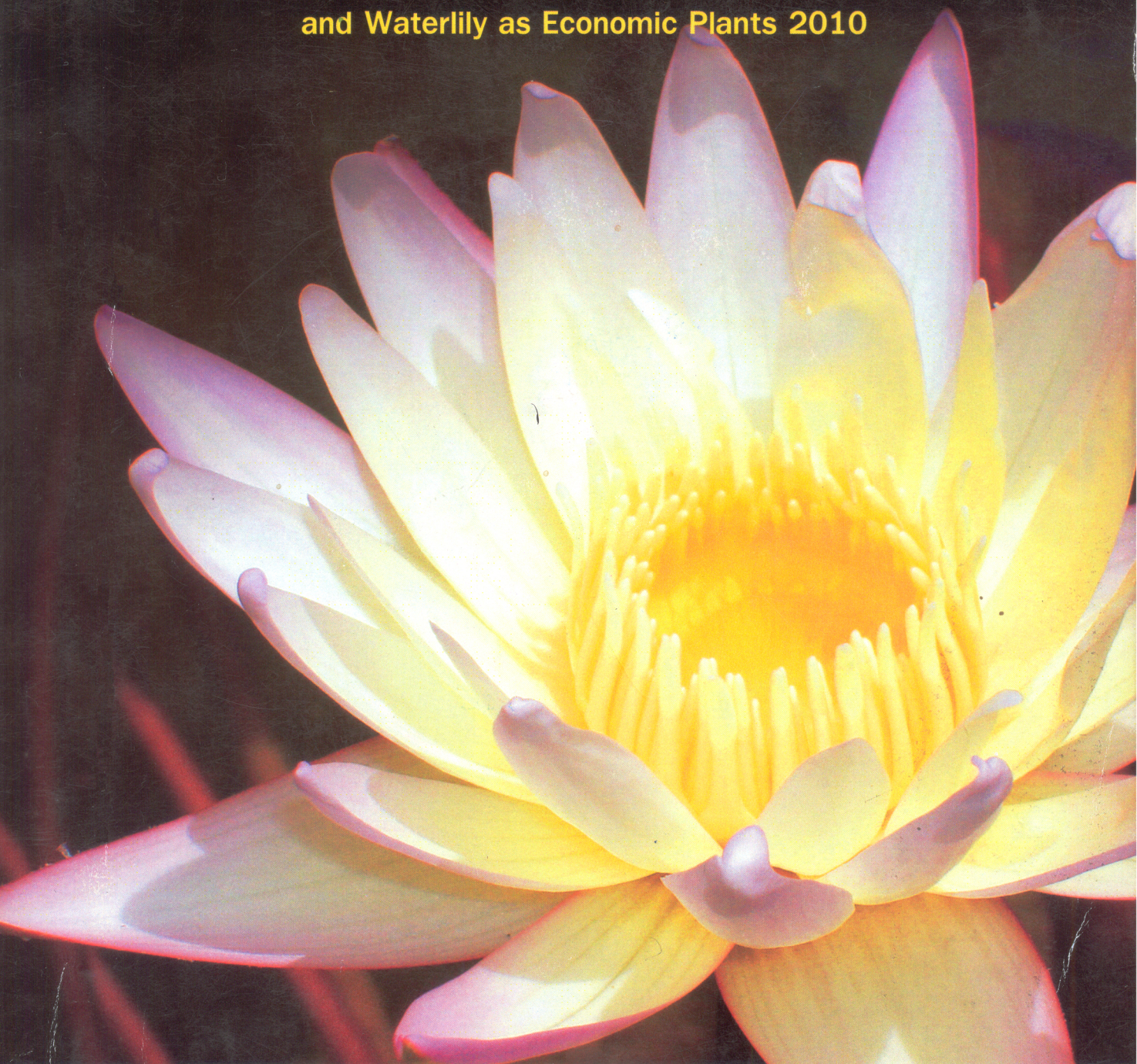




# Proceedings

The 1<sup>st</sup> International Conference on Lotus and Waterlily 2010  
and

The 8<sup>th</sup> Conference on Research and Development of Lotus  
and Waterlily as Economic Plants 2010



October 20 – 24, 2010

Kasetsart University Chalermphrakiat Sakon Nakhon Province Campus, Thailand



# Waterlilies (*Nymphaea*) of India: What do morphology, RAPD, PCR-RFLP and sequence data of the nrDNA ITS region, chloroplast *trnK* intron, *matK* and *rbcL* gene tell us about diversification, hybridization, and adaptation?

