Plant Syst Evol (2012) 298:93–108
DOI 10.1007/s00606-011-0526-z

ORIGINAL ARTICLE

Sequence characteristics and phylogenetic implications of the nrDNA internal transcribed spacers (ITS) in the genus *Nymphaea* with focus on some Indian representatives

Jeremy Dkhar · Suman Kumaria · Satyawada Rama Rao · Pramod Tandon

Received: 9 June 2010 / Accepted: 19 August 2011 / Published online: 10 September 2011

Abstract *Nymphaea*, an aquatic perennial herb with exceptionally beautiful flowers and floating leaves, is well represented globally. Out of ten species reported from India, the internal transcribed spacers (ITS) region of nrDNA was investigated in seven species of *Nymphaea* viz. *N. alba* var. *rubra*, *N. caerulea*, *N.* × *marliacea*, *N. nouchali*, *N. pubescens*, *N. rubra* and *N. tetragona*. Barring *N. pubescens*, whereby double peaks detected in the sequencing chromatograms may be due to random mutations occurring in some of the ITS paralogues, the additional signals detected for *N. alba* var. *rubra* and *N. rubra* are probably influenced by recent hybridization and introgression. Our study on sequence characteristics of ITS 1 and ITS 2 revealed high G + C content (ITS 1, 45.5–48.4%; ITS 2, 50.2–51.5%) and sequence divergence. Percentage of sequence divergence based on substitution and substitution plus indels is 44.15 and 57.19, respectively, for the ITS 1; 29.74 and 47.96% were recorded for the ITS 2. Although highly variable, conserved motifs within the ITS 1 and ITS 2 region of *Nymphaea* were identified and are found to be common throughout the order Nymphaeales. Sequence analysis of the ITS 1 and ITS 2 failed to detect any variation between two morphotypes of *N. nouchali*, namely *N. nouchali* JD 06 and *N. nouchali* JD 07, differing in flower color and found at the same geographical location. However, on comparison with another specimen of *N. nouchali* found at a different location, they showed considerable variation in nucleotide composition. Complemented by sequence data retrieved from GenBank, phylogenetic tree reconstruction of the genus *Nymphaea* based on neighbor-joining, maximum parsimony, maximum likelihood and Bayesian inference methods is presented and discussed.

Keywords *Nymphaea* · ITS 1 · ITS 2 · Sequence divergence · Phylogenetic relationship · Neighbor joining · Maximum parsimony · Maximum likelihood · Bayesian inference

J. Dkhar · S. Kumaria (✉) · P. Tandon
Plant Biotechnology Laboratory, Centre for Advanced Studies in Botany, North-Eastern Hill University, Shillong 793 022, India
e-mail: sumankhatrikumaria@hotmail.com

S. R. Rao
Department of Biotechnology and Bioinformatics, North-Eastern Hill University, Shillong 793 022, India

Introduction

Nymphaea L., an aquatic perennial herb with exceptionally beautiful flowers and floating leaves, is considered to be of evolutionary significance as it represents a group of early evolving flowering plants. It is the most diverse genus of the order Nymphaeales and is well represented globally. Conard's (1905) monograph of *Nymphaea* has classified the genus into five subgenera viz. *Nymphaea* subg. *Anecphya* (7–10 species), subg. *Brachyceras* (14–16 species), subg. *Hydrocallis* (14 species), subg. *Lotos* (3–4 species) and subg. *Nymphaea* (8 species) with each subgenus displaying a characteristic distribution. Ten species of *Nymphaea*, both wild (*N. alba*, *N. candida*, *N. nouchali*, *N. pubescens*, *N. rubra and N. tetragona*) and cultivated (*N. alba* var. *rubra*, *N. caerulea*, *N.* × *marliacea* and *N. micrantha*), and are reported from India (Mitra 1990), representing three subgenera, i.e., *Brachyceras*, *Lotos* and *Nymphaea*. *Nymphaea tetragona*, compared to the other ten species, is restricted in distribution and represented by only one population found at Umjapung Village, Nongkrem, East Khasi Hills District, Meghalaya, North-East India

(25°28′N–91°52′E). It is rare and endangered, and has been included for recovery programs (Ganeshaiah 2005). Interestingly, this species was thought to have been exterminated (Chauhan 1983).

The ITS regions have a definite role to play in the processing of nuclear ribosomal RNAs (nrRNAs). In vivo mutational studies in yeast (*Saccharomyces cerevisiae*) indicated that deletions in certain regions of ITS 1 inhibited production of mature small and large subunit rRNAs (Musters et al. 1990; van Nues et al. 1994), whereas certain deletions and point mutations in ITS 2 prevented or reduced processing of large subunit rRNA (van der Sande et al. 1992). These regions are relatively more conserved as compared to the intergenic spacer and external transcribed spacer regions of rDNA. Furthermore, sequence divergence of the ITS 1 and ITS 2 regions is more than that of nrDNA subunits (26S, 18S, and 5.8S) and is congenial to allowing phylogenetic reconstruction in a good number of genera (Baldwin et al. 1995). Baldwin (1992) initially exploited the phylogenetic utility of the ITS 1 and ITS 2 regions in the family Compositae. Since then, the ITS regions have been used for studying genetic variation at interspecific levels in several taxa, viz. *Nuphar* (Padgett et al. 1999), *Eupatorium* (Schmidt and Schilling 2000), *Inga* (Richardson et al. 2001), *Vigna* (Goel et al. 2002), *Gaertnera* (Malcomber 2002), *Ehrharta* (Verboom et al. 2003), robinoid legumes (Lavin et al. 2003), *Cicer* (Singh et al. 2007), *Avena* (Nikoloudakis et al. 2008), *Ferula* (Kurzyna-Młynik et al. 2008), *Symphyotrichum* (Vaezi and Brouillet 2009), *Ficus* (Baraket et al. 2009), *Citrus* (Kyndt et al. 2010) and many others. The use of the ITS region as a marker for resolving relationships among *Nymphaea* species has been reported (Borsch et al. 1998; Woods et al. 2005a; Liu et al. 2005). The ITS region in *Nymphaea* is reported to be extremely variable, and alignment throughout the genus was difficult. Woods et al. (2005a) could successfully segregate *N. odorata* subsp. *odorata* and *N. odorata* subsp. *tuberosa* into two separate taxonomic units. They also found the ITS region to be variable even at the subspecies level.

Based on the chloroplast *trn*T-*trn*F spacer sequences, Borsch et al. (2007) could resolve *Nymphaea* into three major lineages: subg. *Nymphaea* emerging as the first branch followed by subgg. *Brachyceras-Anecphya* and another clade comprising subg. *Hydrocallis* and subg. *Lotos*. However, relationships among members within these groups remain unclear. In addition to the molecular-based taxonomic revision of four wild Indian representatives of the genus *Nymphaea*, an attempt to resolve the phylogenetic relationship among members of *N.* subg. *Lotos* was conducted based on the ITS region, chloroplast *trn*K intron and *mat*K gene (Dkhar et al. 2010). The study indicated a close relationship between *N. petersiana* and *N. lotus*, while *N. pubescens* and *N. rubra* emerged as a separate clade. Moreover, molecular evidence suggests probable misidentification of *N. tetragona* and one specimen of *N. nouchali*. However, only one individual for each species was investigated. Explorations were re-conducted to increase taxon sampling and take into account the cultivated species of the genus *Nymphaea* found in India.

In the present study, by adopting a molecular phylogenetic approach utilizing nucleotide sequence data of the ITS region, we intend to evaluate the genetic divergence among Indian Nymphaeas and to reconstruct the phylogenetic tree of the genus *Nymphaea* in the light of both present data and the data available in the GenBank database (http://www.ncbi.nlm.nih.gov).

Materials and methods

Plant material and taxon sampling

The seven species of the genus *Nymphaea* examined in this study are *N. alba* var. *rubra*, *N. caerulea*, *N.* × *marliacea, N. nouchali*, *N. pubescens*, *N. rubra* and *N. tetragona*. Exploratory trips were conducted to gather samples of each *Nymphaea* species investigated from the states of Meghalaya and Assam (Northeast India). *Nymphaea* × *marliacea* was found to be growing together with *N. alba* var. *rubra* in one of the ponds found at Mukhla Village, Jaiñtia Hills District, Meghalaya (25°30′N–92°10′E). In case of *N. nouchali*, two sites were identified: one at Paikan, Goalpara District, Assam (26°02′N–90°38′E), and the other at Guwahati, Kamrup District, Assam (26°10′N–91°46′E). In Paikan, individuals with flower color variations (blue and white colored flower) were observed in the same pond. The plant at Guwahati was found to be morphologically distinct from the one found at Paikan. Plants of each *Nymphaea* species were identified at the Botanical Survey of India, Eastern Circle, Shillong, and specimen vouchers were prepared accordingly. The specimen vouchers were deposited at the Herbarium, Centre for Advanced Studies in Botany, North-Eastern Hill University, Shillong. Collected plants are maintained in an artificial pond of the Plant Biotechnology Laboratory, Centre for Advanced Studies in Botany, North-Eastern Hill University, Shillong, India.

Information on the seven species of the genus *Nymphaea* investigated here are summarized in Table 1, and the sequence data information of 18 additional species of the genus *Nymphaea* (*N. alba*, *N. amazonum*, *N. ampla*, *N. atrans*, *N. candida*, *N. capensis*, *N. elegans*, *N. elleniae*, *N. gigantea*, *N. gracilis*, *N. jamesoniana*, *N. lotus* f. *thermalis*, *N. mexicana*, *N. odorata*, *N. odorata odorata*, *N. odorata tuberosa*, *N. oxypetala* and *N. petersiana*) were

retrieved from the GenBank database. The sequence data of *N. caerulea*, *N. nouchali*, *N. rubra* and *N. tetragona* found in GenBank have been incorporated into the analysis. In addition to the above-mentioned *Nymphaea* species, five species representing four genera of Nymphaeaceae viz. *Barclaya*, *Euryale*, *Nuphar*, *Ondinea* and *Victoria* (in-groups) have been included in the study. Three species of genera *Brasenia* and *Cabomba* (Cabombaceae) have been chosen as outgroups. Nucleotide sequence data of these ingroup and outgroup species were taken from GenBank. The information pertaining to the taxa retrieved from GenBank are summarized in Table 2.

