Research in Plant Biology, 2(2): 30-37, 2012 ISSN : 2231-5101
www.resplantbiol.com

Research Article

Cloning of *Casuarina equisetifolia* chloroplast ribosomal RNA (rRNA) genes and its application in phylogenetic studies

Satheeshkumar PK[1#] and Gupta AK[1*]

[1]School of Biotechnology, Madurai Kamaraj University, Madurai 625021, Tamilnadu, India
[#]Present address: Centre for Bioseparation Technology, VIT University, Vellore 632014, Tamilnadu, India
*Corresponding author: E. mail: akg54@sify.com

Chloroplast genes are considered to be highly conserved compared to the nuclear counterparts. This feature has been used extensively in the phylogenic studies of different plant species. In this study, a 3kb sequence consists of 23S rRNA (ribosomal RNA), 4.5S rRNA, ITS 2 (Intergenic Transcribed Spacer 2) and ITS 3 from the chloroplast genome of *C. equisetifolia* has been amplified using specific primers. The genomic origin of the PCR Amplicon was confirmed by southern hybridization. Multiple alignment with a number of chloroplast sequences showed very high homology with some of the species grouped along with Casuarinaceae in the phylogenetic system of classification. A basic level phylogeny study was done using the sequences in different combination with the sequences selected from NCBI. In the analysis, *C. equisetifolia* found to be grouped among the tree species which are closely placed in the classificatory system.

Key words: 23S rRNA, Molecular phylogeny, multiple alignment, intergenic transcribed spacer

Molecular phylogenetics is the branch of modern evolutionary biology where the sequence data in the form of nucleotides and amino acids were used to analyze the phylogeny of an organism. This method is based on the concept that all the plants originated from a common ancestor (monophyletic) or for a group of plants there is a common ancestor (polyphyletic). So the occurrence of the same genes in different groups of plants is possible. The variation in the sequence depends on the amount of incidents like mutations, chromosomal aberrations; exchange of genetic material between different groups and the more related group of organisms will have more close line of changes. These possibilities are being utilized in molecular phylogeny to assess the relationships existing among plants.

The major advances in the "life tree" construction using phylogenetic method were done using the rRNA sequence information. Later on, the internal transcribed spacers also were used successfully in various plant and animal species (Jorgensen *et al*., 2003, Conti *et al*., 1999, Conti *et al*., 2000). Large number of reports appeared describing the use of *rbcL* gene in the molecular phylogeny in the inter and intrageneric level (Geraniaceae- Price and Palmer, 1993, Magnoliaceae- Martin and Dowd, 1984a, Onagraceae- Conti *et al.*, 1993, Saxifragaceae- Soltis *et al.*, 1993). Hasebe *et al.* 1994 used the *rbcL* sequence data to analyze

the evolutionary lineage of leptosporangiate ferns. Goremykin *et al*., (1996) studied the evolutionary relationship of gnetales to other angiosperm families. They have used specific primers to amplify the region spanning from the 3′ terminus of the 23S rRNA gene to 5′ terminus of 5S rRNA gene, which include the 4.5S rRNA gene and two intergenic transcribed spacer regions (*cpITS2, cpITS3*) from a total of 43 species belong to different groups.

The order Casuarinales show a mixture of characters (both primitive and evolved), and because of that there was always a dispute regarding their position in the phylogenetic system of classification. The Casuarinaceae were once considered to be among the most primitive flowering plants because of the extreme reduction in their floral and vegetative features. The resemblance of the cone like fruiting clusters to the cones of some gymnosperms and the outward similarities of the jointed branch system to the stems of *Equisetum* were considered to be especially indicative of an ancient origin (Cronquist, 1981).

In this study, chloroplast DNA sequence (complete sequences of 4.5S rRNA, Intergenic Transcribed Spacers 2 (*cpITS* 2), *cpITS* 3 and 23S rRNA) was amplified from the chloroplast genome of *C. equisetifolia* and used in a very basic phylogenetic analysis using the programme MEGA 3.1 to determine its phylogenetic position.

Materials and Methods:

Plant Material and DNA extraction: DNA extracted from the Seedlings obtained from the Dept of Forest, Govt of Tamilnadu, Madurai, was used to amplify the rRNA sequence in the study. Genomic DNA was extracted from 1gm of fresh needles/tender shoot tips collected from the plant using modified CTAB DNA extraction method of Porebski *et al*., (1997).

