Telopea 13(1–2) 193–217

Towards a complete species tree of *Nymphaea*: shedding further light on subg. *Brachyceras* and its relationships to the Australian water-lilies

Thomas Borsch[1], Cornelia Löhne[1], Mame Samba Mbaye[2] and John Wiersema[3]

[1]*Botanischer Garten und Botanisches Museum Berlin-Dahlem und Institut für Biologie, Freie Universität Berlin, Königin Luise-Straße 6-8, 14195 Berlin, Germany*
[2]*Département de Biologie Végétale, Université Cheikh Anta Diop de Dakar, BP5005, Dakar-Fann, Sénégal*
[3]*Gambrills, MD 21054, USA*
Author for correspondence: t.borsch@bgbm.org

Abstract

The water-lily genus *Nymphaea* exhibits a worldwide distribution with an estimated number of more than 50 extant species. Recent phylogenetic analyses resolved three major lineages, a subg. *Brachyceras*–subg. *Anecphya* clade, also including *Nymphaea ondinea*, a subg. *Hydrocallis*-subg. *Lotos* clade, and the temperate subg. *Nymphaea* as a third clade. This study extends the taxon sampling for *Brachyceras*, previously the least understood subgenus. Maximum Parsimony and Bayesian analysis of nrITS sequence data depict a monophyletic subg. *Brachyceras*-clade and show a New World clade to be nested within African taxa. Plastid *trn*T-*trn*F sequence data are less conclusive. A middle Miocene origin is inferred for the New World *Brachyceras* lineage that must have dispersed out of Africa either via a Beringian migrational route or through immediate long distance dispersal. Within subg. *Brachyceras*, the West African individuals of *Nymphaea guineensis* form a distinct clade in both nuclear and plastid trees to which the Madagascan *Nymphaea minuta* is sister. Central and East African *Brachyceras* species appear well separated, suggesting a separating effect of the Dahomey gap to the evolution of these species. ITS sequences are more powerful in identifying *Nymphaea* species than *trn*T–*trn*F sequences. Nevertheless, about 15% of the known species remain to be sampled for a complete molecular tree of water-lilies. This also requires sampling of multiple populations in order to discover entities with a common evolutionary history and distinct molecular and morphological characters.

Introduction

Among all genera of the water-lily clade (Nymphaeales), *Nymphaea* is the most diverse lineage with more than 50 species. Recent molecular phylogenetic analyses have shown several lineages that represent species radiations on certain continents (Borsch et al. 2007, 2008; Löhne et al. 2007). A lineage comprising all temperate species in Eurasia and North America gains high support from both DNA and morphology, and

 ISSN0312-9764

corresponds to subg. *Nymphaea*. Two lineages with night bloomers appear vicariant for the New and Old World Tropics, respectively. The monophyletic subg. *Hydrocallis* occurs in Mexico, the Caribbean, Central and South America (Wiersema 1987, Löhne et al. 2008b) whereas subg. *Lotos* is native from Africa through the Indian subcontinent to northern Australia. Molecular studies have shown that the Southern African *N. petersiana*, previously classified within subg. *Brachyceras*, in fact belongs to subg. *Lotos* where it is sister to the remainder of the species (Borsch et al. 2007). All phylogenetic analyses hitherto carried out agree on the close relationship between the pantropically distributed subg. *Brachyceras* and an Australian radiation of species constituting subg. *Anecphya*. Complex reticulate evolutionary patterns were involved in the genesis of the approximately 10–16 species of *Anecphya* (Löhne et al. 2008a), many of which were described only recently (Jacobs & Porter 2007; Jacobs & Hellquist 2006; Jacobs & Hellquist 2011, this issue). Morphology and molecules agree on the finding that the monotypic genus *Ondinea* diverged rather recently from ancestors within *Anecphya* (Borsch et al. 2007, Löhne et al. 2009), pointing at rapid changes in floral architecture as well as organ number, probably as a consequence of changes in life form (submerged aquatic in temporal streams). Therefore, Löhne et al. (2009) transferred *Ondinea purpurea*, including its two subspecies, to *Nymphaea*.

Analyses of divergence times and historical biogeography of Nymphaeales (Yoo et al. 2005, Löhne et al. 2008b) revealed two major diversification phases during the evolution of the Nymphaeales crown group. First, there was a rapid differentiation into the three major lineages, Cabombaceae, *Nuphar* and the remaining Nymphaeaceae, during the Palaeocene. Secondly, the radiation of core Nymphaeaceae (*Victoria*, *Euryale*, *Nymphaea* incl. *Ondinea*) took place from the Late Oligocene to Middle Miocene. The scenario outlined by Löhne et al. (2008b) starts with a differentiation of the *Nymphaea* subgenera in the Northern Hemisphere and subsequent migration of subg. *Hydrocallis* to the New World and subg. *Lotos* to the Old World. During the Miocene, the subgenus *Hydrocallis* radiated in South America, whereas ancestors of subgenus *Anecphya* migrated to Australia followed by a rapid radiation. Löhne et al. (2008b) explain the present pattern of tropical disjunctions of closely related lineages as being a result of range expansion in the Northern hemisphere during the Early Tertiary and, from the Oligocene to Miocene, subsequent vicariance due to the formation of migration barriers (oceans and climatic zones). Similar scenarios have been inferred for several other pantropical angiosperm groups, such as Magnoliaceae (Azuma et al. 2001, Doyle et al. 2004), Lauraceae (Chanderbali et al. 2001) and Malphigiaceae (Davis et al. 2002, 2004).

The subgenus *Brachyceras* is the least understood group of water-lilies. Many taxa, especially from the Palaeotropics, have an unclear status and have not been collected for decades. A comprehensive modern treatment of the group is sorely needed to integrate the fragmentary regional information that currently exists. The *trn*T–*trn*F data set of Borsch et al. (2007) included eight *Brachyceras* species, which at that time comprised the largest number of taxa in any phylogenetic study involving the subgenus. There were two African and one New World subclades, although their relationships remained unclear. Evidence from multiple plastid regions, mitochondrial *matR*, nuclear ITS and morphology placed the Neotropical *N. gracilis* and the African *N. micrantha* as sisters (Löhne et al. 2007, Borsch et al. 2008). However, the monophyly of subg. *Brachyceras* itself could not be resolved with any confidence in the more densely sampled *trn*T–*trn*F tree of Borsch et al. (2007). Therefore, it remains to be clarified if the pan-tropical

water-lily species of *Brachyceras* are paraphyletic to subg. *Anecphya* or are in fact monophyletic. In order to test this, dense taxon sampling will be crucial as a random choice of few *Nymphaea* species is likely to lead to biased phylogenetic conclusions on the phylogenetic position of *Nymphaea* and on phylogenetic relationships within the genus (Löhne et al. 2007).

A complete species phylogeny of water-lilies will not only be essential for better understanding of their evolutionary diversification but also is a prerequisite to study species limits and to arrive at a complete modern species assessment. The strategy is to include all morphological entities (species and phenotypic variants) into plastid and nuclear sequence data sets. This allows a determination of all lineages and, through genomic incongruence, also of areas of the tree where hybridisation led to reticulate origin of species. Hybrid speciation in conjunction with allopolyploidisation is evident within *Nymphaea* subgenera *Anecphya* (Löhne et al. 2009) and *Nymphaea* (Volkova et al. 2010). Representation of individuals from throughout the range of species is necessary to detect geographical differentiation within monophyletic species, as evident in the widespread North American *Nymphaea odorata* (Woods et al. 2005), and incomplete lineage sorting. This may be a factor in the Australian subg. *Anecphya* (Löhne et al. 2008a). Reticulation or unequal lineage sorting, however, has not yet been detected within other water-lily lineages but this could simply be caused by limitations in sampling.