DNA extraction, amplification and sequencing

Fresh leaves collected from each plant were cleansed in running tap water prior to DNA extraction. Total genomic DNA was extracted from leaves of each species by following the method of Doyle and Doyle (1987) with the addition of the saturated phenol extraction step prior to ethanol precipitation. Polymerase chain reaction (PCR) was used to amplify the ITS 1 and ITS 2 regions of nrDNA in each of the species. The primers ITS 2 (5′ GCTGCGTTCTTCATCGATGC 3′) and ITS 5 (5′ GGA AGTAAAAGTCGTAACAAGG 3′) were used to amplify the ITS 1 region, and the primers ITS 3 (5′ GCATCG ATGAAGAACGCAGC 3′) and ITS 4 (5′ TCCTCCGCT TATTGATATGC 3′) were used to amplify the ITS 2 region. These primers were used in equal ratio. Amplifications were carried out in 100-µl PCR reactions, each containing 5 µl of sterilized glycerin (SRL, India), 10 µl of 10× Taq Buffer A containing 15 mM $MgCl_2$ (Bangalore Genei, India), 5 µl of 2 mM dNTPs (Bangalore Genei, India), 0.8 µl of *Taq* DNA Polymerase (3 U/µl) (Bangalore Genei, India), 20 µl of 10 mM ITS 4/ITS 2 primer, 20 µl of 10 mM ITS 3/ITS 5 primer (Operon Technologies, Germany), 4 µl of genomic DNA (30 ng/µl) and an appropriate volume of sterilized Millipore water. In some cases, the whole ITS region (ITS 1, 5.8S, ITS 2) was amplified using primers ITS 5 and ITS 4 mentioned above. DNA amplification was performed in an Applied Biosystems GeneAmp® PCR System 2700 programmed for 41 cycles (40 cycles for 1 min at 97°C, 1 min at 48°C and 45 s at 72°C, with an auto extension of 4 s at 72°C followed by an additional incubation for 7 min at 72°C for the 41st cycle). Amplified ITS 1 and ITS 2 (or the ITS region) were

Table 1 *Nymphaea* species with their respective subgenera, specimen vouchers, number of samples, place of collection and GenBank accession numbers of deposited sequences

Species	Subgenera	Specimen vouchers[a]	Samples	Place of collection	GenBank accession no.
N. nouchali Burm.f.[b]	*Brachyceras*	JD 06	2	Paikan, Goalpara District, Assam	EU191033, EU191038, FJ597740
N. nouchali Burm.f.[c]	*Brachyceras*	JD 07	2	Paikan, Goalpara District, Assam	FJ198404, FJ597741
N. nouchali Burm.f.	*Brachyceras*	JD 02	2	Guwahati, Kamrup District, Assam	FJ198405, FJ597742
N. caerulea Savigny	*Brachyceras*	JD 04	3	Shillong, East Khasi Hills District, Meghalaya	EU191032, EU191037, FJ597738, GQ468651
N. pubescens Willd.	*Lotos*	JD 09	1	Guwahati, Kamrup District, Assam	GQ468652
N. pubescens Willd.	*Lotos*	JD 08	2	Paikan, Goalpara District, Assam	FJ597743, GQ468653
N. rubra Roxb. Ex Andrews	*Lotos*	JD 10	2	Nongpoh, Ri-Bhoi District, Meghalaya	GQ468656, FJ597744
N. rubra Roxb. Ex Andrews	*Lotos*	JD 11	2	Chamata, Kamrup District, Assam	GQ468654, GQ468655
N. rubra Roxb. Ex Andrews	*Lotos*	–	2	Goalpara District, Assam	GU199476, GU199477
N. alba var. *rubra* Lönnroth	*Nymphaea*	JD 03	2	Shillong, East Khasi Hills District, Meghalaya	GQ477274, GQ358633, FJ597737, GQ358638
N. alba var. *rubra* Lönnroth	*Nymphaea*	JD 12	2	Smit, East Khasi Hills District, Meghalaya	GQ358634, GQ358635, GQ358639, GQ358640
N. alba var. *rubra* Lönnroth	*Nymphaea*	JD 13	2	Mukhla, Jaintia Hills District, Meghalaya	GQ358636, GQ358637, GQ358641, GQ358642
N. × *marliacea* Latour-Marliac	*Nymphaea*	JD 05	2	Mukhla, Jaintia Hills District, Meghalaya	GQ468657, FJ597739
N. tetragona Georgi	*Nymphaea*	JD 01	2	Nongkrem, East Khasi Hills District, Meghalaya	EU191035, EU191040, FJ597745

[a] Specimen vouchers deposited at the Herbarium, Department of Botany, North-Eastern Hill University, Shillong, India; [b] *N. nouchali* blue-colored flower; [c] *N. nouchali* white-colored flower

Table 2 Additional taxa used in the study retrieved from GenBank

Family	Genus	Species	GenBank accession No.	References
Nymphaeaceae	*Nymphaea*	*Nymphaea alba* Linn.	AJ012308	Unpublished
Nymphaeaceae	*Nymphaea*	*Nymphaea amazonum* Mart. & Zucc.	FM242149	Borsch et al. (2008)
Nymphaeaceae	*Nymphaea*	*Nymphaea ampla* (Salisbury) de Candolle	AY771812	Woods et al. (2005a)
Nymphaeaceae	*Nymphaea*	*Nymphaea atrans* S.W.L. Jacobs	FJ026555	Löhne et al. (2008)
Nymphaeaceae	*Nymphaea*	*Nymphaea caerulea* Savigny	AY707897	Liu et al. (2005)
Nymphaeaceae	*Nymphaea*	*Nymphaea candida* C. Presl	AY707900	Liu et al. (2005)
Nymphaeaceae	*Nymphaea*	*Nymphaea capensis* Thunb.	AY707898	Liu et al. (2005)
Nymphaeaceae	*Nymphaea*	*Nymphaea elegans* Hook.	AY771811	Woods et al. (2005a)
Nymphaeaceae	*Nymphaea*	*Nymphaea elleniae* S.W.L. Jacobs	FJ026560	Löhne et al. (2008)
Nymphaeaceae	*Nymphaea*	*Nymphaea gigantea* Hook.	FJ026566	Löhne et al. (2008)
Nymphaeaceae	*Nymphaea*	*Nymphaea gracilis* Zucc.	FM242151	Borsch et al. (2008)
Nymphaeaceae	*Nymphaea*	*Nymphaea jamesoniana* Planch.	FM242152	Borsch et al. (2008)
Nymphaeaceae	*Nymphaea*	*Nymphaea lotus* Linn. f. *thermalis* (DC.) Tuzson	FM242153	Borsch et al. (2008)
Nymphaeaceae	*Nymphaea*	*Nymphaea mexicana* Zucc.	AY771816	Woods et al. (2005a)
Nymphaeaceae	*Nymphaea*	*Nymphaea nouchali* Burm.f.	AY707896	Liu et al. (2005)
Nymphaeaceae	*Nymphaea*	*Nymphaea odorata* Aiton	AY771858	Woods et al. (2005a)
Nymphaeaceae	*Nymphaea*	*Nymphaea odorata odorata* Britton & Brown	AY771846	Woods et al. (2005a)
Nymphaeaceae	*Nymphaea*	*Nymphaea odorata tuberosa* (Paine) Wiersema & Hellquist	AY771853	Woods et al. (2005a)
Nymphaeaceae	*Nymphaea*	*Nymphaea oxypetala* Planch.	FM242150	Borsch et al. (2008)
Nymphaeaceae	*Nymphaea*	*Nymphaea petersiana* Klotzsch	FM242156	Borsch et al. (2008)
Nymphaeaceae	*Nymphaea*	*Nymphaea rubra* Roxb. Ex Andrews	AY707902	Liu et al. (2005)
Nymphaeaceae	*Nymphaea*	*Nymphaea tetragona* Georgi	AY707899	Liu et al. (2005)
Nymphaeaceae	*Nymphaea*	*Nymphaea tetragona* Georgi	EU428057	Volkova et al. (2010)
Nymphaeaceae	*Nymphaea*	*Nymphaea tetragona* Georgi	EU428053	Volkova et al. (2010)
Nymphaeaceae	*Barclaya*	*Barclaya longifolia* Wall.	FM242140	Borsch et al. (2008)
Nymphaeaceae	*Euryale*	*Euryale ferox* Salisb.	AY858637	Unpublished
Nymphaeaceae	*Nuphar*	*Nuphar japonica* DC.	AF067595	Padgett et al. (1999)
Nymphaeaceae	*Nuphar*	*Nuphar variegata* Durand	AF067578	Padgett et al. (1999)
Nymphaeaceae	*Ondinea*	*Ondinea purpurea* Hartog	FJ026600	Löhne et al. (2008)
Nymphaeaceae	*Victoria*	*Victoria amazonica* (Poepp.) Sowerby	AY858642	Unpublished
Nymphaeaceae	*Victoria*	*Victoria cruziana* Orbigny	AY858643	Unpublished
Cabombaceae	*Brasenia*	*Brasenia schreberi* Gmel.	EF526394	Unpublished
Cabombaceae	*Cabomba*	*Cabomba caroliniana* A. Gray	AY620424	Liu et al. (2005)
Cabombaceae	*Cabomba*	*Cabomba furcata* Schult & Schult.f.	AY620425	Liu et al. (2005)

fractionated by gel electrophoresis on 3% standard agarose gels in 1× TAE buffer and stained with ethidium bromide (Fig. 1). The eluted fragment was purified using a QIA-Quick gel extraction kit (Qiagen, Germany) and sequenced at Bangalore Genei Pvt Ltd., India, and Axygen Pvt Ltd., India.