Amplification of chloroplast DNA: DNA sequence consists of partial 23S, ITS 2, 45S rRNA and ITS 3 was amplified from chloroplast using the primers, 23S forward (5′- GAAGATTGGGAGCTCTGTGC- 3′) and 23S reverse (5′- ACAAGAAGCTGAGCCG ATGT- 3′). The PCR parameters were; 94°C initial denaturation for 3 min followed by 35 cycles of 94°C- 30 sec, 58°C-45 sec and 72°C 45 sec and a 7min- 72°C follows for the final extension. The PCR product was cloned into pGEM-T Easy vector and sequenced. All the sequencing experiments were carried out with the automated DNA sequencer (ABI Prism 377, PE applied Biosystems) at the DST-FIST DNA sequencing facility, School of Biotechnology, Madurai Kamaraj University, Tamil Nadu, India.

Southern analysis: 10µg of genomic DNA was taken and digested with appropriate restriction enzymes, run on a 0.8% agarose gel and the DNA was transferred to the nylon membrane through capillary method. The membrane was probed with the PCR Amplicon labeled with ^{32}P. The protocol followed was described elsewhere.

Chromosome walking to complete 23S rRNA sequence: To complete 23S rRNA gene sequence, degenerate primers were designed using the *Alnus incana* 23S rRNA sequence as template. Two sets of primers designed, namely C23S F1 (5′- GGA AAG GCT TAC GGA TAC C- 3′), C23S R1 (5′- CAA ACC TCC TGG ATG TCT- 3′), C23S F2 (5′- GAC AGC CAG GTT TGC- 3′) and C23S R2 (5′- CCG AGA CAG TGC CCA GA- 3′) were used to amplify the sequence upstream to partial 23S rRNA sequence.

Collection of sequences for phylogenetic analysis: To study the molecular phylogenetic relationships, a total of 31 sequences were selected for 23S rRNA gene from plants belonging to different phylogenetic groups

like thallophyta (4), bryophyta (3) pteridophyta (3), gymnospermae (1) and angiospermae (20) (Table - 1). These sequences were used in multiple alignment and phylogenetic tree construction using the clustalW and Guide Tree programme available in the EMBL website (www. ebi.ac.uk) and the neighborhood joining (NJ) method in the Mega 3.1 programme (Kumar *et al.*, 2004) with a bootstrap index of 100.

Internet resources and software's used: Similarity search of the nucleotide sequences was carried out at the server of National Center for Biotechnology Information using the blastn (www.ncbi.nlm.nih.gov/blast) programme. Identification of open reading frames on the sequences was done at www.searchlauncher.bcm.tmc.edu/translate. Primer designing was done using the software available from www.frodo.wi.mit .edu/cgi-bin/primers3 and the primer designing programme available in the GCG software. Melting temperature and GC content of the designed primers were checked using the programme obtained from www.primerbiosoft.com/netprimer. The sequence data used in the phylogenetic analysis was collected from the NCBI server (www.ncbi.nlm.nih.gov/nucleotide). Multiple alignments were performed at www.ebi.ac.uk/tools/clustalw. Phylogenetic trees were made using the programmes guide tree, available in the EMBL website or MEGA 3.1 (Kumar *et al.*, 2004) obtained from the site, www.megasoftware.asu.edu.

Results:

Amplification of 23S rRNA gene from chloroplast of *C. equisetifolia*

DNA extracted from the shoot tips/needles was used in the PCR experiments to amplify 23S rRNA sequence using the primers 23S forward and 23S reverse. A 1.08Kb DNA amplified in the PCR was cloned and sequenced. The sequence showed a maximum of 99% identity with ITS 3 (226bp), 4.5S (99bp), ITS 2 (99bp) and part (658bp) of 23S rRNA gene of chloroplast from *Alnus incana* (data not given).

Southern analysis

Genomic origin of 1.08 Kb PCR fragment was confirmed by southern analysis. For this, 12µg of total DNA, extracted from the needles of *Casuarina* using the modified CTAB protocol, was digested with restriction enzymes, *Bam*HI, *Hind*III and *Xho*I. The DNA transblotted nylon membrane was then probed with radiolabelled 1.08 Kb PCR fragment. The membrane was exposed to X-ray film under dark conditions. Two fragments of 9.2Kb and 0.9Kb (Figure - 1, lane 3), single fragments of 9.4Kb (lane 1) and 1.8Kb (lane 2) were generated by *Xho*I, *Bam*HI and *Hind*III respectively confirming the genomic origin of 1.08Kb PCR amplicon.