Here we use the plastid *trn*T–*trn*F and nuclear ITS regions to construct a comprehensive sequence data set with the aim of including as many taxa as possible rather than limiting the size of the taxon set in order to increase the number of characters. Insufficiently resolved clades or subclades within the respective trees can then be studied later with additional approaches including genomic fingerprinting techniques and algorithms not *a priori* relying on dichotomous evolutionary patterns. Aims of this study are therefore to include more samples of subg. *Brachyceras* and to test (1) how the various lineages of *Brachyceras* are related to each other and to the Australian *Anecphya* clade, and (2) the extent to which hitherto unsampled taxa of *Brachyceras* can be distinguished by DNA sequence data.

Material and Methods

Taxon sampling

The *trn*T–*trn*F dataset used in this study comprises 86 species of Nymphaeales, representing both genera of the Cabombaceae (*Brasenia*, *Cabomba*) as outgroup taxa, each genus of the Nymphaeaceae (*Barclaya*, *Euryale*, *Nuphar*, *Nymphaea*, *Victoria*), and within the genus *Nymphaea*, each of the five subgenera (*Anecphya*, *Brachyceras*, *Hydrocallis*, *Lotos*, *Nymphaea*). *Nymphaea ondinea*, previously separated from *Nymphaea* as a distinct genus, is also included in our data sets.

A subset of 43 of these taxa was employed for analysing the relationships of subg. *Brachyceras* and *Anecphya* using the nuclear ITS marker. Here, two species of subg. *Hydrocallis* (*N. amazonum*, *N. jamesoniana*) were used as outgroup taxa, with all available taxa of *Anecphya* and *Brachyceras* forming the ingroup. New samples were sequenced for both *trn*T–*trn*F and ITS in this study. Additionally, published sequences from the authors' previous studies were used to complement the data sets. All taxa used

in this study, including information on origin of the material, voucher specimens and EMBL/Genbank accession numbers, are listed in Appendix 1.

DNA isolation, amplification and sequencing

DNA was isolated from silica-gel-dried leaf tissue using the triple extraction method of Borsch et al. (2003). The chloroplast genomic region *trn*T–*trn*F is widely used for phylogenetic analyses and spans the intergenic spacer between *trn*T and *trn*L, the *trn*L gene including its intron as well as the spacer between *trn*L and *trn*F. We amplified this region in two fragments using the primers rps4–5R (Sauquet et al. 2003) and trnL110R (Borsch et al. 2003) for the 5'-fragment and the primers C and F (Taberlet et al. 1991) for the 3'-fragment. In addition to the amplification primers, the internal sequencing primers D and E (Taberlet et al. 1991) were used to read through long poly-TA stretches in the p8 stem loop of the *trn*L intron (see Borsch et al. 2007 for a detailed analysis of this region).The nuclear marker region ITS spans the internal transcribed spacer 1 (ITS1) between 18S and 5.8S rDNA, the 5.8S rDNA itself, and the internal transcribed spacer 2 (ITS2) between 5.8S and 26S rDNA. Amplification and sequencing of this region was done using the standard primers ITS4 and ITS5 (White et al. 1990) and following the procedure outlined in Löhne et al. (2008a).

Alignment and indel coding

Sequences were aligned manually with PhyDe® version 0.9.95 (Müller et al. 2007) following the rules outlined in Löhne and Borsch (2005). For ITS, sequences of *Amborella*, *Austrobaileya*, *Illicium*, *Schisandra* and *Kadsura* could not be aligned with the sequences of Nymphaeales (at least for the major parts of the region). Therefore, only ITS sequences of representatives of the subgenera *Brachyceras* and *Anecphya*, with two representatives of subg. *Hydrocallis* as outgroup, were aligned and analysed. Mutational hotspots (after Borsch et al. 2003) were excluded from analysis. All length mutations in ITS were coded automatically in a "01"-matrix with SeqState version 1.4 (Müller 2005), applying the "simple indel coding" strategy after Simmons and Ochoterena (2000). The indel matrix was appended to the sequence matrix for Maximum Parsimony and Bayesian analyses. Thus, the final ITS matrix contained 767 characters (707 nucleotide characters and 60 indels).

Alignment and indel coding of the *trn*T–*trn*F regions was done in the same way. Additionally, the p8 stem-loop region within the *trn*L intron (originally excluded as a hotspot from the main matrix, see also Borsch et al. 2007) was aligned separately for each subgenus and appended to the matrix. This allowed the inclusion of nucleotide positions and length mutations, which are potentially informative within the subgenera. The final *trn*T–*trn*F matrix comprised 2077 characters (1953 nucleotide positions from *trn*T–*trn*F, including the p8 region, and 124 indels).

Phylogenetic and dating analyses

The two data sets of *Nymphaea s.l.* were analysed separately. Maximum parsimony (MP) ratchet analyses were conducted with command files generated by PRAP (Müller 2004) and then executed with PAUP* version 4.0b10 (Swofford 2002). The ratchet settings were 200 iterations, weight 2, weighted = 25%, and 10 random addition cycles. Heuristic search parameters were simple stepwise addition, saving only one of the shortest trees per random-addition cycle and increasing the maximum number of trees

automatically by 100. Node support was estimated through jackknifing (JK) 10,000 replicates (simple addition, keeping 1 tree per replicate, deleting 36.8% of characters in each replicate).

For Bayesian analysis, models of molecular evolution in ITS and the *trn*T–*trn*F region, respectively, were determined using MrModeltest version 2.2 (Nylander 2004) according to the Akaike information criterion. The following models were selected: GTR+G for ITS and the *trn*L intron, GTR+I for the *trn*L–*trn*F spacer, GTR for the *trn*T–*trn*L spacer, JC for the *trn*L gene, and F81 for the p8 region within the *trn*L intron. Bayesian analyses of the total evidence data sets (substitutions plus indels) were performed using MrBayes version 3.1.2 (Ronquist & Huelsenbeck 2003), with the binary model applied to the indel partition. Analyses settings were as follows: four runs with four chains and 1,000,000 generations each, saving one tree per 100 generations. During the calculation, tree probabilities converged to a stable value after 38,000 generations for ITS and after 15,000 generations for *trn*T-*trn*F; thus, the burn in was set to 380 and 150, respectively.

Divergence times were inferred using the integrated relaxed clock approach implemented in the program BEAST version 1.5.4 (Drummond & Rambaut 2007). Instead of rate priors, two calibration points were used for each data set. For *trn*T–*trn*F the minimum age for the *Nuphar* crown group was set as 52 ± 1 Mya (Chen et al. 2004; *Nuphar wutuensis* is the oldest known fossil record of *Nuphar*). The *Nymphaea* crown group was set to have a minimum age of 33 ± 1 Mya. This is based on *Nymphaea liminis* as the oldest fossil seeds that can be unambiguously assigned to *Nymphaea* (dated as Upper Eocene/Lower Oligocene; Collinson 1980). The time of the Eocene/Oligocene boundary is drawn from Berggren et al. (1995). This set of taxa assigned to *Nymphaea* was not specified as monophyletic for the calculations. For the ITS data set, minimum ages for the crown group of the *Brachyceras-Anecphya*-clade were set as 24.7 ± 5.8 Mya and for the *Anecphya* clade as 19.6 ± 6.5 Mya based on the dated *trn*T–*trn*F tree of this study (Fig. 2). Models of molecular evolution were chosen for each partition as depicted by MrModeltest (see above). However, since the programme BEAST does not include a model for binary matrices we excluded the indel characters from both data sets. Thus, the matrices containing only nucleotide characters were analysed with BEAST. Rates for each branch were drawn independently from a lognormal distribution (Drummond et al. 2006). A Yule speciation model was assumed and a random starting tree was used. Analyses were performed for 10,000,000 generations, saving one tree every 1000 generations. The burn in was set to 10% yielding 9000 trees. The maximum clade credibility tree was calculated with TreeAnnotator v. 1.4.8 (Drummond & Rambaut 2007), setting the posterior probability limit to 0.7. Trees were rooted with Cabombaceae in the *trn*T–*trn*F analysis and with *Nymphaea amazonum* and *N. jamesoniana* in the ITS analysis.