Sequence alignment and indel coding

The published ITS 1 and ITS 2 sequences of *Nymphaea* (Woods et al. 2005a) were used to determine the boundaries of these regions in the species presently investigated. The ITS 1 and ITS 2 sequences, thus obtained, were subjected to multiple sequence alignments using Clustal X program (Thompson et al. 1997) available at ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/. A default setting with a fixed gap penalty of 15, a floating gap penalty of 6.66 and DNA transition weight of 0.5 in the multiple alignment parameter option were selected for alignment using the Clustal X program. In total, four multiple sequence alignments were performed: ITS 1 and ITS 2 multiple sequence alignments of six *Nymphaea* species (generated during this study; ITS of *N. alba* var. *rubra* are partially sequenced and were not included) for sequence characteristics and phylogenetic analysis, and multiple sequence alignments of ITS 1 and ITS 2 of all the data sets (present data plus data available in GenBank) for phylogenetic reconstruction of the genus *Nymphaea*. This alignment was further examined

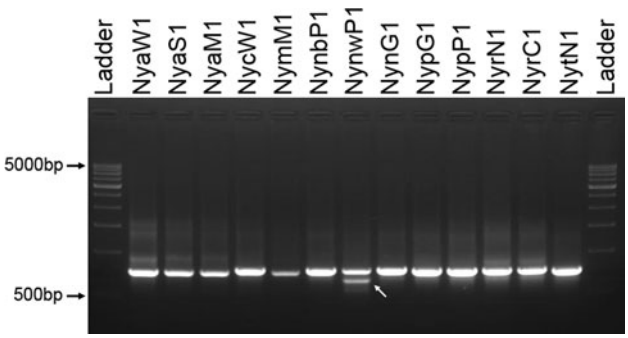

Fig. 1 Amplified ITS region of seven Indian representatives of the genus *Nymphaea*. DNA sample numbers of random individuals collected from different populations of each species are indicated above each lane. Amplified ITS region of the fungal endophyte, *Choanephora conjuncta*, in *N. nouchali* JD 07 is indicated by a *white arrow*

and modified manually applying the 'staggering alignment' rule as suggested by Barta (1997) and Morrison (2006). Morrison (2006) proposes that non-homologous residues should not be aligned to any other residues, unless there is evidence to do so. If they lack homology, they should be staggered (Barta 1997; in case of insertions). For highly divergent sequences, as is the case with the ITS region of *Nymphaea*, staggering will produce gaps different from those generated by insertions or deletions. To test the effectiveness of such an alignment procedure, we have either included or excluded gaps from the analysis. Gaps were coded in a binary matrix using SeqState v 1.21 (Müller 2005) applying the simple indel coding strategy (Simmons and Ochoterena 2000).

Phylogenetic analyses

Prior to phylogenetic analysis, sequence characteristics (sequence divergence, G + C content, number of transitions and transversions) of ITS 1 and ITS 2 were calculated for the six *Nymphaea* species presently studied using MEGA version 4 (Tamura et al. 2007; http://www.megasoftware.net). The multiple sequence alignment matrixes of ITS 1 and ITS 2 were incorporated into MEGA version 4 and analyzed separately.

Neighbor-joining (NJ), maximum parsimony (MP) and maximum likelihood (ML) methods were used to analyze the aligned sequence data matrix (ITS 1 + ITS 2). All trees were constructed using PHYLIP ver 3.69 (Felsenstein 2004) accessible at http://evolution.gs.washington.edu/phylip.html. Bootstrap analysis was carried out with 999 random seeds and 1,000 replicates (Felsentein 1985). For NJ, the distance matrix was estimated following the Kimura 2-parameter model (Kimura 1980). Rate variation among sites for all data was estimated using Diverge 2.0 (Gu 2001).

Bayesian inference (BI) of phylogeny was also conducted for the combined data set (ITS 1 + ITS 2) using MrBayes v.3.1.2 (Ronquist and Huelsenbeck 2003). BI analyses were performed for 1,000,000 generations applying the default settings (MCMC, two runs with four chains each, heating temperature 0.2, saving one tree every 100 generations). The best model of molecular evolution for the combined dataset was determined using jModelTest 0.1 (Posada 2008). The GTR model of molecular evolution with gamma-distributed rate variation across sites was assigned to the concatenated data. The binary (restriction site) model was applied to the indel partition. All trees created by Phylip and MRBAYES were viewed with the program Tree View 1.5 (Page 1996) available online at http://taxonomy.zoology.gla.ac.uk/rod/rod.html.

Outgroup selection

The *Nymphaea* species included in this study comprises groups that represent all subgenera within the genus Nymphaea, viz. subg. *Anecphya*, subg. *Brachyceras*, subg. *Hydrocallis*, subg. *Lotos* and subg. *Nymphaea.* Recent molecular analysis (Löhne et al. 2007) could resolve *Victoria* and *Euryale* as forming a sister clade to *Nymphaea* subgenera *Hydrocallis* and *Lotos*. Furthermore, *Ondinea purpurea* is nested within *N.* subg. *Anecphya* (Borsch et al. 2007; Löhne et al. 2008). In view of these recent findings, we included *Victoria*, *Euryale* and *Ondinea* in our data sets, and chose *Cabomba* and *Brasenia* as outgroups. Borsch et al. (2007) had suggested the two genera plus *Nuphar* as possible outgroups for *Nymphaea.*

Results

Sequencing signals in ITS

Direct sequencing of the ITS region of *N. alba* var. *rubra*, *N.* × *marliacea*, *N. pubescens* and *N. rubra* revealed additional signals in their respective chromatograms, hinting at the presence of divergent ITS paralogues (Fig. 2). Such deviating paralogues of the ITS region have been reported in several species of *Nymphaea* (Löhne et al. 2008).

Sequence length and base composition of six *Nymphaea* species

The ITS 1 and ITS 2 regions of the six *Nymphaea* species presently investigated showed variable sequence lengths and G + C content (%). The sequence lengths of ITS 1 for all the six species ranged from 276 to 283 bp, while ITS 2 sequence lengths ranged from 231 to 258 bp (Table 3).

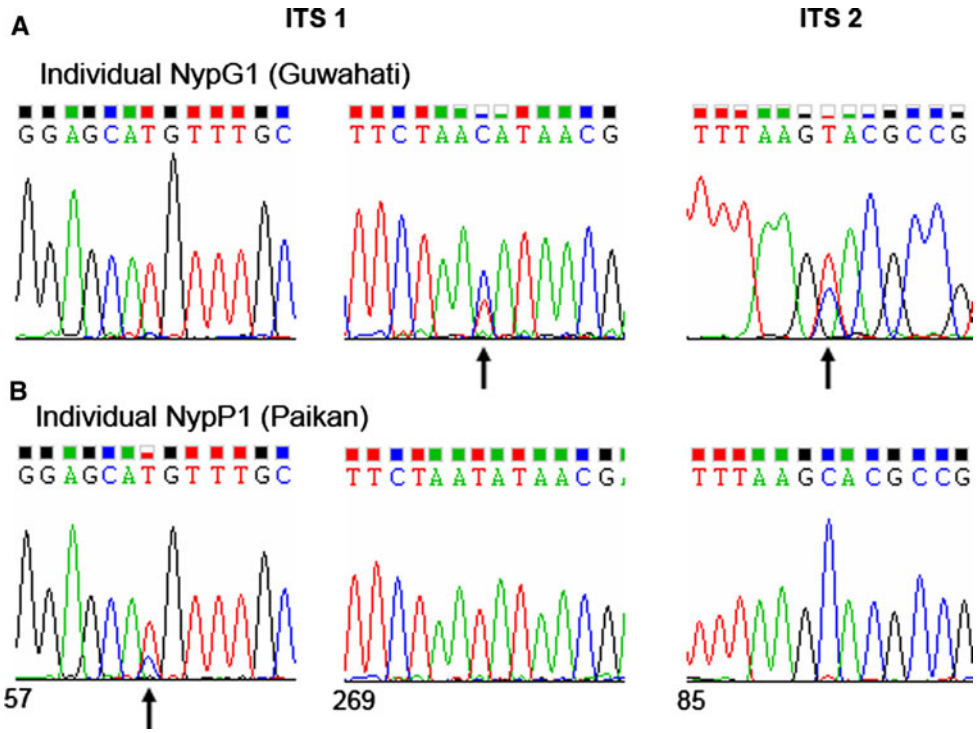

Fig. 2 Sequencing signals of portions of the ITS 1 and ITS 2 regions in two individuals of *Nymphaea pubescens*. *Arrows* indicate sites with additional signals. Numbers at the bottom indicate nucleotide sites

However, the G + C content (%) of ITS 2 was found to be slightly higher as compared to the ITS 1 region. The ITS 2 region recorded an average G + C content of 50.5%, and an average G + C content of 46.16% was recorded for ITS 1.

Sequence alignment

For determining sequence statistics, 299 and 269 characters were aligned for ITS 1 and ITS 2, respectively. The addition of *Nuphar advena*, selected as outgroup for phylogenetic relationship among Indian species of the genus *Nymphaea*, resulted in an aligned length of 279 and 264 characters for ITS 1 and ITS 2, respectively. The incorporation of sequence data retrieved from GenBank, aimed at reconstructing the phylogenetic tree of the genus *Nymphaea*, produced 562 aligned characters for ITS 1 and 369 aligned characters for ITS 2. The use of the 'staggering alignment' rule resulted in such a high number of aligned characters. Despite the high sequence variability, three conserved motifs were identified within the ITS 1 and ITS 2 regions (Fig. 3). The conserved motifs are common throughout the order Nymphaeales. Portions of the alignment matrix (concatenated ITS 1 and ITS 2) are illustrated in Fig. 4, depicting nucleotide base substitutions that infer the groupings of respective clades. The multiple sequence alignments for phylogenetic tree reconstruction are available on request from authors.

Sequence divergence among Indian Nymphaeas

The ITS 2 region is more conserved as is evident from the number of conserved sites (140 out of 269, 52.04%) shown in Table 3. Subsequently, relatively higher sequence divergence is recorded for the ITS 1. Percentage of sequence divergence based on substitution and substitution plus indels is 44.15 and 57.19, respectively, for the ITS 1; 29.74 and 47.96% sequence divergence were recorded for the ITS 2. The number of synapomorphic sites for ITS 1 is 74 (24.75%), whereas ITS 2 recorded synapomorphic sites of 54 (20.07%) in number. Similarly, autapomorphic sites are more in number for ITS 1 (26, 8.70%) than ITS 2 (15, 5.58%). ITS 1 records the highest percentage (32.44%) of parsimony informative sites. Both regions recorded frequent transitions rather than transversions. The number of indels for ITS 1 and ITS 2 is 21 and 26, respectively.