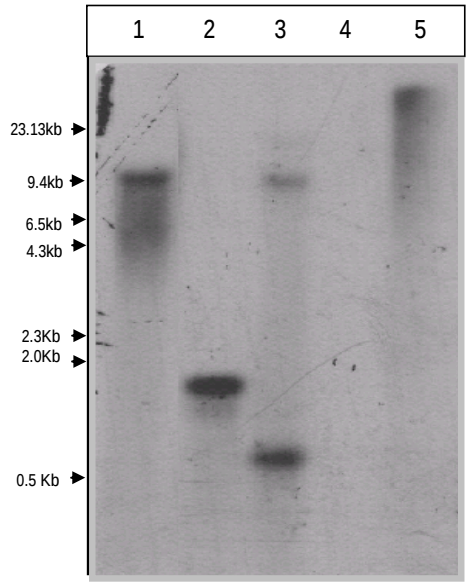

Figure 1. Southern analysis. 10 µg of total DNA was digested with different restriction enzymes, *Bam*HI (lane - 1), *Hind*III (lane - 2) and *Xho*I (lane - 3). The DNA was then transblotted on to a nylon membrane and probed with the 1.08 Kb PCR fragment. Lane - 4 - Blank and lane - 5 – Undigested DNA.

Completion of the sequence

Out of 1.08kb DNA from *C. equisetifolia*, 556bp showed 99% identity with the 3′ end of *Alnus incana* 23S rRNA. The primers designed based on Alnus sequence (C23S F1, C23S R1 and C23S F2, C23S R2) amplified 1.04Kb and 0.93Kb regions running from 5′ to 3′. Sequencing of the amplified and cloned DNA confirmed the 23S rRNA sequence.

Table 1. Sequences of 23S rRNA genes from plants belonging to different phylogenetic groups selected for phylogenetic analysis

Sl. No	Source plant	Accession No.	Sequence length (bp)
1	*Alnus incana* (Ang)	M75722	2811
2	*Amborella trichopoda*	AJ506156	2816
3	*Arabidopsis thaliana*	AP000423	2810
4	*Atropa belladonna*	AJ316582	2811
5	*Bruguiera gymnorrhiza*	AF355767	2867
6	*Calycanthus fertilis*	AJ428413	2804
7	*Casuarina equisetifolia*	AF525938	2643
8	*Ceriops tagal*	AF355762	2873
9	*Lotus corniculatus*	AP002983	2816
10	*Nicotiana tabacum*	Z00044	2810
11	*Nymphaea alba*	AJ627251	2817
12	*Oenothera elata*	AJ271079	2809
13	*Oryza sativa*	X15901	2884
14	*Panax ginseng*	AY582139	2809
15	*Pisum sativum*	X55033	2812
16	*Populus deltoides*	AY029747	2836
17	*Rhizophora stylosa*	AF355761	2894
18	*Saccharum officinarum*	AE009947	2888
19	*Spinacia oleracea*	AJ400848	2810
20	*Triticum aestivum*	AB042240	2888
21	*Zea mays*	Z00028	2888
22	*Anthoceros formosae* (Bry)	AB087478	2819
23	*Marchantia polymorpha*	X04465	2811
24	*Physcomitrella patens*	AP005762	2817
25	*Adiantum capillus-veneris* (Pter)	AY178864	2811
26	*Huperzia lucidula*	AY660566	2810
27	*Psilotum nudam*	NC003386	2819
28	*Pinus koraiensis* (Gymno)	AY228468	2801
29	*Chlorella vulgaris* (Algae)	NC003888	3616
30	*Cyanidium caldarium*	NC001840	2919
31	*Porphyra purpurea*	NC000925	2888
32	*Guillardia theta*	NC000926	2887