Jeremy Dkhar, Suman Kumaria, Pramod Tandon

Plant Biotechnology Laboratory, Centre for Advanced Studies in Botany, North Eastern Hill University, Shillong 793022, India

## ABSTRACT

An investigation of diversification, hybridization, and adaptation in seven Indian representatives of the genus *Nymphaea* viz. *N. alba* var. *rubra*, *N. caerulea*, *N. × marliacea*, *N. nouchali*, *N. pubescens*, *N. rubra* and *N. tetragona* was conducted using morphology, RAPD, PCR-RFLP and sequence data of the ITS region, chloroplast *trnK* intron, *matK* and *rbcL* gene. Phenotypic features with multiple character states (more than two), which may be useful for identification, were recorded for rhizome shape, leaf margin, petal number and color, and pollen shape. Barring *N. caerulea* and *N. nouchali*, RAPD analysis of randomly selected individuals from each species showed low genetic variation at both levels (intra and inter populations). The variability detected for *N. caerulea* may be due to gene flow ( $N_m=0.0737$ ) from another coexisting species (*N. alba* var. *rubra*) or through cross-pollination. In *N. nouchali*, considerable morphological variation accounts for the high genetic variability (68.09%). PCR-RFLP of the ITS region revealed additional fragments, exceeding the expected size of the ITS region when totaled, in *N. alba* var. *rubra*, *N. × marliacea*, *N. rubra* and *N. pubescens*. Sequencing of the ITS region indicated that the extra fragments, manifested as additional signals in the sequencing chromatogram, in *N. pubescens* may be due to random mutations occurring in some of the ITS paralogues. However, recent hybridization and introgression may be the reason for the additional signals depicted in the chromatograms of *N. alba* var. *rubra* and *N. rubra*. This suggestion is supported by the exact nucleotide sequence matches of the chloroplast *trnK* intron, *matK* and *rbcL* gene with their respective putative maternal parent. Furthermore, *matK* gene of *N. tetragona* revealed higher number of non-synonymous substitutions. Molecular evolutionary analysis indicated that three of these sites may be under mild selective pressures. Such adaptive changes at the DNA and protein sequence level of *matK* gene may have been associated with the colonization of *N. tetragona*, suggesting that it could have migrated from China.

## INTRODUCTION

*Nymphaea* is the most diverse genus in the order Nymphaeales and is well represented globally. In India, ten species of *Nymphaea*, both wild (*N. alba*, *N. candida*, *N. nouchali*, *N. pubescens*, *N. rubra* and *N. tetragona*) and cultivated (*N. caerulea*, *N. × marliacea*, *N. micrantha* and *N. alba* var. *rubra*), are reported (Mitra, 1990). However, the classification of *Nymphaea* in India has been reported to be confusing with some names inaccurately used (Cook, 1996). To vindicate Cook's observation, we conducted a molecular based taxonomic revision of some Indian representatives of the genus *Nymphaea* (Dkhar *et al.*, 2010). Interestingly, sequencing signals of the ITS region of *N. rubra* and *N. alba* var. *rubra* revealed several additional signals spanning throughout the



chromatogram, hinting at the presence of divergent ITS paralogues. Could *N. rubra* and *N. alba* var. *rubra* have originated through hybridization? In addition, another interesting observation made from the previous study was the relatively higher number of nonsynonymous substitutions as compared to synonymous substitutions detected in the *matK* gene of *N. tetragona*. The higher rate of nonsynonymous substitutions is usually inferred as an indication of selective pressures acting on protein-coding sequences.