In the investigated plants of *N. nouchali*, sequence analysis of ITS 1 and ITS 2 did not reveal any variations between *N. nouchali* JD 06 and *N. nouchali* JD 07. However, on comparison with *N. nouchali* JD 02, nucleotide base substitutions at position numbers 46 (T↔C), 136 (T↔C) and 249 (G↔A), and an insertion (CAA) at position number 187 were recorded for ITS 1; nucleotide variations at position numbers 47 (C↔G), 173 (A↔G) and an insertion (A) at position number 226 were detected for ITS 2. Interestingly, gel electrophoresis of the amplified ITS region of *N. nouchali* JD 07 resulted in two separate

Table 3 Sequence characteristics of the ITS 1 and ITS 2 regions of six Indian representatives of the genus *Nymphaea*

Taxa	Length (nt)	G + C (%)	Aligned length (nt)	Sequence divergence (%)		Number of							Ti	Tv
				Substitution	Substitution + indels	Deletions/ indels	VS (%)	CS (%)	SS (%)	AS (%)	PIS (%)	MSS (%)		
ITS 1			299	44.15	57.19		132 (44.15)	128 (42.81)	74 (24.75)	26 (8.70)	97 (32.44)	32 (10.70)	59	22
Nymphaea														
N. caerulea	283	48.4				09/21								
N. x *marliacea*	275	–				08/21								
N. nouchali JD 02	282	46.1				10/21								
N. nouchali JD 06	279	45.5				11/21								
N. nouchali JD 07	279	45.5				11/21								
N. pubescens	280	–				10/21								
N. rubra	280	–				10/21								
N. tetragona	276	45.3				07/21								
Nymphaea + *Nuphar*			279	39.43	72.76	/34	110 (39.43)	76 (27.24)	51 (18.28)	21 (7.53)	75 (26.88)	31 (11.11)	49	23
ITS 2														
Nymphaea			269	29.74	47.96		80 (29.74)	140 (52.04)	54 (20.07)	15 (5.58)	60 (22.30)	11 (4.09)	42	27
N. caerulea	257	50.2				06/26								
N. x *marliacea*	232	–				15/26								
N. nouchali JD 02	258	50.4				05/26								
N. nouchali JD 06	257	50.2				06/26								
N. nouchali JD 07	257	50.2				06/26								
N. pubescens	245	–				07/26								
N. rubra	245	–				07/26								
N. tetragona	231	51.5				16/26								
Nymphaea + *Nuphar*			264	32.20	65.15	/41	85 (32.20)	92 (34.85)	39 (14.77)	25 (9.47)	53 (20.08)	21 (7.95)	45	19

–: Exact G + C content (%) cannot be accounted for because of the presence of ambiguity codes in sequence data

VS variable sites, *CS* conserved sites, *SS* synapomorphic sites, *AS* Autapomorphic sites, *PIS* parsimony informative sites, *MSS* multiple substitution sites, *Ti* transitions, *Tv* transversions

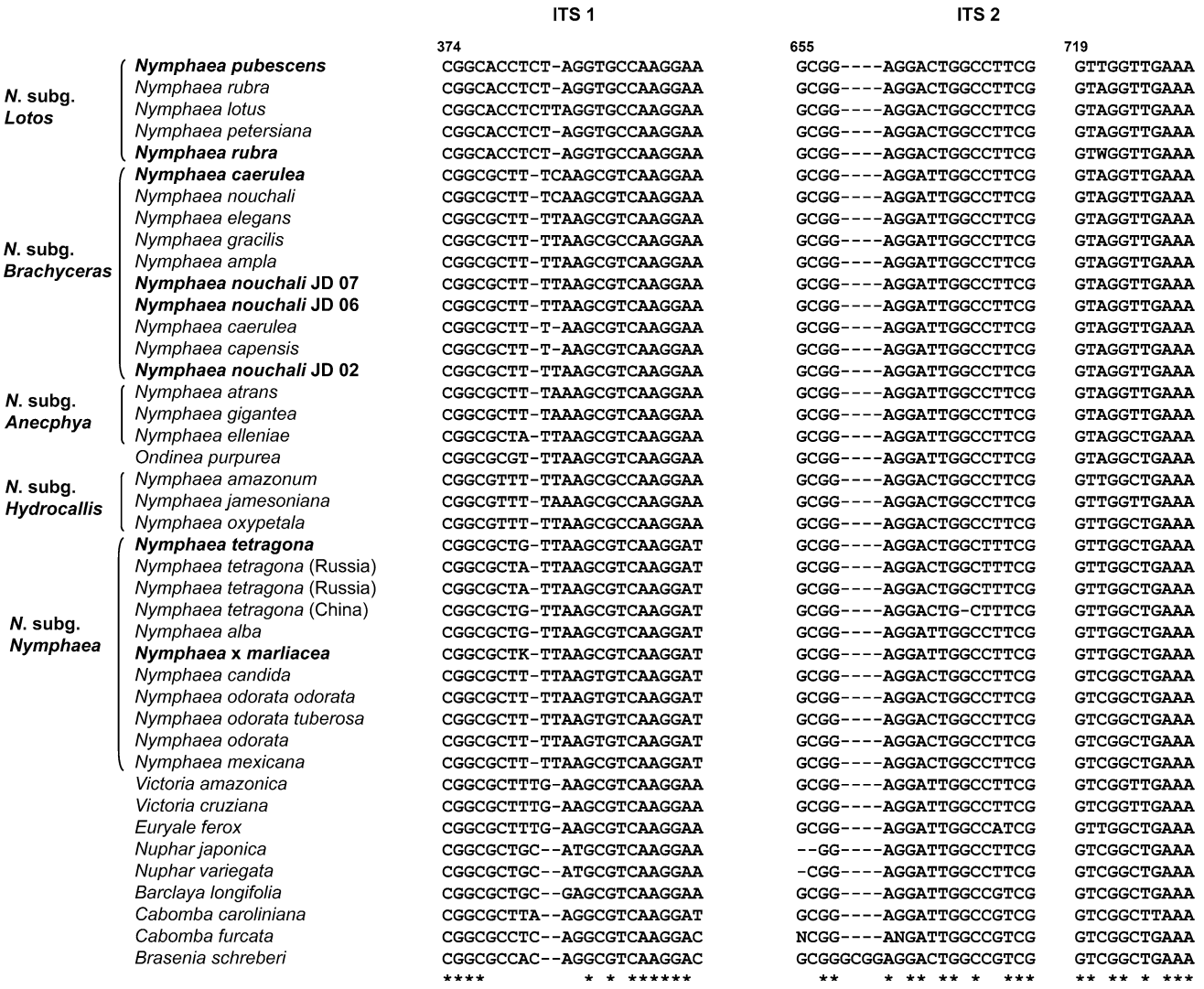

Fig. 3 Conserved motifs identified within the ITS1 and ITS2 region present throughout the order Nymphaeales. *Asterisks* represented below nucleotide base indicate identity, and *dashes* represent gaps. *Bolded* names of *Nymphaea* species represent some of the Indian representatives used in this study

bands (Fig. 1). Direct sequencing followed by nucleotide BLAST of the additional amplified fragment (small size) showed 99% sequence homology with the fungus *Choanephora conjuncta* (accession no. AJ278358), belonging to the Choanephoraceae family.

Phylogenetic relationship among Indian representatives

All four phylogenetic methods were used to analyze relationship among Indian representatives of the genus *Nymphaea* based on the concatenated ITS 1 and ITS 2 sequences (Fig. 5). Barring MP analysis, all other methods could resolve subg. *Nymphaea* as first branching followed by a clade comprised of subg. *Lotos* and subg. *Brachyceras*. No clear relationship among subgenera is indicated in the MP-based tree. Within subg. *Brachyceras*, all methods depict *N. nouchali* JD 02 as a separate clade. NJ and ML trees showed members of subg. *Lotos*, represented here by two species viz. *N. pubescens* and *N. rubra*, as separate clades, whereas MP and BI analysis collapsed both species into a single group.

Phylogenetic tree reconstruction of the genus *Nymphaea*

The NJ, MP, ML and BI methods were used to reconstruct the phylogenetic tree of the genus *Nymphaea* based on the combined nucleotide sequence data of ITS 1 and ITS 2. Barring ML analysis, all other phylogenetic methods support the placement of subg. *Nymphaea* at the base of the