Gymno-Gymnosperms, Bry-Bryophytes, Pter-Pteridophytes, Ang-Angiosperms

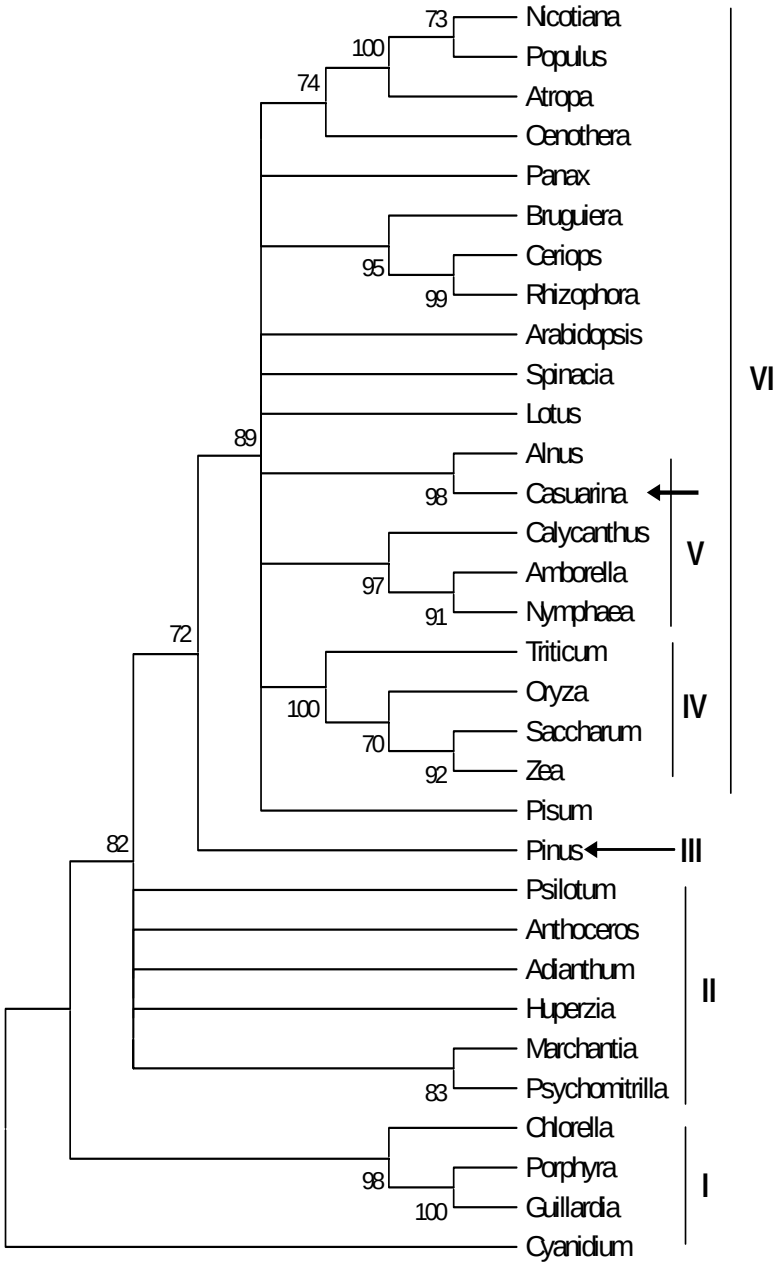

Figure- 2 Phylogenetic tree constructed with the nucleotide sequence of 23S rRNA sequence. 31 sequences were retrieved from the database (Table-1) with the sequence length ranging from 2539-3616bp and the multiple sequence alignment and phylogeny tree construction were performed with the neighbor joining method using the programme MEGA 3.1. The values significant statistically were given in the cladogram. The members in the first group represents the Thalophytes (I), second Cryptogams (II), third Gymnosperms (III), fourth monocots (IV), fifth lower dicots (V) and sixth with both monocots and dicots. Bold arrow represents the position of *Casuarina equisetifolia.*

Multiple alignment and phylogenetic tree construction with 23S rRNA gene

The 23S rRNA gene (1-2643bp) sequence was used as a query against the nucleotide sequence database. From the hits, a total of 31 sequences were selected for 23S rRNA gene from plants belonging to different phylogenetic groups like thallophyta (4), bryophyta (3) pteridophyta (3), gymnospermae (1) and angiospermae (20) (Table - 1). The sequence length varies from 2643bp (*Casuarina*) to 3613bp (*Chlorella*) with an average sequence length of 2.8Kb.Multiple alignment performed with the 23S rRNA gene sequences obtained from 31 different plant genera along with the 23S rRNA gene of *C. equisetifolia* showed considerable degree of similarity throughout the sequence. The *Casuarina* 23S rRNA gene found to be aligning for a stretch of 1-2569bp with all the other sequences but the remaining 104bp showed very less identity with the rest of the sequences. Among the sequences, positions between 936-960, 1459-1502 and 2208-2252 were found to be highly variable among different group of plants (data not shown). For *Casuarina*, a pairwise alignment score of 93 and 92 was obtained with *Alnus incana* and *Calycanthus fertilis* respectively.