In spite of considerable variation that exist among *Nymphaea* species, morphology-based phylogeny of the genus *Nymphaea* has never been analyzed. An attempt was made by Borsch *et al.* (2008), representing each subgenus by 2-3 species. *N.* subg. *Nymphaea* emerges as first branch followed by subg. *Anecphyra*, and another clade comprising of subg. *Brachyceras*, *Hydrocallis* and *Lotos*. Furthermore, variation and systematics of *N. odorata* was conducted using morphological characters (Woods *et al.*, 2005). Three morphological features i.e. mean leaf blade length-to-width ratio, petiole striping, and lobe apex shape were proposed to distinguish *N. odorata* subsp. *odorata* and *N. odorata* subsp. *tuberosa*.

Studies on genetic variation among populations of *Nymphaea* species using molecular markers such as RAPD, AFLP, etc are limited. The only study to our knowledge was reported by Woods *et al.* (2005) using inter-simple sequence repeats (ISSR) for elucidating the patterns of genetic variation among populations of *N. odorata*.

The objectives of this study are (1) to evaluate the patterns of morphological variability, (2) to assess the genetic diversity at the intra and inter populations' levels, (3) to provide molecular evidence for hybrid origin of *N. rubra* and *N. alba* var. *rubra*, (4) to test whether there was selective pressure in *matK* of *N. tetragona*, and (5) to evaluate the phylogenetic relationships among *Nymphaea* species studied integrating both molecular and non-molecular characters.

## MATERIALS AND METHODS

### *Plant material and taxon sampling*

Out of ten species reported from India, seven *Nymphaea* species viz. *N. alba* var. *rubra*, *N. caerulea*, *N. x marliacea*, *N. nouchali*, *N. pubescens*, *N. rubra* and *N. tetragona* were found.

### *Morphological Analyses*

Thirty three morphological features comprising of 27 qualitative and 6 quantitative characters were examined. Qualitative characters were scored as binary states (present/absent) or unordered multistate characters, whereas quantitative characters were scored as continuous characters. A maximum parsimony based phylogenetic tree was constructed using NONA ver 1.6 (Goloboff, 1993).

### *RAPD analysis and PCR-RFLP of the ITS region*

For RAPD analysis, six primers with clear and reproducible bands were chosen for the amplification reactions of all individuals. Nei's genetic diversity and Shannon's index were used to estimate genetic diversity within and among populations. Partitioning of RAPD variation among and within populations was also estimated by analysis of molecular variance (AMOVA; Excoffier *et al.*, 1992). PCR was used to amplify the ITS region of each species investigated, using universal primers ITS 5 and ITS 4 of White *et*



*al.* (1990). The amplification products were digested with the following ten restriction enzymes: *AluI*, *EcoRI*, *HaeIII*, *HhaI*, *HindIII*, *HinfI*, *MboI*, *MspI*, *RsaI* and *TaqI*. The pair-wise genetic similarities among all pairs of samples were estimated with Jaccard's coefficient (Jaccard, 1908) in NTSYS program (Rohlf, 1998), and a dendrogram was constructed using UPGMA method.

#### *Amplification and sequencing of the ITS region, trnK intron, matK and rbcL gene*

PCR was used to amplify the ITS region, *trnK* intron (including the entire *matK* gene) and *rbcL* gene. The PCR primers ITS 4 and ITS 5 (White *et al.*, 1990) were used to amplify the ITS region. For *trnK* intron amplification, primers *trnK3914F* and *trnK2R* as described by Johnson and Soltis (1995) were used. Amplification of the *rbcL* gene was done using the primers 1-F and 1460-R (Fay *et al.*, 1998).

#### *Cloning of the ITS Region*

The purified PCR amplified product from the ITS region of *N. rubra* and *N. alba* var. *rubra* were cloned into plasmids using CloneJET™ PCR Cloning Kit (Fermentas, Canada). The cloned ITS region was amplified using primers pJET1.2 F and pJET1.2 R supplied along with the kit.