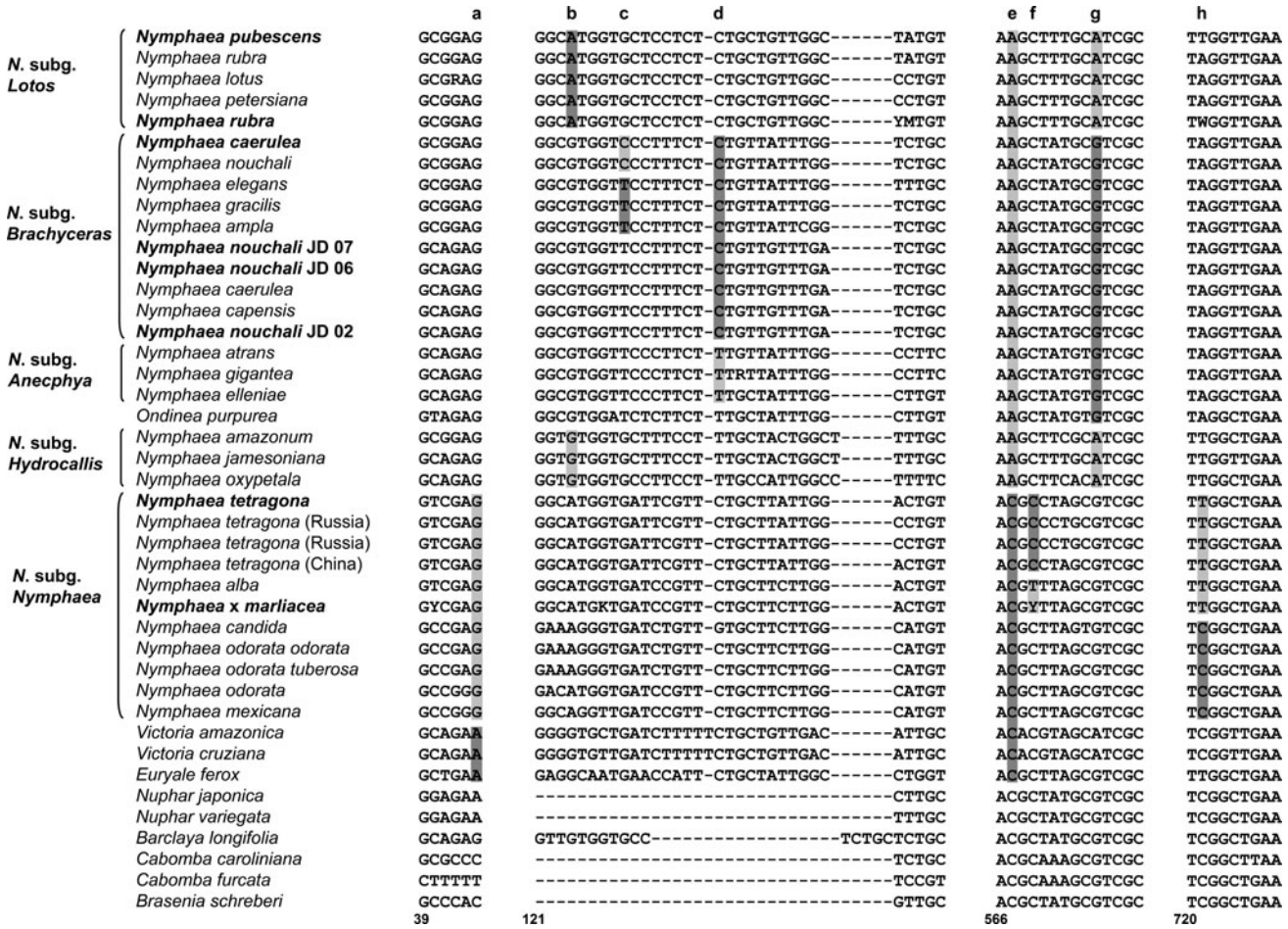

Fig. 4 Portions of the alignment matrix of ITS1 and ITS 2 showing nucleotide base substitutions that infer the grouping of respective clades (represented by *lower case letters* above the alignment matrix) in the Bayesian inference tree (Fig. 7). Numbers at the *bottom* indicate nucleotide sites. *Bolded* names of *Nymphaea* species represent some of the Indian representatives used in this study

tree, followed by a clade comprising subg. *Brachyceras* and subg. *Anecphya*, and another clade comprising subgg. *Hydrocallis* and *Lotos* (Figs. 6, 7). The association between the two clades are indicated but with varying support. In the ML-based tree (figure not shown), the relationship among major clades is ambiguous, but the association between subg. *Hydrocallis* and subg. *Lotos* is maintained. Furthermore, subg. *Brachyceras* and *Anecphya* emerged as sister groups. *Ondinea purpurea* showed a close relationship with members of subg. *Anecphya*, corroborating earlier reports (Borsch et al. 2007; Löhne et al. 2007, 2008). Interestingly, the position of the *Euryale-Victoria* clade is contradictory, with almost all methods favoring different placement within the phylogenetic tree.

All four methods could resolve members of subg. *Nymphaea* into two clades (>84% bootstrap value; 100% posterior probability): one comprising *N. mexicana*, *N. odorata* and *N. candida*, and two subspecies of *N. odorata*; and the other comprises *N. alba*, *N.* × *marliacea* and *N. tetragona*. *N. tetragona* found in India is closely related to the Chinese material rather than that found in Russia. Within subg. *Brachyceras*, *N. nouchali* JD 02 emerged as a separate clade showing close association to *N. capensis* and *N. caerulea* (GenBank). *N. caerulea* found in India formed a sister group with a clade comprising *N. ampla*, *N. elegans* and *N. gracilis*. Nucleotide sequence data of *N. nouchali* retrieved from GenBank are identical to those of the *N. caerulea*. *Nymphaea* subg. *Lotos* is represented here by all known species viz. *N. lotus*, *N. pubescens*, *N. rubra* and the recently added *N. petersiana*. All methods support the association between *N. petersiana* and *N. lotus*, as was reported in our previous study (Dkhar et al. 2010). However, increasing sampling size for *N. rubra* did not provide a clear relationship with respect to *N. pubescens*. The NJ and ML trees could resolve *N. pubescens* as a separate group, whereas MP and BI analyses grouped *N. pubescens* and *N. rubra* together. With the exception of subg. *Anecphya*, the inclusion of gaps as characters in BI analysis did not alter tree topology, but rather increased bootstrap support.

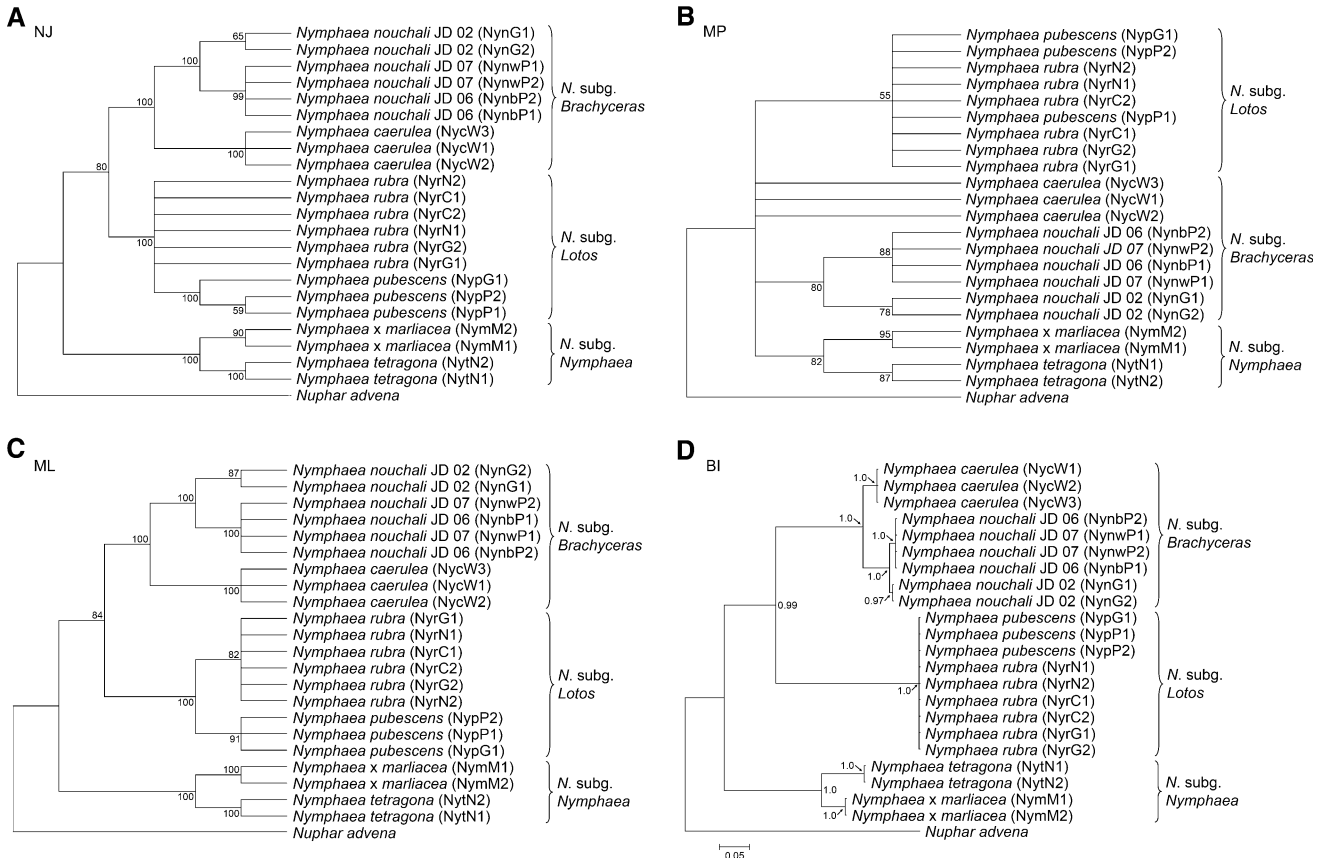

Fig. 5 Fifty percent majority-rule consensus phylogenetic trees for inferring relationships among Indian representatives of the genus *Nymphaea* obtained using four phylogenetic methods namely neighbor joining (NJ), maximum parsimony (MP), maximum likelihood (ML) and Bayesian inference (BI) of the concatenated ITS 1 and ITS 2 sequence data. MP tree is obtained from three equally parsimonious trees of 582 steps. BI tree is a consensus of 15,000 trees obtained from four runs of Bayesian analysis. Bootstrap values and posterior probabilities are given at the nodes. Branch lengths reflect number of changes per site. *Letters* followed by numerical within parentheses indicate DNA sample number

Discussion

Sequence characteristics of ITS 1 and ITS 2

The findings reported here identify ITS 1 and ITS 2 regions of *Nymphaea* as having high sequence divergence and G + C content (%). Although ITS 1 is longer than ITS 2 in sequence length (Table 3), the G + C content in ITS 2 (50.2–51.5%) is relatively higher compared to ITS 1 (45.5–48.4%). This is in conformity with an earlier report that identified angiosperm nrDNA ITS 2 sequences as GC rich (Hershkovitz and Zimmer 1996). A similar case was observed in *Nuphar* (belonging to Nymphaeaceae, Padgett et al. 1999) and *Hedysarum* (Chennaoui et al. 2007) where both ITS 1 and ITS 2 recorded more than 50% G + C content, with ITS 2 having a relatively higher G + C proportion. Remarkably, presumed functional ITS paralogs of *Cycas* species recorded 64.4 and 65.2% G + C content for ITS 1 and ITS 2, respectively (Xiao et al. 2010). Recent studies on fig cultivars (*Ficus carica* L.), however, revealed a relatively higher percentage of G + C content for ITS 1 (64.25%) as compared to ITS 2 (62.41%; Baraket et al. 2009). The higher level of sequence divergence in ITS 1 has been reported in a variety of plant species (Baldwin 1993; Moller and Cronk 1997; Kollipara et al. 1997). A similar case is also observed in this study where ITS 1 sequence divergence due to substitution and substitution plus indels is relatively much higher compared to ITS 2. However, this does not seem to be the case in *Nuphar*, where ITS 1 and ITS 2 of published sequences (Padgett et al. 1999) recorded low levels of ~3.87 and 7.05% sequence divergence, respectively. Low levels of sequence divergence in ITS 1 and ITS 2 were also recorded among species of *Cucumis* (Garcia-Mas et al. 2004) and *Brachycome* (Field et al. 2006).