In the phylogenetic tree constructed, all the angiosperms clustered into a single cluster and other groups form separate clades (Figure - 2). The single gymnosperm *Pinus* used in the study, separated from both cryptogams and angiosperms. Thallophytes (*Chlorella, Porphyra, Guillardia* and *Cyanidium*) found to be forming the base of the tree. The bootstrap values obtained with pteridophytes were somehow not statistically significant. However, *Casuarina* found to be clustered along with *Alnus incana*, - belong to the same subgroup in which *Casuarina* is placed- with a significant bootstrap value. This clade appeared with the basal angiosperms *like Nymphaea, Lotus, Calycanthus* and *Amborella*.

Discussion

Casuarinaceae are a Gondwanic family with a unique combination of morphological characters not comparable to any other family. The phylogenetic system of classifications made by using the morphological and other features are found to be incomplete as there always been a conflict with regard to some of the plant groups. Molecular phylogeny uses the macromolecular data accumulated in the form of DNA and amino acid sequence for the analysis of evolutionary relationship exist between different plant groups (Wang *et al.*, 2000, Templeton 2001). The whole idea of molecular phylogeny is based on the observation that, by comparing homologous molecules from different organisms it is possible to establish their degree of similarity, thereby revealing a hierarchy of relationships among them (Penny, 2002). There are numerous reports showing the significance of molecular phylogeny as it provided an additional tool to confirm the present position of a particular plant/group or to solve the disagreement over the positioning of a particular species.

Out of, nuclear and plastid gene sequences that could have been used, the later have been widely used for phylogenetic analysis since the plastid genomes are believed to have a common ancestry. (Raven and Allen, 2003, Price and Palmer, 1993, Conti *et al.*, 1993, Hasebe *et al.*, 1994, Odintosova and Yurina, 2003). Among various chloroplast genes, rRNA genes were the first being used for phylogenetic studies (Martin and Dowd, 1984a, Martin and Dowd 1984b, Zimmer *et al.*, 1989, Gielly and Taberlet, 1994, Troitsky *et al.*, 1991, Goremykin *et al.*, 1996). Primikirios (2000) used the 16S and 23S rRNA sequence separately to study the phylogenetic relationship exists in the family vitaceae. The results obtained from nine sequences belonged to different genera of vitaceae were

supporting the existing system of classification. Similar results were obtained when Santose *et al*., (2002) used 23S rRNA sequences to study the phylogeny in symbiotic dinoflagellates. In this study, when the 23S rRNA sequence was used for the phylogenetic analysis, *Casuarina* found to be clustered along with the closest member used in the study, *Alnus incana*. The clade where *Casuarina* appeared was one among the lower dicots.

In conclusion, the analyses produced the results which showed a very high resemblance with the existing phylogenetic system of classification. The plants placed under the subgroup Hamamelidae (eg. *Alnus*, *Quercus* etc) found to be aligning with this species in the analysis performed. The separation of this species was so distinguishable from other groups of plants like pteridophytes or gymnosperms. Therefore it can be concluded that, the results obtained were supporting the existing system of classification by Cronquist (Cronquist, 1988).

Acknowledgements

AKG is gratefully acknowledging the financial support from the Department of Forest, Govt. of Tamilnadu, and Dept of Biotechnology Govt. of India. SPK acknowledge CSIR, Govt. of India for providing a research fellowship.

References

Conti, E, Fischbach, A, Sytsma, KJ. 1993. Tribal relationships in Onagraceae: implications from *rbcL* sequence data. Ann. Mo. Bot. Garden. 80: 672- 685.

Conti, E, Soltis, DE., Hardig, TM, Schneider, J. 1999. Phylogenetic relationships of the silver saxifrages (*Saxifraga*, sect. Ligulatae Haworth): implications for the evolution of substrate specificity, life histories and biogeography. Mol. Phylo. Evol. 13: 536–555.

Conti, E, Suring, E, Boyd, D, Jorgensen, J, Grant, J, Kelso, S. 2000. Phylogenetic relationships and character evolution in *Primula* L.: the usefulness of ITS sequence data. Plant Biosystems. 134: 385–392.

Cronquist, A. 1981. *An Integrated System of Classification of Flowering Plants*. Columbia Univ. Press: New York

Cronquist, A. 1988. A botanical critique of cladism. Bot. Rev. 53: 107-128.