#### *Molecular phylogenetic and evolutionary analyses*

Prior to phylogenetic analyses, sequences obtained were subjected to multiple sequence alignment using Clustal X program (Thompson *et al.*, 1997) with default settings. Neighbor joining (NJ), Maximum parsimony (MP) and Maximum likelihood (ML) methods were used to analyze the aligned sequence data matrix. All trees were constructed using Phylip ver 3.69 (Felsenstein, 1989). In addition, Bayesian inference (BI) of phylogeny was conducted for both separated and combined dataset using MRBAYES v.3.1.2 (Ronquist and Huelsenbeck, 2003). All trees were viewed with the program Tree View 1.5 (Page, 1996). *Nuphar advena*, reported as a first branch of family Nymphaeaceae (Les *et al.*, 1999), has been chosen as outgroup. Molecular adaptation tests on the chloroplast *matK* codon sites were performed using maximum likelihood models and programs included in PAML ver. 4.3 (Yang, 2007).

## RESULT AND DISCUSSION

#### *Morphological and genetic diversity*

As was reported earlier, morphological observation showed considerable variation among the *Nymphaea* species investigated. Rhizome shape, a character to our knowledge not mentioned in earlier studies, could differentiate among subgenus. Members of subg. *Brachyceras* have triangular shaped rhizomes; globular rhizomes are observed in subg. *Lotos*, whereas subg. *Nymphaea* possessed rhizomes that are cylindrical in shape. The morphological feature with considerable variation, as was observed by Wiersema (1988), is petal color; white being the most predominant. Barring *N. rubra*, a pink-flowered waterlily, species of subg. *Lotos* and *Hydrocallis* have white or nearly white petals (Wiersema, 1988). Interestingly, the petal color in *N. rubra* appears in *N. pubescens* as pink tinge at the apex of the petal. This intriguing character may provide additional evidence to our findings, suggesting that *N. rubra* may have originated through hybridization. In addition to earlier reports that suggest them to be night-blooming, *N.*



*pubescens* and *N. rubra* have white colored pollens, probably crucial in attracting pollinators.

The breeding system greatly affects the diversity maintained in a particular population. Selfing tends to produce individuals with low genetic variation. Our studies using RAPD revealed high genetic variability among individuals of *N. caerulea*. This may be attributed to the high percentage of seed formation (100%), suggesting it to be an outcrossing species, or gene flow from a sympatric species, *N. alba* var. *rubra*. The high genetic variability recorded for *N. nouchali* may be due to a probable misidentification of *N. nouchali* JD 02. *N. nouchali* JD 02 showed considerable variation in both molecular and non-molecular characters from the remaining two specimen of *N. nouchali*, i.e. *N. nouchali* JD 06 and *N. nouchali* JD 07. Interestingly, sequence analysis of the ITS, *trnK* intron, *matK* and *rbcL* gene did not reveal any sequence divergence between *N. nouchali* JD 06 and *N. nouchali* JD 07, in spite of the difference in flower color. This manifestation may be brought about by the presence of fungal endophyte, isolated from leaf samples of *N. nouchali* JD 07. In *N. pubescens*, the RAPD markers displayed relatively lower genetic variability as compared to *N. caerulea* and *N. nouchali*. A possibility of gene flow between *N. pubescens* and *N. nouchali*, sympatric species found at Paikan, is indicated ( $N_m = 0.0236$ ). RAPD failed to detect genetic variation among individuals of *N. tetragona*. PCR-RFLP further substantiated our findings, supporting the co-existence of divergent ITS paralogues contributed by both putative parental species. The presence of such divergent ITS paralogues is accountable for the additional restricted fragments observed in *N. alba* var. *rubra*, *N. × marliacea*, and *N. rubra*.