Conserved motifs within the ITS 1 and ITS 2 region were identified in this study. The ITS 2 region is more conserved as compared to ITS 1. A characteristic conserved sequence GGCRY- (4 to 7n) -GYGYCAAGGAA (where Y = C or T; R = A or G) of ITS 1 presumed to be

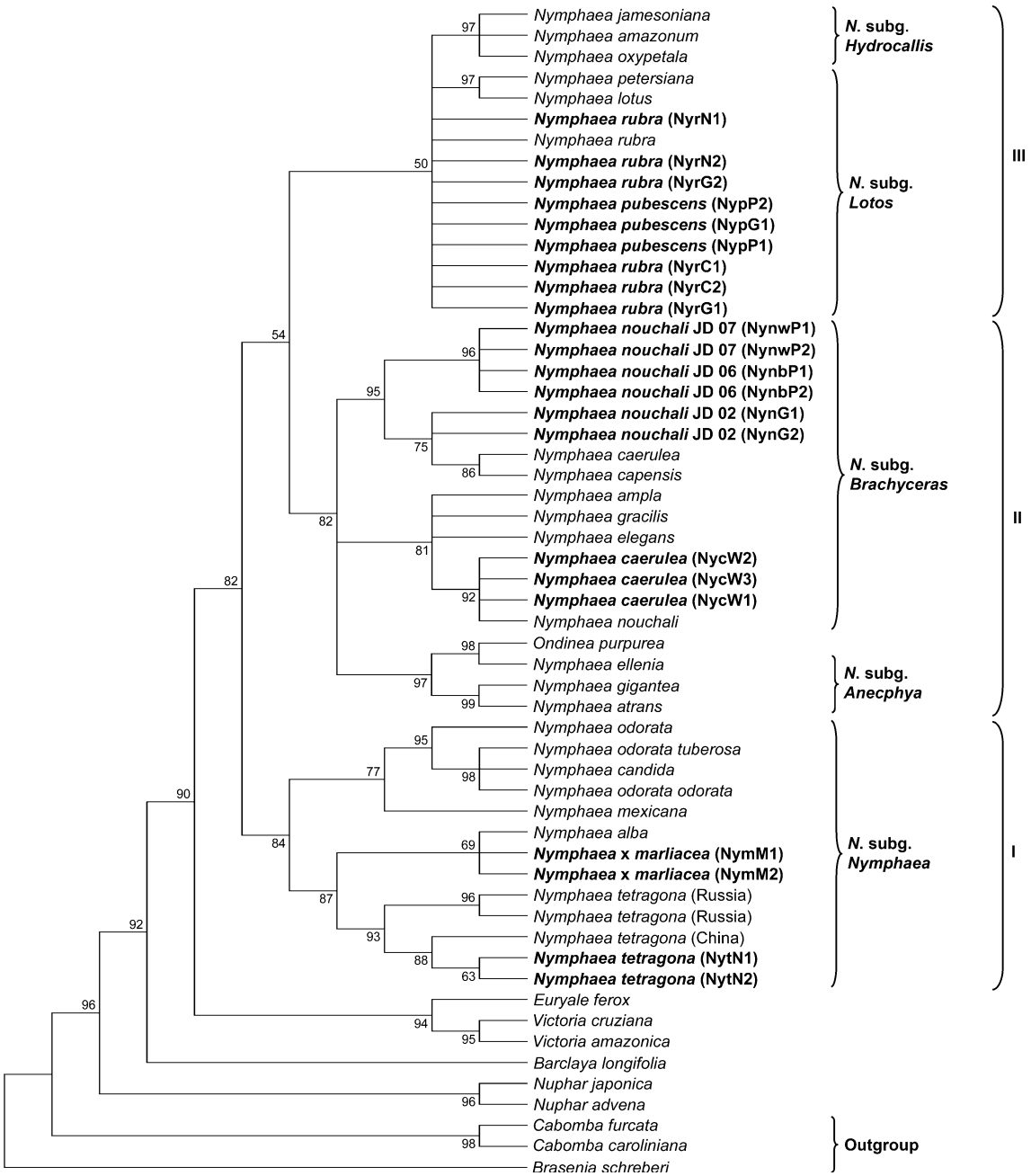

Fig. 6 Fifty percent majority-rule consensus of 1,152 trees (equally parsimonious of 2,223 steps) obtained from maximum parsimony analysis of concatenated ITS 1 and ITS 2 sequence data for inferring relationships among *Nymphaea* species. Bootstrap values are given at the nodes. *Letters* followed by numerical within parentheses indicate DNA sample number

a recognition site for processing of a primary transcript into the structural rRNA (Liu and Schardl 1994) was found in the ITS 1 region, common throughout the order Nymphaeales. The conserved motifs GCGGAGGAYTGGCYDTCG (where Y = C or T; D = A, T or G) and GTHGGYTKAAA (where H = C, T or A; Y = C or T; K = G or T) are located in the ITS 2 region (Fig. 3). The identification of conserved sequences within the ITS region is beneficial as accurate sequence alignments of the matrix are required for meaningful tree inferences.

Intraspecific variations

Morphological characteristics of *Nymphaea* have been studied by Moseley (1961), focusing on floral vasculature, and by Wiersema (1988), based on reproductive features. Recently, Volkova and Shipunov (2008) have analyzed morphological variation of *Nymphaea* confined to European Russia. These studies reported a substantial amount of morphological variation at the interspecific levels. The first report on intraspecific variations in *Nymphaea* was

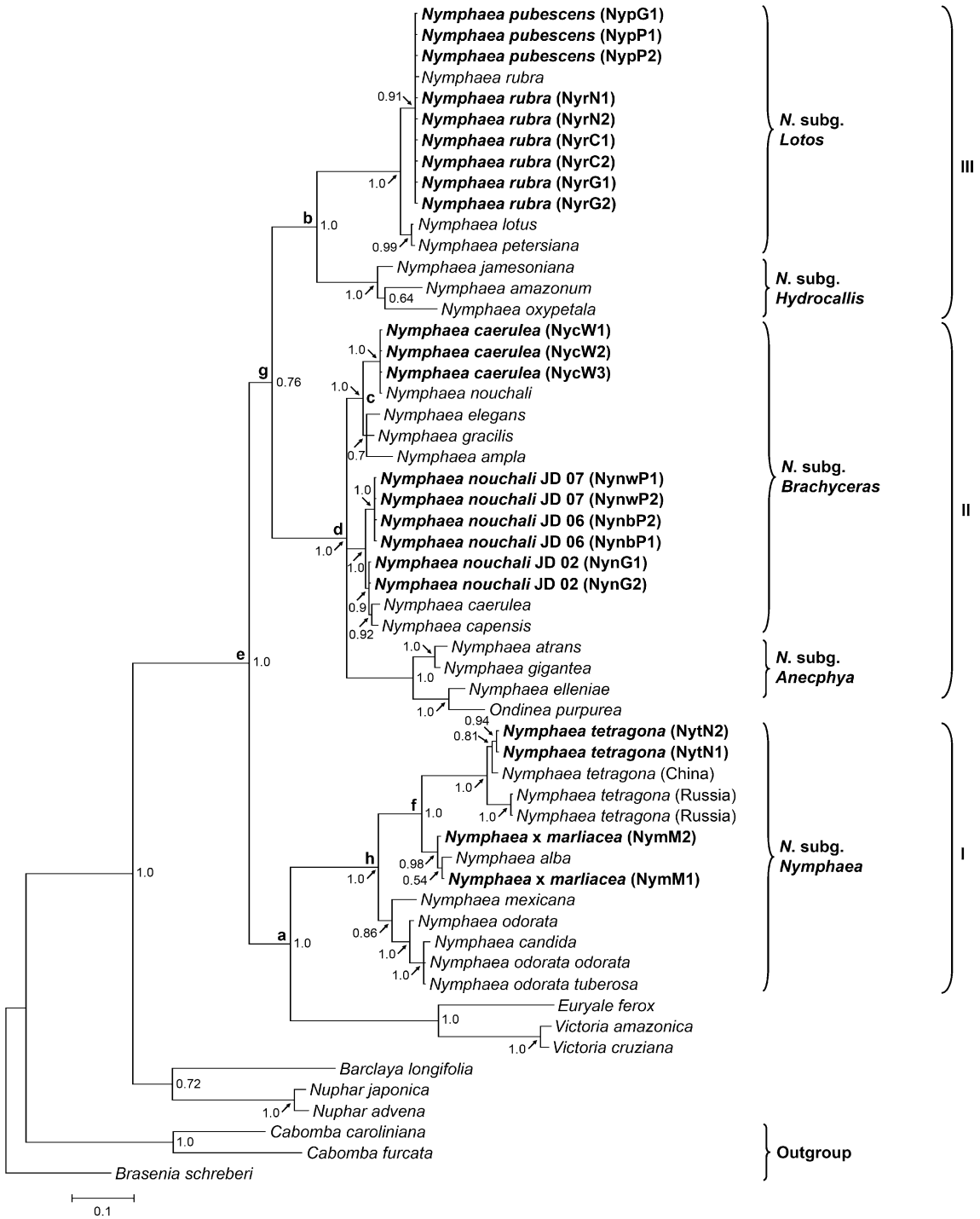

Fig. 7 Fifty percent majority-rule consensus of 14,937 trees obtained from four runs of Bayesian analysis of concatenated ITS 1 and ITS 2 data implementing the GTR + G model for inferring relationships among *Nymphaea* species. *Numbers* at nodes indicate posterior probabilities. *Branch lengths* represent the number of changes per site. DNA sample numbers are indicated within parentheses

conducted by Woods et al. (2005a, b). They studied non-molecular and molecular variations in *N. odorata* and complemented the data for segregation of *N. odorata* subsp. *odorata* and *N. odorata* subsp. *tuberosa* as separate entities. In our study, distinct morphotypes were identified for *N. nouchali*. In one particular location (Paikan, Goalpara District, Assam), individuals with variations in flower color (blue, *N. nouchali* JD 06 and white, *N. nouchali* JD 07) were identified. Another plant of *N. nouchali* JD 02 from Guwahati showed distinct variations in flower color,

number of petals and stamens. Sequence analysis of the ITS 1 and ITS 2 regions did not reveal any nucleotide variations between *N. nouchali* JD 06 and *N. nouchali* JD 07. In addition, chloroplast sequence data of the *trn*K intron, *mat*K and *rbc*L gene failed to show any difference between the two specimens. However, when compared to *N. nouchali* JD 02, nucleotide base substitutions and indels were detected, indicating distinctness between the two specimens.