Results

Trees obtained from *trn*T–*trn*F

Maximum Parsimony (MP) analysis of the chloroplast marker *trn*T–*trn*F yielded 114 shortest trees with a tree length of 543 steps. Figure 1 shows the strict consensus of these 114 trees with Jackknife support values above branches. The consensus tree of the Bayesian Analysis using MrBayes (BA) is not shown since it is largely congruent

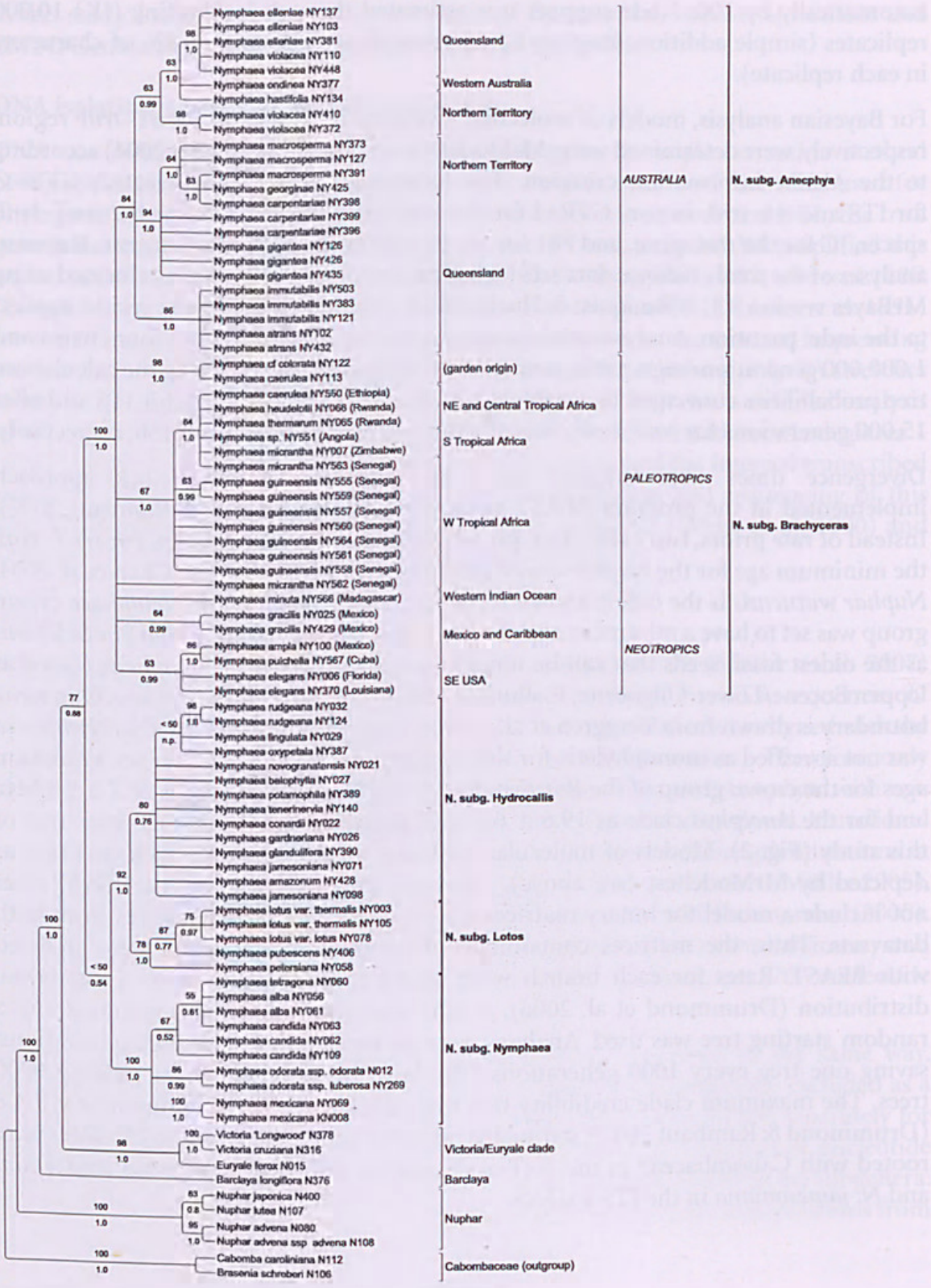

Fig. 1. Phylogenetic relationships within *Nymphaea* and Nymphaeaceae as inferred from the plastid marker *trn*T–*trn*F. The tree shown is the strict consensus of 114 shortest trees obtained with Maximum Parsimony analysis. Jackknife values are given above branches. Posterior probabilities as obtained from Bayesian analysis with MrBayes (tree not shown) are given below branches. Nodes, which are not present in the Bayesian tree, are marked with *n.p.* (not present).

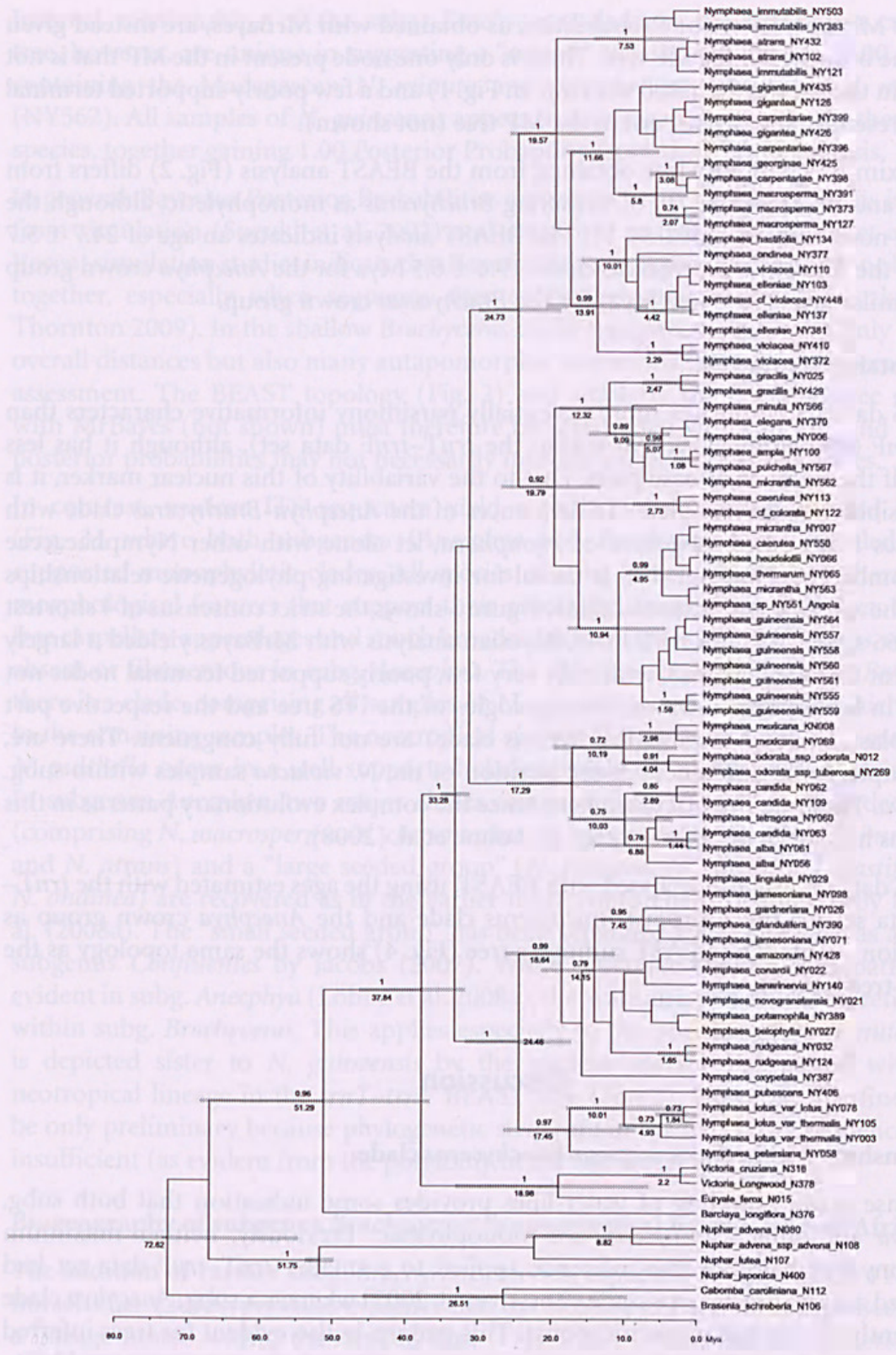

Fig. 2. Summary tree from BEAST analysis of the plastid marker *trn*T–*trn*F. Log-normal prior of 52.0 ± 1 million years was set for the most recent common ancestor (MRCA) of *Nuphar* and 33.0 ± 1 million years for the MRCA of *Nymphaea*, according to the fossil record (see text). Posterior probability values above 0.7 are presented above each node. The inferred age of each node is given below the respective branch, while the bars represent 95% confidence intervals of the inferred ages.