Gielly, L. Taberlet, P. 1994. The use of chloroplast DNa to resolve plant phylogenies: noncoding versus *rbcL* sequences. Mol. Biol. Evol. 11(5): 769-777.

Goremykin, V, Bobrova, V, Pahnke, J, Troitsky, A, Antonov, A, Martin, W. 1996. Non coding sequences from the slowly evolving chloroplast inverted repeat in addition to *rbcL* data donot support Gnetalian affinities of angiosperms. Mol. Biol. Evol. 13: 383-396.

Hasebe, M, Omori, T, Nakazawa, M, Sano, T, Kato, M. 1994. *rbcL* gene sequence provide evidence for the evolutionary lineages of leptosporangiate ferns. Proc. Natl. Acad. Sci. 91: 5730-5734.

Jorgensen, JL., Stehlik, I, Brochmann, C, Conti, E. 2003. Implications of ITS 2 sequences and RAPD markers for the taxonomy and biogeography of the *Oxytropis campestris* and *O. arctia* (Fabaceae) complexes in Alaska. Am. J. Bot. 90(10): 1470-1480.

Kumar, S, Tamura, K, Nei, M. 2004. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Briefings in Bioinformatics. 5: 150-163.

Martin, PG, Dowd, JM. 1984a. The study of plant phylogeny using amino acid sequences of Ribulose-1, 5- Bisphosphate Carboxylase. 5. Magnoliaceae,

Polygonaceae and the concept of primitiveness. Aust. J. Bot. 14: 301-309.

Martin, PG, Dowd, JM. 1984b. The study of plant phylogeny using amino acid sequences of Ribulose-1, 5- Bisphosphate Carboxylase. 3. Addition of Malvaceae and Ranunculaceae to the phylogenetic tree. Aust. J. Bot. 14: 283-290

Odintsova, MS, Yurina, NP. 2003. Plastid genomes of higher plants and algae: structure and functions. Mol. Biol. 37(5): 649- 662.

Penny, D. 2002. Molecular Evolution: Introduction. In *Encyclopedia of Life Sciences*. Nature Publishing Group, Macmillan.

Porebski, S, Beileym, LG, Baum, BR. 1997. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant. Mol. Biol. Reptr. 15(1): 8-15.

Price, RA, Palmer, JD. 1993. Phylogenetic relationships of the Geraniaceae and Geraniales from *rbcL* sequence comparisons. Ann. Mo. Bot. Garden. 80: 661-671.

Primikirios, NI. 2000. Classification of Vitis vinifera (L) in spermatophytes molecular phylogeny based on chloroplastic 16S and 23S rDNA sequences, (electronic version obtained from the net- The Greek *Vitis* Database).

Raven, JA, Allen, JF. 2003. Genomics and chloroplast evolution: what did cyanobacteria do for plants?, Genome Biol. 4 (3) : 209.1-209.5.

Santose, SR, Taylor, DJ, Kinzie, RA, Hidaka, III, Sakai, M, Coffroth, MA. 2002. Molecular phylogeny of symbiotic dinoflagellates inferred from partial chloroplast large subunit 23S rDNA sequences. Mol. Phylogen. Evol. 23: 97-111.

Soltis, DE, Morgan, DR, Grable, A, Soltis, PS, Kuzoff, R. 1993. Molecular systematics of Saxifragaceae sensu stricto. Am. J. Bot. 80: 1056-1081.

Templeton, AL. 2001. Use of phylogeographic analysis of gene trees to test species status and processes. Mol. Ecol. 10: 779-791.

Troitsky, AV, Melekhovets, YF, Rakhimova, GM, Bobrova, VK, Valiejo-Roman, KM, Antonov, AS. 1991. Angiosperm origin and early stages of seed plant evolution deduced from rRNA sequence comparisons. J. Mol. Evol. 32: 253– 261.

Wang, X, Szmidt, AE, Nguyen, HN. 2000. The phylogenetic position of the endemic flat-needle pine *Pinus krempfii* (Lec., Pinaceae) from Vietnam, based on PCR-RFLP analysis of chloroplast DNA. Plant Systematics and Evol. 220: 21- 36.

Zimmer, EA, Hamby, RK, Arnold, ML, Leblanc, DA, Theriot, EC. 1989. Ribosomal RNA phylogenies and flowering plant evolution. In *The Hierarchy of Life,* Eds. B Fernholm, K Bremer, H Jornvall, pp. 205–214. Amsterdam: Elsevier.