Polymorphisms among ITS paralogues have been reported in several species of *Nymphaea* (Löhne et al. 2008). Such deviating paralogues may be due to random mutations occurring in some copies of the ITS region (Baldwin et al. 1995) or a case of recent hybridization and introgression. Additional signals detected in the sequencing chromatogram of *N. pubescens* may be due to random mutations occurring in both ITS 1 and ITS 2 (Fig. 2). However, additional signals detected in *N. alba* var. *rubra* and *N. rubra* could be due to recent hybridization and introgression. This assumption was confirmed by molecular cloning techniques, whereby distinct alleles were isolated from both species (Dkhar et al. 2011).

The addition of *N.* × *marliacea* JD 05 and *N. alba* var. *rubra*

In one of the exploratory trips to Jaiñtia Hills District, Meghalaya, we came across a population of water lilies having individuals with red- and yellow-colored flowers. At first sight, the two variants seemed to belong to the same plant with different flower colors. However, critical analyses of individuals with yellow-colored flowers revealed variations on the leaf (reddish-brown blotches), and that they became somewhat cup shaped and rose above the water when crowded, and had a lower number of carpels (11–14), characteristics of *N.* × *marliacea* (Mitra 1990). *Nymphaea* × *marliacea* (Chromatella) was the first hybrid produced by J. B Latour-Marliac through a cross between *N. alba* var. *rubra* (Swedish Red Waterlily) and *N. mexicana* (Latour-Marliac 1899). The lack of sequence data on *N. alba* var. *rubra* prevents the demonstration of a genetic contribution from both parental species. Nevertheless, phylogenetic analyses support the placement of *N.* × *marliacea* JD 05 as a close relative of *N. alba*. The red-flowered individuals resemble those found in Ward's Lake, Shillong, where Mitra (1990) would have had access to, and identify them as *N. alba* var. *rubra*. However, we found out that the Indian material identified as *N. alba* var. *rubra* is a hybrid originating from *N. alba* and *N. odorata*, with the latter species forming the maternal parent (Dkhar et al. 2011). It is unlikely that the plant would have originated through natural hybridization, because *N. alba* is not found in the northeastern part of India, and *N. odorata* is not available in India. So, the likely explanation for the origin of the Indian plant called *N. alba* var. *rubra* is through artificial hybridization, probably representing one of the hybrids produced by Latour-Marliac (1885–1890) that was subsequently introduced in Ward's Lake, Shillong, during its beautification by William Ward in 1892–1893 (Hussain 2005).

Interspecific relationships among *Nymphaea* species

The pioneering efforts of classification of *Nymphaea* by Conard (1905), which was verified by recent molecular studies of Borsch et al. (2007), into five subgenera has now been widely accepted. Although major lineages within *Nymphaea* are well resolved, the interspecific relationships within these groups remained uncertain. The placement of *N.* subg. *Nymphaea* at the base of the phylogenetic tree, followed by a clade comprising subg. *Brachyceras* and *Anecphya*, and another clade comprising subg. *Hydrocallis* and subg. *Lotos*, corroborates earlier reports (Borsch et al. 2007; Löhne et al. 2007; Borsch et al. 2008). Furthermore, the phylogenetic position of *Euryale-Victoria* clade within the genus *Nymphaea* (as is indicated by NJ tree), and as sister to all species of *Nymphaea* (as is indicated by MP tree), has been reported previously. However, the emergence of this clade as sister to *N.* subg. *Nymphaea* (both ML and BI trees) has never been reported. Based on the MP and BI analyses of the ITS region (Borsch et al. 2008), the *Euryale-Victoria* clade is nested within the genus *Nymphaea*, but with weak support. The difference in positioning of the *Euryale-Victoria* clade between the present study and that of Borsch et al. (2008) may be due to the difference in character numbers [547 characters of Borsch et al. (2008) against 931 characters in our study]. However, both trees of Borsch et al. (2008) were congruent, whereas our phylogenetic trees based on MP and BI showed dissimilarity. This could be due to the presence of a large number of gaps as a result of the use of the 'staggering alignment' rule. MP analysis using PHYLIP treats gaps as the fifth nucleotide state, counting one change whenever a gap arises (Felsenstein 2004). Other programs like Dnaml (ML), Dnadist (NJ) of PHYLIP and the software MrBayes consider gaps as missing data (Felsenstein 2004; Ronquist and Huelsenbeck 2003). As a result of such treatment, both ML and BI analysis generated identical tree topologies, whereas MP analysis produced an incongruent tree.

Based on the chloroplast *trn*T-*trn*F spacers, Borsch et al. (2007) could identify three lineages within *N.* subg. *Nymphaea*. However, our study supports the separation of subg. *Nymphaea* into two clades: *N. alba*, *N. tetragona* and *N.* × *marliacea* form one clade, whereas *N. mexicana*, *N. odorata* and *N. candida* form another. Among the

samples of *N. tetragona*, the Indian material is closely related to the plant found in China. This observation is supported by the morphological similarities shared between the two materials (eFloras 2008; Dkhar et al. 2010). Surprisingly, *N. candida* is grouped together with the two subspecies of *N. odorata*. Sequence analysis of the chloroplast *trn*T-*trn*F spacers failed to show any nucleotide variation between *N. alba* and *N. candida*, albeit a difference in morphology (Borsch et al. 2008). Recently, *N. candida* was reported to be an allopolyploid that originated through hybridization between *N. alba* and *N. tetragona* (Volkova et al. 2010), with the former species forming the maternal parent. But the ITS region of the allopolyploid was homogenized in the direction of *N. alba* (Volkova et al. 2010). These findings do not support the placement of *N. candida* as sister to *N. odorata* subsp., but rather indicated a probable misidentification of this taxon (GenBank accession no. AY707900).

NJ analysis could resolve members of subg. *Brachyceras* into two clades, but with weak bootstrap support (53%). Also, its divergence from subg. *Anecphya* is well supported with a bootstrap value of 100%. However, all other phylogenetic methods failed to provide clear branching between the two subgenera. The *trn*T-*trn*F-based study of Borsch et al. (2007) could resolve species of subg. *Brachyceras* into three clades: *N. gracilis*, *N. elegans* and *N. ampla* (New World species) form one clade; *N. micrantha*, *N. heudelotii* and *N. thermarum* form the second clade; and *N. caerulea* and *N. colorata* form another. The placement of the New World members as a separate clade is maintained in the present investigation, but formed a sister group with *N. caerulea* found in India. *N. nouchali* JD 02 is resolved in a different clade as compared to *N. nouchali* JD 06 and *N. nouchali* JD 07, showing close association with *N. capensis*. The present study corroborated the earlier findings that proposed *N. nouchali* JD 02 as a distinct species (Dkhar et al. 2010).

Nymphaea subg. *Lotos*, the smallest subgenus among the five subgenera, is represented here by all known species, viz. *N. lotus*, *N. pubescens*, *N. rubra* and the recently added *N. petersiana*. The morphology-based classifications of *N. petersiana* have placed this taxon as synonymy under *N. capensis* and *N. nouchali* of subg. *Brachyceras*, respectively (Conard 1905; Verdcourt 1989). Our previous study indicated no close association between *N. petersiana* and *N. nouchali* (Dkhar et al. 2010), but rather substantiated earlier findings that placed *N. petersiana* within subg. *Lotos*. *N. petersiana* emerged as a close relative of *N. lotus*, contrary to what was indicated in the *trn*T-*trn*F-based tree of Borsch et al. (2007). The additional signals detected in the sequencing chromatogram of the ITS region of *N. rubra* were replaced by ambiguity codes in the corresponding sequence data. On comparison with sequence data of other members of subg. *Lotos*, the additional signals occur at sites differentiating *N. pubescens* and *N. lotus*, suggesting that *N. rubra* may have originated through hybridization. Molecular cloning techniques confirmed this assumption, and the exact sequence matches of the chloroplast marker identify *N. pubescens* as the maternal parent (Dkhar et al. unpbl data). Lemmon et al. (2009) demonstrated the effect of ambiguous data ('?' or 'N') in phylogenetic analysis. They showed that such data can produce misleading estimates of topology and branch lengths. However, partially ambiguous characters, character taking the state of ambiguity codes, were not considered. The incorporation of such characters into the data produced contrasting topology when subjected to different phylogenetic methods. *N. pubescens* with respect to *N. rubra* emerged as a separate clade (NJ and ML trees), whereas MP and BI analysis grouped them together.

Although a substantial amount of morphological variation has been reported among *Nymphaea* species (Moseley 1961; Wiersema 1988; Volkova and Shipunov 2008), recent studies have focused mostly on molecular markers. Combining molecular and non-molecular characters could provide further insight into the existing phylogenetic relationships.

Acknowledgments The authors acknowledge two anonymous reviewers for their valuable comments made on an earlier version of the manuscript. The authors are grateful to Dr. S. Goel, Delhi University, for generous help and suggestions. JD is thankful to Dr. R. Gogoi, Botanical Survey of India, Shillong, for necessary help in procuring *N. nouchali* JD 02. This work is part of a project 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). JD is thankful to the University Grants Commission for awarding him the Rajiv Gandhi National Fellowship.