#### *Hybridization in N. rubra and N. alba* var. *rubra*

Sequencing signals of the ITS region of all samples of *N. rubra* and *N. alba* var. *rubra* revealed double peaks in their respective chromatograms, hinting at the presence of divergent ITS paralogues. Such deviating paralogues of the ITS region have been reported in several species of *Nymphaea* (Löhne *et al.*, 2008). The additional signals detected in the chromatogram may be due to random mutations occurring in some of the ITS paralogues (Baldwin, 1995). However, this explanation seems unlikely because on comparison with sequence data of putative parents, the additional signals occur at sites differentiating the suspected parental species. This was further confirmed by means of molecular cloning techniques whereby distinct alleles were isolated from the putative hybrid. Based on the evidence provided by the nrDNA ITS and the chloroplast *trnK* intron, *matK* and *rbcL* gene, we suggest that the Indian material, which has been identified as *N. alba* var. *rubra*, represents a hybrid originating from *N. alba* and *N. odorata*. Similarly, molecular evidence provided here suggests that *N. rubra* is a hybrid produced through a cross between *N. lotus* and *N. pubescens*, with the latter species forming the maternal parent.

#### *Selective pressure during colonization of N. tetragona*

*N. tetragona* is well distributed globally, found throughout China, Japan, Finland, and Russia; but its location in India is restricted to one particular population. Our previous study indicated that the Indian plant is closely related to plants found in China (Dkhar *et al.*, 2010). Molecular evolutionary analysis suggested that the plant taxon migrated from China, and the migratory processes involved might have brought about



some changes at both the DNA and protein sequence levels. This adaptive response would have rendered slight advantage for the plant to survive in new habitat. Such adaptive changes at the DNA and protein sequence levels of *rbcL* gene have been reported in *Schideia* (Kapralov and Filatov, 2006) and *Potamogeton* (Iida *et al.*, 2009).

#### *Phylogenetic relationships*

Borsch *et al.* (2007) conducted the first molecular phylogenetic analysis of *Nymphaea* based on chloroplast *trnT-trnF* region. In their study, three well-supported major lineages within *Nymphaea* are resolved with subg. *Nymphaea* emerging as the first branch followed by a clade comprising subg. *Hydrocallis* and subg. *Lotos*, and another clade comprising subg. *Anecphya* and *Brachyceras*. Barring MP analysis, all other methods of phylogenetic analysis of combined datasets (ITS, *trnK*, *matK*, *rbcL*) corroborated the findings of Borsch *et al.* (2007), placing subg. *Nymphaea* as first branch, with subg. *Lotos* and subg. *Brachyceras* emerging as sister clades. This evidence is further supported by complementing with morphological characters. Within subg. *Nymphaea*, individual data could not resolve the relationships among the three members. On combining all chloroplast markers, *N. alba* var. *rubra* emerged as a separate clade; *N. tetragona* showed close association with *N. × marliacea*. Within subg. *Brachyceras*, represented here by two species i.e. *N. caerulea* and three specimens of *N. nouchali*, incongruence exist between chloroplast and nuclear marker for the phylogenetic placement of *N. nouchali* JD 02. All chloroplast markers support the association of *N. nouchali* JD 02 with *N. caerulea* whereas the ITS data showed it to be closer to the remaining specimen of *N. nouchali*. Molecular evidence provided in our previous study is in disagreement about the taxonomic identity of *N. nouchali* JD 02, suggesting it to be *N. capensis* (Dkhar *et al.*, 2010).

### ACKNOWLEDGEMENT

This work is part of a project that has been sanctioned to PT and SK by the Department of Biotechnology, Ministry of Science and Technology, India (File no. BT/PR-7055/BCE/08/437/2006).

### REFERENCE

- Baldwin, B. G., Sanderson, M. J., Porter, J. M., Wojciechowski, M. F., Campbell, C. S., Donoghue, M. J. 1995. The ITS region of Nuclear Ribosomal DNA: a valuable source of evidence on Angiosperm Phylogeny. *Ann. Miss. Bot. Gard.* 82, 247–277.
- Borsch, T., Löhne, C., Wiersema, J., 2008. Phylogeny and evolutionary patterns in Nymphaeales: integrating genes, genomes and morphology. *Taxon* 57, 1052–1081.
- Borsch, T., Hilu, K.W., Wiersema, J.H., Löhne, C., Barthlott, W., Wilde, V., 2007. Phylogeny of *Nymphaea* (Nymphaeaceae): evidence from substitutions and microstructural changes in the chloroplast *trnT-trnF* region. *Int. J. Plant Sci.* 168, 639–671.
- Cook, C.D.K., 1996. Aquatic and wetland plants of India. Oxford University Press, Oxford.
- Dkhar, J., Kumaria, S., Rama Rao, S., Tandon, P. 2010. Molecular phylogenetics and taxonomic reassessment of four Indian representatives of the genus *Nymphaea*. *Aquat. Bot.* 93, 135–139.