with the MP tree. Posterior Probabilities, as obtained with MrBayes, are instead given below the branches of the MP tree. There is only one node present in the MP that is not present in the BA tree (marked with *n.p.* in Fig. 1) and a few poorly supported terminal nodes present in the BA but not in the MP tree (not shown).

The maximum credibility tree obtained from the BEAST analysis (Fig. 2) differs from the MP and BA trees (Fig. 1) by depicting *Brachyceras* as monophyletic, although the relevant node receives only 0.92 PP. The BEAST analysis indicates an age of 24.7 ± 5.7 Mya for the *Anecphya-Brachyceras* clade, 19.6 ± 6.5 Mya for the *Anecphya* crown group and a similar age (19.8 ± 6.7 Mya) for the *Brachyceras* crown group.

Trees obtained from ITS

The ITS data set comprises more potentially parsimony informative characters than *trn*T–*trn*F (250 within ITS, 190 within the *trn*T–*trn*F data set), although it has less than half the number of base pairs. Due to the variability of this nuclear marker, it is not possible to align complete ITS sequences of the *Anecphya-Brachyceras* clade with sequences from other subgenera of *Nymphaea*, let alone with other Nymphaeaceae or Cabombaceae. However, ITS is useful for investigating phylogenetic relationships within the *Anecphya-Brachyceras* clade. Figure 3 shows the strict consensus of 4 shortest trees (466 steps). As with *trn*T–*trn*F, Bayesian analysis with MrBayes yielded a largely congruent tree (not shown), with only very few, poorly supported terminal nodes not present in both trees. However, the topologies of the ITS tree and the respective part of the plastid tree (*Anecphya-Brachyceras* clade) are not fully congruent. There are, for example, some differences in the position of the *N. violacea* samples within subg. *Anecphya*. These are not discussed here, since the complex evolutionary patterns in this subgenus have been discussed earlier by Löhne et al. (2008).

The ITS data set was also analysed with BEAST, using the ages estimated with the *trn*T–*trn*F data set for the *Anecphya-Brachyceras* clade and the *Anecphya* crown group as calibration points. The BEAST summary tree (Fig. 4) shows the same topology as the MP/BA tree (Fig. 3).

Discussion

Relationships within the *Anecphya-Brachyceras* clade

The dense taxon sampling of water-lilies provides some indication that both subg. *Anecphya* and subg. *Brachyceras* are monophyletic. Previously, neither maximum parsimony nor Bayesian tree inference applied to a smaller *trn*T–*trn*F data set had recovered a subg. *Brachyceras* clade (Borsch et al. 2007) whereas a subg. *Anecphya* clade consistently received maximum support. This pattern is also evident for trees inferred from the extended *trn*T–*trn*F data set in this study, although the maximum credibility tree calculated with BEAST (Fig. 2) depicts *Brachyceras* as monophyletic. It has to be noted that this tree gives an alternative hypothesis which is based on a different approach to summarising trees based on the criterion of highest product of posterior probabilities as implemented in the BEAST package. In any case the monophyly of subg. *Brachyceras* needs further testing using additional plastid data.

Internal relationships of the subg. *Brachyceras* clade in the maximum credibility tree, however, are unique in suggesting a "mixed" New World lineage (1.00 PP) also containing the Madagascan *N. minuta* and one deviant sample of *N. micrantha* (NY562). All samples of *N. guineensis* appear in a lineage sister to most other African species, together gaining 1.00 Posterior Probability from the BEAST analysis.

In general, Bayesian Posterior Probabilities are prone to over credit nodes as is evident from simulation (Suzuki et al. 2002) and empirical analyses (Simmons et al. 2004). Recent simulation studies indicate that Bayesian Inference may also group long branches together, especially when sequence sites evolve heterogeneously (Kolaczkowski & Thornton 2009). In the shallow *Brachyceras* clade, sequences exhibit not only very low overall distances but also many autapomorphic mutations that hamper correct model assessment. The BEAST topology (Fig. 2) and similarly the Bayesian tree obtained with MrBayes (not shown) must therefore be interpreted with caution, and the high posterior probabilities may not necessarily indicate correctly resolved nodes.

In contrast, nuclear ITS sequences yield a well-resolved and supported MP tree (Fig. 3), where both subgenera (*Anecphya* and *Brachyceras*) are revealed as well supported monophyletic clades. All species of subg. Brachyceras share characteristic morphological features that support their separation from subg. *Anecphya*: there are free carpellary appendages and much broader filaments in subg. *Brachyceras*, which are absent or filamentous in subg. *Anecphya*. The ITS tree shows that within *Brachyceras* there is a clade, comprising all samples of *N. guineensis* and *N. minuta*, which is sister to the remaining samples. The neotropical species *N. ampla*, *N. elegans*, *N. gracilis* and *N. pulchella* occur in a well supported clade completely nested within African taxa. In subgenus *Anecphya*, two major clades corresponding to the "small seeded group" (comprising *N. macrosperma*, *N. carpentariae*, *N. georginae*, *N. gigantea*, *N. immutabilis* and *N. atrans*) and a "large seeded group" (*N. violacea*, *N. elleniae*, *N. hastifolia* and *N. ondinea*) are recovered as in the earlier more comprehensive analysis by Löhne et al. (2008a). The "small seeded group" has been separated from *Anecphya* as a distinct subgenus *Confluentes* by Jacobs (2007). While reticulate evolutionary patterns are evident in subg. *Anecphya* (Löhne et al. 2008a), there are also some hints on reticulation within subg. *Brachyceras*. This applies especially to the position of *N. minuta*, which is depicted sister to *N. guineensis* by the nuclear marker but nested within the neotropical lineage in the *trn*T–*trn*F BEAST tree (Fig. 2). However, this finding can be only preliminary because phylogenetic structure in the *trn*T–*trn*F sequences is still insufficient (as evident from the polytomy in the MP tree, Fig. 1).

Biogeography of subgenus *Brachyceras*: the neotropical lineage "out-of-Africa"

The addition of further subg. *Brachyceras* species, especially to the nuclear ITS tree of Borsch et al. (2008), provides evidence that the Neotropical species of *Brachyceras* form a lineage nested within the African taxa (Figs 3 & 4). Previous phylogenetic studies of *Nymphaea* with a smaller taxon sampling showed *N. heudelotii* (Africa) sister to *N. gracilis* plus *N. ampla* (Löhne et al. 2007, Borsch et al. 2008). The plastid *trn*T–*trn*F trees of this study (Figs 1 & 2), however, do not allow any insights on historical plant migration patterns within *Brachyceras* due to limited resolution and support. However, the ITS tree suggests an origin "out-of-Africa" of the neotropical *Brachyceras* sublineage (Fig. 3).