References

Baldwin BG (1992) Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. Mol Phylogenet Evol 1:3–16

Baldwin BG (1993) Molecular phylogenetics of *Calcydenia* (Compositae) based on ITS sequences of nuclear ribosomal DNA: chromosomal and morphological evolution reexamined. Am J Bot 80:222–238

Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ (1995) The ITS region of Nuclear Ribosomal DNA: a valuable source of evidence on Angiosperm Phylogeny. Ann Miss Bot Gard 82:247–277

Baraket G, Saddoud O, Chatti K, Mars M, Marrakchi M, Trifi M, Salhi-Hannachi A (2009) Sequence analysis of the internal transcribed spacers (ITSs) region of the nuclear ribosomal DNA (nrDNA) in fig cultivars (*Ficus carica* L.). Sci Hortic 120:34–40

Barta JR (1997) Investigating phylogenetic relationships within the Apicomplexa using sequence data: the search for homology. Methods 13:81–88

Borsch T, Hilu KW, Wilde V, Barthlott W (1998) Inferring the phylogeny of *Nymphaea*: evidence from the *trn*T-*trn*F region of the chloroplast genome. Am J Bot 85:116–117

Borsch T, Hilu KW, Wiersema JH, Löhne C, Barthlott W, Wilde V (2007) Phylogeny of *Nymphaea* (Nymphaeaceae): evidence from substitutions and microstructural changes in the chloroplast *trn*T-*trn*F region. Int J Plant Sci 168:639–671

Borsch T, Löhne C, Wiersema J (2008) Phylogeny and evolutionary patterns in Nymphaeales: integrating genes, genomes and morphology. Taxon 57:1052–1081

Chauhan AS (1983) Dwindling taxa of Meghalaya. In: Jain SK, Rao RR (eds) An assessment of threatened plants of India. Botanical Survey of India, Kolkata, pp 142–145

Chennaoui H, Marghali S, Marrakchi M, Trifi-Farah N (2007) Phylogenetic relationships in the North African genus *Hedysarum* as inferred from ITS sequences of nuclear ribosomal DNA. Genet Resour Crop Evol 54:389–397

Conard HS (1905) The waterlilies: a monograph of the genus *Nymphaea*. Lord Baltimore Press, Baltimore

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

Dkhar J, Kumaria S, Tandon P (2011) *Nymphaea alba* var. *rubra* is a hybrid of *N. alba* and *N. odorata* as evidenced by molecular analysis. Ann Bot Fennici 48:317–324

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

eFloras (2008) Published on the Internet http://www.efloras.org. Missouri Botanical Garden, St. Louis, MO and Harvard University Herbaria, Cambridge, MA. Accessed 16 March 2010

Felsenstein J (2004) PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle

Felsentein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Field BL, Houben A, Timmis JN, Leach CR (2006) Internal transcribed spacer sequence analyses indicate cytoevolutionary patterns within *Brachycome* Cass. (Asteraceae). Plant Syst Evol 259:39–51

Ganeshaiah KN (2005) Recovery of endangered and threatened species. Developing a national priority list of plants and insects. Curr Sci 89:599–600

Garcia-Mas J, Monforte AJ, Arús P (2004) Phylogenetic relationships among *Cucumis* species based on the ribosomal internal transcribed spacer sequence and microsatellite markers. Plant Syst Evol 248:191–203

Goel S, Raina SN, Ogihara Y (2002) Molecular evolution and phylogenetic implications of internal transcribed spacer sequences of nuclear ribosomal DNA in the *Phaseolus-Vigna* complex. Mol Phylogenet Evol 22:1–19

Gu X (2001) Maximum-likelihood approach for gene family evolution under functional divergence. Mol Biol Evol 18:453–464

Hershkovitz MA, Zimmer EA (1996) Conservation patterns in angiosperm rDNA ITS2 sequences. Nucleic Acids Res 24:2857–2867

Hussain I (2005) From residency to Raj Bhavan. A history of the Shillong Government House. Regency Publications, New Delhi

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

Kollipara KP, Singh RJ, Hymowitz T (1997) Phylogenetic and genomic relationship in the genus *Glycine* Wild. based on sequences from the ITS region rDNA. Genome 40:57–68

Kurzyna-Młynik R, Oskolski AA, Downie SR, Kopacz R, Wojewódzka A, Spalik K (2008) Phylogenetic position of the genus *Ferula* (Apiaceae) and its placement in tribe Scandiceae as inferred from nrDNA ITS sequence variation. Plant Syst Evol 274:47–66

Kyndt T, Dung TN, Goetghebeur P, Toan HT, Gheysen G (2010) Analysis of ITS of the rDNA to infer phylogenetic relationships among Vietnamese *Citrus* accessions. Genet Resour Crop Evol 57:183–192

Latour-Marliac JB (1899) Hardy hybrid water lilies. The Garden, March 18

Lavin M, Wojciechowski MF, Gasson P, Hughes C, Wheeler E (2003) Phylogeny of robinioid legumes (Fabaceae) revisited: *Coursetia* and *Gliricidia* recircumscribed, and a biogeographical appraisal of the Caribbean endemics. Syst Bot 28:387–409

Lemmon AR, Brown JM, Stanger-Hall K, Lemmon EM (2009) The effect of ambiguous data on phylogenetic estimates obtained by maximum likelihood and Bayesian inference. Syst Biol 58:130–145

Liu JS, Schardl CL (1994) A conserved sequences in internal transcribed spacer 1 of plant nuclear rRNA genes. Plant Mol Biol 26:775–778

Liu YL, Xu LM, Ni XM, Zhao J-R (2005) Phylogeny of the Nymphaeaceae inferred from ITS sequences. Acta Phytotax Sin 43:22–30

Löhne C, Borsch T, Wiersema JH (2007) Phylogenetic analysis of Nymphaeales using fast-evolving and noncoding chloroplast markers. Bot J Linn Soc 154:141–163

Löhne C, Borsch T, Jacobs SWL, Hellquist CB, Wiersema JH (2008) Nuclear and plastid DNA sequences reveal complex reticulate patterns in Australian water-lilies (*Nymphaea* subgenus *Anecphya,* Nymphaeaceae). Aust Syst Bot 21:229–250

Malcomber ST (2002) Phylogeny of *Gaertnera* Lam. (Rubiaceae) based on multiple DNA markers: evidence of a rapid radiation in a widespread, morphologically diverse genus. Evolution 56: 42–57

Mitra RL (1990) Nymphaeaceae. In: Nayar MP, Thothathri K, Sanjappa M (eds) Fascicles of flora of India, Fascicles 20. Botanical Survey of India, Kolkata, pp 11–25

Moller M, Cronk QCB (1997) Origin and relationships of *Saintpaulia* (Gesneriaceae) based on ribosomal DNA internal transcribed spacer (ITS) sequences. Am J Bot 84:956–965

Morrison DA (2006) Multiple sequence alignment for phylogenetic purposes. Aust Syst Bot 19:479–539

Moseley MF (1961) Morphological studies of the Nymphaeaeceae II. The flower of *Nymphaea*. Bot Gaz 112:233–259

Müller K (2005) SeqState: primer design and sequence statistics for phylogenetic DNA datasets. Appl Bioinformatics 4:65–69

Musters W, Boon K, van der Sande CAFM, Heerikhuizan HV, Planta RJ (1990) Functional analysis of transcribed spacers of yeast ribosomal DNA. EMBO J 9:3989–3996

Nikoloudakis N, Skaracis G, Katsiotis A (2008) Evolutionary insights inferred by molecular analysis of the ITS1–5.8S-ITS2 and IGS Avena sp. sequences. Mol Phylogenet Evol 46:102–115

Padgett DJ, Les DH, Crow GE (1999) Phylogenetic relationships in *Nuphar* (Nymphaeaceae): evidence from morphology, chloroplast DNA, and nuclear ribosomal DNA. Am J Bot 86: 1316–1324

Page RD (1996) TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12:357–358

Posada D (2008) jModelTest: phylogenetic model averaging. Mol Biol Evol 25:1253–1256

Richardson JE, Pennington RT, Pennington TD, Hollingsworth PM (2001) Rapid diversification of a species-rich genus of Neotropical rain forest trees. Science 293:2242–2245

Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574

Schmidt GJ, Schilling EE (2000) Phylogeny and biogeography of *Eupatorium* (Asteraceae: Eupatorieae) based on nuclear ITS sequence data. Am J Bot 87:716–726

Simmons MP, Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. Syst Biol 49:369–381

Singh A, Devarumath RM, RamaRao S, Singh VP, Raina SN (2007) Assessment of genetic diversity, and phylogenetic relationships based on ribosomal DNA repeat unit length variation and internal transcribed spacer (ITS) sequences in chickpea (*Cicer arietinum*) cultivars and its wild species. Genet Resour Crop Evol 54:810–825

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Thompson JD, Gibson TJ, Plewniak F, Mongin FJ, Higgins DG (1997) Clustal X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Vaezi J, Brouillet L (2009) Phylogenetic relationships among diploid species of *Symphyotrichum* (Asteraceae: Astereae) based on two nuclear markers, ITS and GAPDH. Mol Phylogenet Evol 51:540–553

van der Sande CAFM, Kwa M, van Nues RW, van Heerikhuizen H, Raué HA, Planta RJ (1992) Functional analysis of internal transcribed spacer 2 of *Saccharomyces cerevisae* ribosomal DNA. J Mol Biol 223:899–910

van Nues RW, Rientjes JMJ, van der Sande CAFM, Zerp SF, Sluiter C, Venema J, Planta RJ, Raué HA (1994) Separate structural elements within internal transcribed spacer 1 of *Saccharomyces cerviasiae* precursor ribosomal RNA direct the formation of 17S and 26S rRNA. Nucleic Acids Res 22:912–919

Verboom GA, Linder HP, Stock WD (2003) Phylogenetics of the grass genus *Ehrharta*: evidence for radiation in the summer-arid zone of the South African Cape. Evolution 57:1008–1021

Verdcourt B (1989) Nymphaeaceae. In: Polhill RM (ed) Flora of tropical East Africa. Balkema, Rotterdam, pp 1–12

Volkova PA, Shipunov AB (2008) Morphological variation of *Nymphaea* (Nymphaeaceae) in European Russia. Nord J Bot 25:329–338

Volkova PA, Pavel T, Christian B (2010) Evolutionary dynamics across discontinuous freshwater systems: rapid expansions and repeated allopolyploid origins in the Palearctic white water-lilies (*Nymphaea*). Taxon 59:483–494

Wiersema JH (1988) Reproductive biology of *Nymphaea* (Nymphaeaceae). Ann Missouri Bot Gard 75:795–804

Woods K, Hilu KW, Borsch T, Wiersema JH (2005a) Pattern of variation and systematics of *Nymphaea odorata*: II. Sequence information from ITS and *trn*L-*trn*F. Syst Bot 30:481–493

Woods K, Hilu KW, Borsch T, Wiersema JH (2005b) Pattern of variation and systematics of *Nymphaea odorata*: I. Evidence from morphology and inter-simple sequence repeats (ISSRs). Syst Bot 30:471–480

Xiao L-Q, Möller M, Zhu H (2010) High nrDNA ITS polymorphism in the ancient extant seed plant *Cycas*: incomplete concerted evolution and the origin of pseudogenes. Mol Phylogenet Evol 55:168–177