- Excoffier, L., Smouse, P. E., Quattro, M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes; application to human mitochondrial DNA restriction site data. *Genetics* 131, 479–491.
- Fay, M. F., Bayer, C., Alverson, W. S., de Bruijin, A. Y. & Chase, M. W. 1998. Plastid *rbcL* sequence data indicate a close affinity between *Diegodendron* and *Bixa*. *Taxon* 47, 43–50.
- Felsenstein, J., 1989. Phylip-Phylogeny inference package (version 3.2). *Cladistics* 5, 164–166.
- Goloboff, P. A. 1993. NONA v. 1.6. Inst. Miguel Lillo, Tucumán, Argentina.
- Iida, S., Miyagi, A., Aoki, S., Ito, M., Kadono, Y., Kosuge, K. 2009. Molecular Adaptation of *rbcL* in the Heterophyllous Aquatic Plant *Potamogeton*. *PLoS ONE* 4, e4633. doi:10.1371/journal.pone.0004633.
- Johnson, L. A., Soltis, D. E. 1995. Phylogenetic inference in Saxifragaceae sensu stricto and *Gilia* (Polemoniaceae) using *matK* sequences. *Ann. Mo. Bot. Gard.* 82, 149–175.
- Kapralov, M. V., Filatov, D. A. 2006. Molecular adaptation during adaptive radiation in the Hawaiian endemic genus *Schiedea*. *PLoS ONE* 1, e8. doi:10.1371/journal.pone.0000008.
- Les, D. H., Schneider, E. L., Padgett, D. J., Soltis, P. S., Soltis, D. E., Zanis, M. 1999. Phylogeny, classification and floral evolution of water lilies (Nymphaeaceae; Nymphaeales): a synthesis of non-molecular, *rbcL*, *matK*, and 18S rDNA data. *Syst. Bot.* 24, 28–46.
- Löhne, C., Borsch, T., Jacobs, S. W. L., Hellquist, C. B., Wiersema, J. H. 2008. Nuclear and plastid DNA sequences reveal complex reticulate patterns in Australian water-lilies (*Nymphaea* subgenus *Anecphyra*, Nymphaeaceae). *Aust. Syst. Bot.* 21, 229–250.
- Mitra, R. L., 1990. Nymphaeaceae. In: Nayar, M.P., Thothathri, K., Sanjappa, M., (eds.), Fascicles of flora of India, Fascicles 20. Botanical Survey of India, pp. 11–25.
- Page, R. D. 1996. TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12, 357–358.
- Rohlf, F. J. 1998. NTSYS-pc: Numerical taxonomy and multivariate analysis system, v. 2.0. Exeter Software. Setauket, New York.
- Ronquist, F., Huelsenbeck, J. P. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Mongin, F. J., Higgins, D. G. 1997. Clustal X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876–4882.
- White, T.J., Bruns, T., Lee, S., Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, D. G. M., Sninsky, J., White, T., (eds.) PCR Protocols: A guide to methods and applications. Academic Press, California. pp. 315–322.
- Wiersema, J. H. 1988. Reproductive biology of *Nymphaea* (Nymphaeaceae) *Ann. Missouri Bot. Gard.* 75, 795–804.
- Woods, K., Hilu, K. W., Borsch, T. & Wiersema, J. H. 2005. Pattern of variation and systematics of *Nymphaea odorata*: I. Evidence from morphology and inter-simple sequence repeats (ISSRs). *Syst. Bot.* 30, 471–480.
- Yang, Z. 2007. PAML 4: a program package for phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591.