Molecular dating based on a *mat*K tree of 25 Nymphaeales taxa (Löhne et al. 2008b) suggested an age for the *Anecphya-Brachyceras* crown group of 9.9 ± 6.5 Mya. Even if the closest palaeotropical relatives of the New World *Brachyceras* lineage are not yet precisely known, it is obvious that their split must be younger. Our estimate with the fossil calibrated minimum ages of *Nuphar* (52 Mya; *Nuphar wutuensis*) and *Nymphaea* (33 Mya, *Nymphaea liminis*) using a much more densely sampled *trn*T–*trn*F data set and BEAST (Fig. 2) increases the hypothesised age of the *Anecphya-Brachyceras* crown group to 24.7 ± 5.8 Mya. Differences may be caused by the denser taxon sampling in this study as better taxon sampling tends to yield older nodal estimates (Linder et al. 2005). Accordingly, our *trn*T–*trn*F analysis hypothesises a split of a *Brachyceras* sublineage mainly composed of New World species to have already taken place 19.6 ± 6.5 Mya. However, this topology should be viewed with caution (see above). A dating approach using the ITS data set of *Anecphya-Brachyceras* has some limits because there are no fossils known for this lineage, and previously calculated node ages including their Credibility Intervals need to be used as calibration points. The revealed minimum

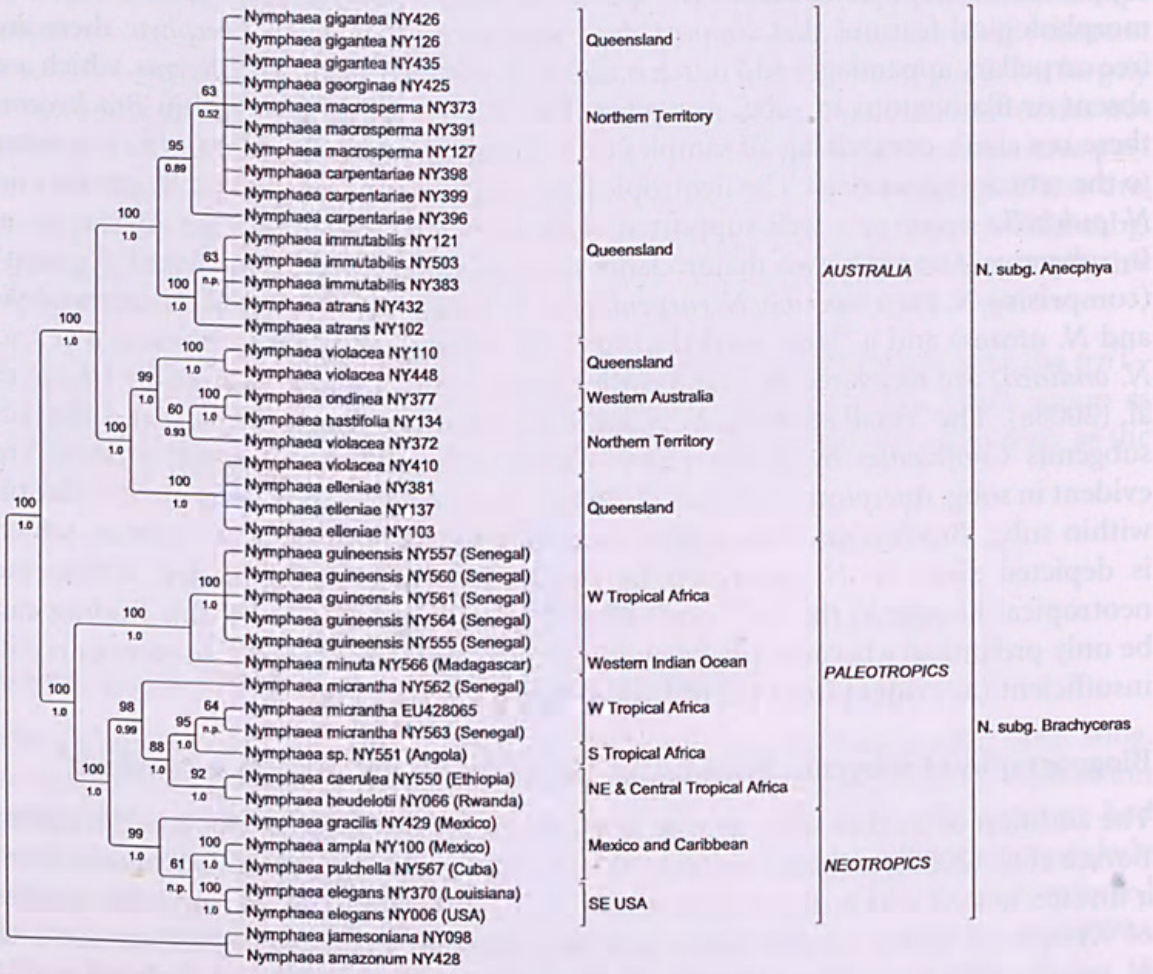

Fig. 3. Phylogeny of *Nymphaea* subg. *Brachyceras* and subg. *Anecphya* as inferred from the nuclear marker ITS. Tree tree shown is the strict consensus of four shortest trees obtained by Maximum Parsimony analysis. Jackknife values are given above branches. Posterior probabilities as obtained from Bayesian analysis with MrBayes (tree not shown) are given below branches. Nodes, which are not present in the Bayesian tree, are marked with *n.p.* (not present).

ages within *N.* subg. *Brachyceras* and *Anecphya* are about twice as old as calculated using a *mat*K tree that, in comparison, sampled only 30% of the taxa (Löhne et al. 2008b). Nevertheless, the split of a clear New World *Brachyceras* lineage is found to have occurred 10 ± 6 Mya (Fig. 4), and thus roughly falls into the same time range as expected from plastid data by Löhne et al. (2008b).

Plant migrations from Africa to the Neotropics are known from various angiosperm families and appear to be even more important in explaining African-Neotropical biogeographic relations than vicariance (Erkens et al. 2009). This is indicated by inferred origins of lineages considerably younger than 100–110 Mya, when the opening of the Atlantic Ocean was initiated (McLoughlin 2001). In several families, early to middle Eocene diversification is reconstructed and migration of angiosperms

