1997). Molecular phylogeny of extant gymnosperms and seed plant evolution: Analysis of nulear 18S rRNA sequences. Mol. Biol. Evol. 14: 56-68.

Cheng Y, Nicolson RG, Tripp K, Chaw SM (2000). Phylogeny of Taxaceae and Cephalotaxaceae genera inferred from chloroplast matK gene and nuclear rDNA ITS region. Mol. Phyl. Evol. 14(3): 353-365.

Gadek PA, Alper DL, Heslewood MM, Quinn GJ (2000). Relationships within Cupressaceae sensu lato a combined morphological and molecular approach. Am. J. Bot. 87: 1044-1057.

Hiep NT, Loc PK, Luu NDT, Thomas PL, Farjon A, Averyanov L, Regalado JJr (2004). Conifers of Vietnam: conservation status 2004, Hanoi (in Vietnamese).

Juke TH, Cantor CR (1969). *Evolution of protein moleculas. In Mammalian protein metabolism* Munro HN (ed.). Academic Press, New York, 21: 132.

Kimura M (1980). A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16: 111-120.

Kumar S, Gadagkar SR (2001). Disparity Index: A simple statistic to measure and test the homogeneity of substitution patterns between molecular sequences. Genetics, 158: 1321-1327.

Kusumi J, Tsumara Y, Yoshimaru H, Tachida H (2002). Molecular evolution of nuclear genes in Cuprassaceae a group of conifer trees. Mol. Biol. Evol. 19(5): 736-747.

Little DP (2006). Evolution and circumscription of the true Cypresses (Cupressaceae: Cupressus). Syst. Bot. 31(3): 461-480.

Little DP, Schwarzbach AE, Adams RP, Hsieh CF (2004). The circumscription and phylogenetic relationships of Callitropsis and the newly described genus Xanthocyparis (Cupressaceae). Am. J. Bot. 91: 1872-1881.

Luu NDT, Thomas PI (2004). Conifers of Vietnam, Foreign languages publishing House Hanoi.

Nei M, Gojobori T (1986). Simple methods for estimating the number of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3: 418-426.

Nicholas KB, Nicholas HBJ (1997). GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by the authors.

Stefanovic 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(5): 688-697.

Swofford, DL (1993). PAUP*: phylogenetic analysis using parsimony, version 3.1.1. Champaign, IL: Illinois Natural History Survey,

Tajima F (1993). Simple methods for testing molecular clock hypothesis. Genetics, 135: 599-607.

Tamura K, Dudley J, Nei M, Kumar S (2007). MEGA4: molecular evolutionary genetic analysis (MEGA) sofware version 4.0. Mol. Biol. Evol. 10: 1093.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997). The ClustalIX Windows interface: flexible stragies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research, 25: 4876-4882.

Xavier JL, Karine LA (2000). rapid method for plant genomic instability using Unanchored-Microsatellite primers. Plant Mol. Biol. Report. 18: 283a-283g.

Xiang Q, Li J (2005). Derivation of Xanthocyparis and Juniperus from within Cupressus evidence from sequences of nrDNA internal transcribed spacer region. Harward papers, Bot. 9: 375-382.