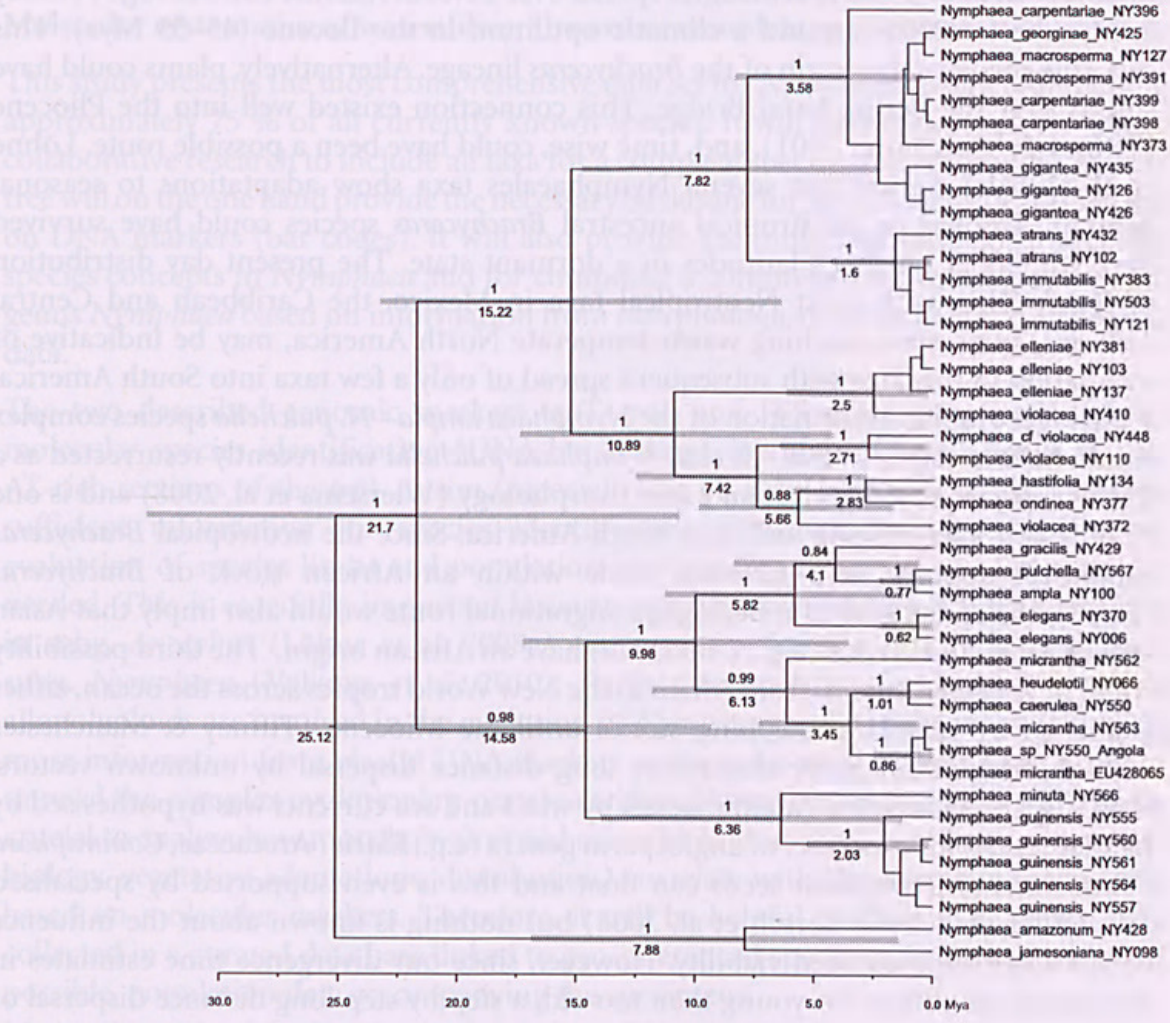

Fig. 4. Summary tree from BEAST analysis of the nuclear marker ITS. Log-normal prior of 24.7 ± 5.8 Mya was set for the MRCA of the ingroup (the *Brachyceras-Anecphya* clade) and 19.6 ± 6.5 Mya for the MRCA of subg. *Anecphya*. These dates were taken from the previous analysis of the *trn*T–*trn*F dataset, in order to allow at least a rough estimation of ages within this group. Posterior probability values above 0.7 are presented above each node. The inferred age of each node is given below the respective branch, while the bars represent 95% confidence intervals of the inferred ages.

adapted to tropical climates proceeded via the North Atlantic or the Bering Land Bridges. This possibility finally ceased during the Oligocene with the cooling of the northern hemisphere (Morley 2003). Meliaceae were shown to have dispersed in the Eocene via Beringia from Africa into the New World (Müllner et al. 2006). For Malpighiaceae, a "Laurasian migration route" via a boreotropical connection between North America and Eurasia during the Eocene was hypothesised (Davis et al. 2002). For Melastomataceae, Renner et al. (2001) reported the arrival of boreotropical taxa in southern continental New World areas in the Oligocene and Miocene. Such a scenario may be true for the ancestor of the Neotropical night blooming water-lilies (Löhne et al. 2008b) that subsequently produced the radiation of the 14 species of subg. *Hydrocallis* distributed from the Caribbean and Central America to South America.

How did the ancestor of the Neotropical *Brachyceras* species reach the Western hemisphere? Theoretically, there are three possible migrational routes (Morley 2003, Erkens et al. 2009). The first assumes spread over a North Atlantic Land Bridge (Tiffney & Manchester 2001) around a climatic optimum in the Eocene (45–55 Mya). This certainly predates the origin of the *Brachyceras* lineage. Alternatively, plants could have spread over the Bering Land Bridge. This connection existed well into the Pliocene (Tiffney & Manchester 2001), and, time wise, could have been a possible route. Löhne et al. (2008b) argued that several Nymphaeales taxa show adaptations to seasonal habitats. Tropical or paratropical ancestral *Brachyceras* species could have survived dark winters at northern latitudes in a dormant state. The present day distribution of *Brachyceras*, with most Neotropical taxa in Mexico, the Caribbean and Central America, some even reaching warm-temperate North America, may be indicative of a radiation in this area with subsequent spread of only a few taxa into South America. A phylogeographic examination of the *Nymphaea ampla* - *N. pulchella* species complex would be interesting in this context. *Nymphaea pulchella* was recently resurrected as a distinct species based on leaf and seed morphology (Wiersema et al. 2008) and is one of only two taxa ranging well into South America. Since the neotropical *Brachyceras* sublineage appears as a terminal clade within an African stock of *Brachyceras* (Fig. 3), the acceptance of a Beringian migrational route would also imply that Asian taxa of subg. *Brachyceras* (e.g. *N. nouchali*) have an African origin. The third possibility could have been dispersal from Africa to the New World tropics across the ocean, either facilitated by islands as stepping stones until the Miocene (Tiffney & Manchester 2001, Morley 2003) or by immediate long distance dispersal by unknown vectors. Seed dispersal across the Atlantic ocean by wind and sea currents was hypothesised by Renner (2004) for a number of angiosperm genera (e.g., *Elaeis* / Arecaceae, *Commiphora* / Burseraceae). *Nymphaea* seeds can float and this is even supported by specialised arils (Wiersema 1987, Borsch et al. 2008) but nothing is known about the influence of salt water on their seed viability. However, since our divergence time estimates in *Nymphaea* are rather too young than too old, a step by step long distance dispersal of *Brachyceras* species into the New World via the Beringian Land Bridge in a still suitable climate may well be a possible explanation.

Another interesting pattern is the close relationship of *N. guineensis* from West Africa and the Madagascan endemic *N. minuta* inferred by ITS (Fig. 3). This would imply a trans-African disjunction of closely related lineages, as observed in several other plant groups (see Sanmartín et al. 2010 for overview). However, in the *trn*T–*trn*F BEAST tree (Fig. 2), *Nymphaea minuta* is nested within a clade with the Neotropical species; but this position is not recovered by the MP analysis (Fig. 1). Thus, further investigation

of phylogenetic relationships, preferably at population levels, will be necessary to draw sound conclusions on the biographic history of *Nymphaea* subg. *Brachyceras* in Africa.

Further sequence data from all genomic compartments will be needed to test the monophyly of subg. *Brachyceras*, and its internal relationships. Tree inference using combined data sets of genomic regions selected for high phylogenetic structure will yield improved node confidence (Borsch & Quandt 2008). A larger number of concatenated genomic regions is also expected to produce better time estimates with smaller credibility intervals using relaxed clock methods (Battistuzzi et al. 2010). Our data also show a trend to relatively higher Credibility Intervals for younger nodes. Dating of shallow phylogenies, as for example is the case within *Brachyceras*, will certainly benefit from a more representative sampling of sequence characters that will allow calculation of more accurate posterior intervals (Brown & Yang 2010). Increasing the molecular data sets of *Nymphaea* will therefore be an important future task.

Molecular systematics of water-lilies: Towards a complete species tree

This study presents the most comprehensive data set for *Nymphaea* to date, comprising approximately 75 % of all currently known species. It will be a major task for future collaborative research to include all taxa for a complete species tree. A complete species tree will on the one hand provide the necessary backbone for species identification based on DNA markers (bar codes). It will also provide the molecular basis for evaluating species concepts in *Nymphaea* and for compiling a comprehensive monograph of the genus *Nymphaea* based on information from morphological, molecular and ecological data.

The two described genomic markers *trn*T–*trn*F and ITS show great potential for molecular species identification (DNA bar-coding). Within *trn*T–*trn*F, there are the AT-rich sections of the *trn*L intron (especially the p8 stem-loop region) that provide sufficient information for species identification. However, as a prerequisite an evaluation of species limits and population level studies across the species' ranges is needed. This is especially important because reticulate evolution has been identified in subg. *Anecphya* (Löhne et al. 2008a) and recently also shown for the temperate subg. *Nymphaea* (Volkova et al. 2010). Further research is also needed to see if allopolyploids are involved in the evolution of *Nymphaea* and its subgenera. Including more information from plastid DNA markers seems to be most promising in order to unravel the complex evolutionary patterns within *Nymphaea*. However, it will also be crucial to analyse how morphological and other biological characters (e.g., pollination biology, vegetative adaptations, distribution) correlate with the complete species tree based on molecular markers. Therefore, it will be helpful to have all sequence data collected in a curated data base linked to geo-referenced specimen information and, if possible, population data accompanying the specimens.

After more than a century since the seminal work of Conard (1905), exploration and description of new water-lily species has still not finished. As evident from Fig. 5, one quarter of the currently accepted species were described in the last 30 years. Main contributions in this sense were the monograph of *Nymphaea* subg. *Hydrocallis* by Wiersema (1987, see also Wiersema et al. 2008) and the contributions by Surrey W. L. Jacobs on *Nymphaea* subg. *Anecphya* (Jacobs 1992, 1994, 2007, Jacobs & Hellquist 2006, Jacobs & Porter 2007). There are still several taxa of uncertain status, especially within the least understood subg. *Brachyceras* but also in subg. *Lotos*, where the identity

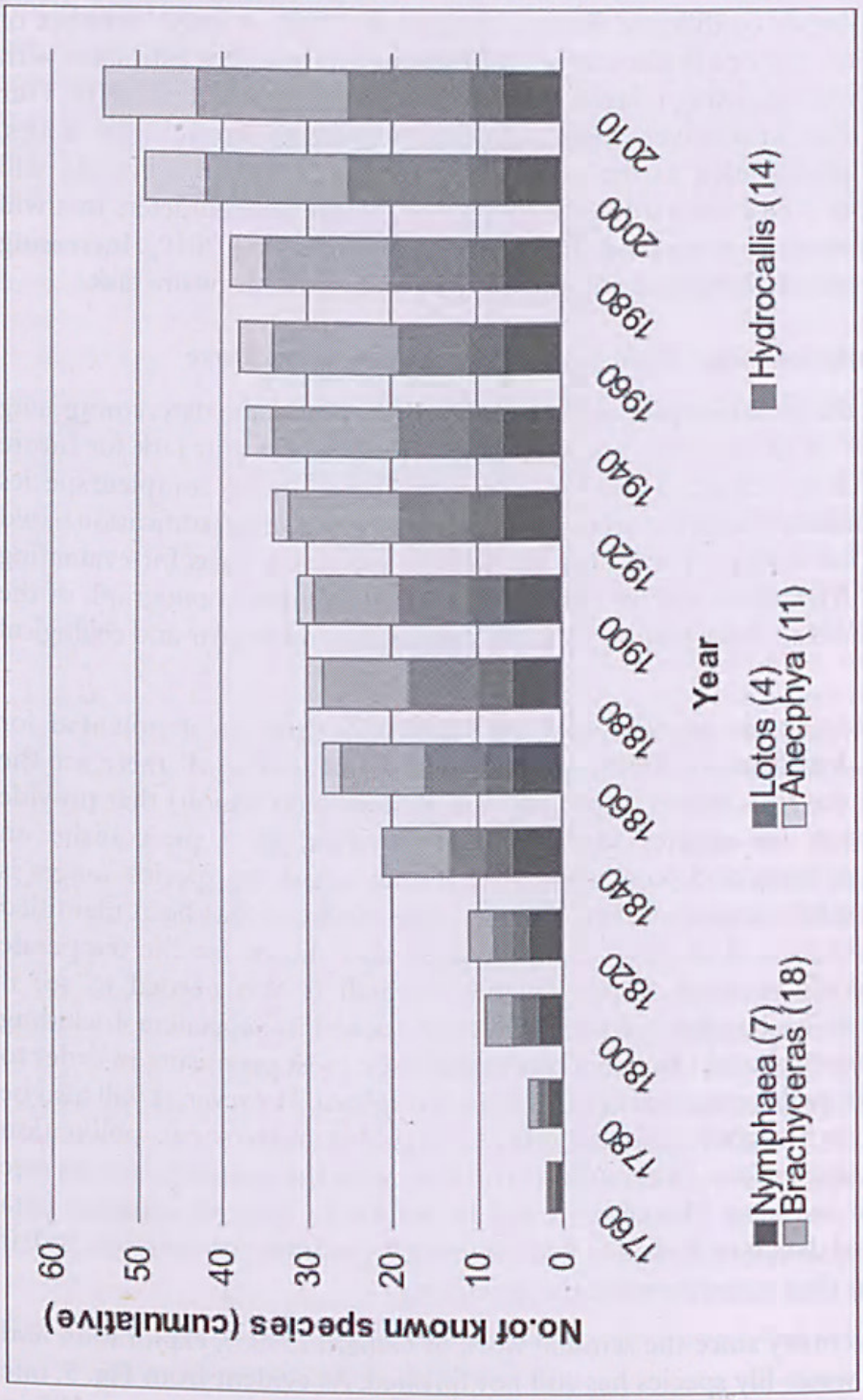

Fig. 5. Number of known species in the genus *Nymphaea* and it's five subgenera *Anecphya, Brachyceras, Hydrocallis, Lotos* and *Nymphaea* since Linnean times. The figure is based on the list of currently accepted species (based on the online database of the USDA Agricultural Research Service, http://www.ars-grin.gov/cgi-bin/npgs/html/splist.pl?8329) and the year of first description of the respective species names. The number behind the subgeneric names indicates the number of currently accepted species. In contrast to the current classification, *Nymphaea petersiana* is here considered as part of subg. *Lotos* because of evidence from DNA (see, e.g., Borsch et al. 2008).

of *N. rubra* needs to be verified (Mitra & Subramanyam 1982, Venu et al. 2003). Thus, it might be expected that future work on these two subgenera will again alter (and probably increase) the total number of species in the genus *Nymphaea*. Additionally, our present study adds further interesting questions, especially with respect to the biogeographical history of subg. *Brachyceras*.

In conclusion, it must be stated that a comprehensive monograph of *Nymphaea* with data on species distribution, ecology and conservation status is very much needed. Such an undertaking should be and actually can only result from the joint effort and expertise from an international working group on *Nymphaea* and the Nymphaeales.

Acknowledgments

We appreciate the help of Bettina Giesicke (Berlin) who generated the new sequences in the lab; and the Botanic Garden Berlin-Dahlem, especially Ms. Daniela Bunde, for taking care of the *Nymphaea* living collection. Furthermore, we thank Carlos Magdalena, Royal Botanic Gardens Kew, for providing seeds of *Nymphaea minuta*.

References

Azuma H, García-Franco JG, Rico-Gray V & Thien LB (2001) Molecular phylogeny of the Magnoliaceae: The biogeography of tropical and temperate disjunctions. *American Journal of Botany* 88: 2275–2285.

Battistuzzi FU, Filipski A, Hedges SB & Kumar S (2010) Performance of relaxed-clock methods in estimating evolutionary divergence times and their credibility intervals. *Molecular Biology and Evolution* 27: 1289–1300.

Berggren WA, Kent DV, Swisher CC & Aubry PP (1995) A revised Cenozoic geochronology and chronostratigraphy. Pp. 129–212 in Berggren WA, Kent DV, Swisher CC, Aubry PP & Hardenbol J (eds) *Geochronology, time scales and global stratigraphic correlation.* (SEPM Spec. Publ.: Tulsa)

Borsch T, Hilu KW, Quandt D, Wilde V, Neinhuis C & Barthlott W (2003) Non-coding plastid *trnT–trnF* sequences reveal a well resolved phylogeny of basal angiosperms. *Journal of Evolutionary Biology* 16: 558–576.

Borsch T, Hilu KW, Wiersema JH, Löhne C, Barthlott W & Wilde V (2007) Phylogeny of *Nymphaea* (Nymphaeaceae): evidence from substitutions and microstructural changes of the chloroplast *trnT–F* region. *International Journal of Plant Sciences* 168: 639–671.

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

Borsch T & Quandt D (2008) Mutational dynamics and phylogenetic utility of non-coding chloroplast DNA. *Plant Systematics and Evolution* 282: 169–199.

Brown RP & Yang Z (2010) Bayesian dating of shallow phylogenies with a relaxed clock. *Systematic Biology* 59: 119–131.

Chanderbali AS, Van der Werff H & Renner SS (2001) Phylogeny and historical biogeography of Lauraceae: evidence from the chloroplast and nuclear genomes. *Annals of the Missouri Botanical Garden* 88: 104–134.

Chen I, Manchester SR & Chen Z (2004) Anatomically preserved seeds of *Nuphar* (Nymphaeaceae) from the Early Eocene of Wutu, Shandong Province, China. *American Journal of Botany* 91: 1265–1272.

Collinson ME (1980) Recent and Tertiary seeds of the Nymphaeaceae sensu lato with a revision of *Brasenia ovula* (Brong.) Reid and Chandler. *Annals of Botany* 46: 603–632.

Conard HS (1905) The waterlilies: a monograph of the genus *Nymphaea*. *Publications of the Carnegie Institutions of Washington* 4: 1–279.

Davis CC, Bell CD, Mathews S & Donoghue MJ (2002) Laurasian migration explains Gondwanan disjunctions: evidence from Malpighiaceae. *Proceedings of the National Academy of Sciences of the United States of America* 99: 6833–6837.

Davis CC, Fritsch PW, Bell CD & Mathews S (2004) High-latitude tertiary migrations of an exclusively tropical clade: evidence from Malpighiaceae. *International Journal of Plant Sciences* 165: S107–S121.

Doyle JA, Sauquet H, Scharaschkin T & Le Thomas A (2004) Phylogeny, molecular and fossil dating, and the biogeographic history of Annonaceae and Myristicaceae (Magnoliales). *International Journal of Plant Sciences* 165: S55–S67.

Drummond AJ & Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7: 214

Drummond AJ, Ho SY, Phillips MJ & Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PloS Biology* 4: e88

Erkens RHJ, Maas JW & Couvreuer TLP (2009) From Africa via Europe to South America: migrational route of a species-rich genus of Neotropical lowland rain forest trees (*Guatteria*, Annonaceae). *Journal of Biogeography* 36: 2338–2352.

Jacobs SWL (1992) New species, lectotypes and synonyms of Australasian *Nymphaea* (Nymphaeaceae). *Telopea* 4: 635–641.

Jacobs SWL (1994) Further notes on *Nymphaea* (Nymphaeaceae) in Australasia. *Telopea* 5: 703–706.

Jacobs SWL (2007) *Nymphaea* subg. *Confluentes*. P. 458 in Wilson AJG (ed.) *Flora of Australia*, vol. 2. (CSIRO and Australian Biological Resources Study: Melbourne)

Jacobs SWL & Hellquist CB (2006) Three new species of *Nymphaea* (Nymphaeaceae) in Australia. *Telopea* 11: 155–160.

Jacobs SWL & Hellquist CB (2011) New species, possible hybrids and intergrades in Australian *Nymphaea* (Nymphaeaceae) with a key to all species. *Telopea* 13: 233–243.

Jacobs SWL & Porter CL (2007) Nymphaeaceae. Pp. 259–275 in Wilson AJG (ed.) *Flora of Australia*, vol. 2. (CSIRO and Australian Biological Resources Study: Melbourne)

Kolaczkowski B & Thornton JW (2009) Long-Branch attraction bias and inconsistency in Bayesian Phylogenetics. *PloS One* 4: e7891.

Landon K, Edwards RA & Nozaic PI (2006) A new species of waterlily (*Nymphaea minuta*, Nymphaeaceae) from Madagascar. *Sida* 22: 887–893.

Linder HP, Hardy CR & Rutschmann F (2005) Taxon sampling effects in molecular lock dating: an example from the African Restionaceae. *Molecular Phylogenetics and Evolution* 35: 569–582.

Löhne C & Borsch T (2005) Molecular evolution and phylogenetic utility of the *petD* group II intron: A case study in basal angiosperms. *Molecular Biology and Evolution* 22: 317–332.

Löhne C, Borsch T & Wiersema JH (2007) Phylogenetic analysis of Nymphaeales using fast-evolving and non-coding chloroplast markers. *Botanical Journal of the Linnean Society* 154: 141–163.

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

Löhne C, Yoo M-J, Borsch T, Wiersema JW, Wilde V, Bell CD, Barthlott W, Soltis DE & Soltis PS (2008b) Biogeography of Nymphaeales: Extant patterns and historical events. *Taxon* 57: 1123–1146.

Löhne C, Wiersema JH & Borsch T (2009) The unusual *Ondinea*, actually just another Australian water-lily of *Nymphaea* subg. *Anecphya* (Nymphaeaceae). *Willdenowia* 39: 55–58.

McLoughlin S (2001) The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. *Australian Journal of Botany* 49: 271–300.

Mitra RL & Subramanyam K (1982) Is *Nymphaea rubra* Roxb. ex Andrews an apomict? *Bulletin of the Botanical Survey of India* 24: 83–86.

Morley RJ (2003) Interplate dispersal paths for megathermal angiosperms. *Perspectives in Plant Ecology, Evolution and Systematics* 6: 5–20.

Müller K (2004) PRAP — computation of Bremer support for large data sets. *Molecular Phylogenetics and Evolution* 31, 780–782.

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

Müller J, Müller KF, Neinhuis C & Quandt D (2007) PhyDE - Phylogenetic Data Editor. Programme distributed by the author. http://www.phyde.de.

Müllner AN, Savolainen V, Samuel R & Chase MW (2006) The mahogany family « out-of-Africa »: Divergence time estimation, global biogeographic patterns inferred from plastid *rbcL* DNA sequences, extant, and fossil distribution of diversity. *Molecular Phylogenetics and Evolution* 40: 236–250.

Nylander JAA (2004) MrModeltest 2.2. Programme distributed by the author. Evolutionary Biology Centre, Uppsala University.

Renner SS (2004) Plant dispersal across the tropical Atlantic by wind and sea currents. *International Journal of Plant Sciences* 165: S23–S33.

Renner SS, Clausing G & Meyer K (2001) Historical biogeography of Melastomataceae: the roles of Tertiary migration and long-distance dispersal. *American Journal of Botany* 88: 1290–1300.

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

Sanmartín I, Anderson CL, Alarcon M, Ronquist F & Aldasoro JJ (2010) Bayesian island biogeography in a continental setting: the Rand Flora case. *Biology Letters* 6: 703–707.

Sauquet H, Doyle JA, Scharaschkin T, Borsch T, Hilu KW, Chatrou LW & Le Thomas A (2003) Phylogenetic analysis of Magnoliales and Myristicaceae based on multiple data sets: implications for character evolution. *Botanical Journal of the Linnean Society* 142: 125–186.

Simmons MP & Ochoterena H (2000) Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology* 49: 369–381.

Simmons MP, Pickett KM & Miya M (2004) How meaningful are Bayesian support values? *Molecular Biology and Evolution* 21: 188–199.

Swofford DL (2002) PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4.0b10., Sinauer Associates, Sunderland, Massachusetts.

Suzuki Y, Glazko GV & Nei M (2002) Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. *Proceedings of the National Academy of Sciences of the United States of America* 99: 16138–16143.

Taberlet P, Gielly L, Pautou G & Bouvet J (1991) Universal primers for amplification of 3 noncoding regions of chloroplast DNA. *Plant Molecular Biology* 17: 1105–1109.

Tiffney BH & Manchester SR (2001) The use of geological and paleontological evidence in evaluating plant phylogeographic hypotheses in the northern hemisphere Tertiary. *International Journal of Plant Sciences* 162: S3–S17.

USDA, ARS, National Genetic Resources Program. *Germplasm Resources Information Network — (GRIN)* [On-line Database]. National Germplasm Resources Laboratory, Beltsville, Maryland. (http://www.ars-grin.gov/cgi-bin/npgs/html/queries.pl) (Accessed 06 June 2010)

Venu P, Rajakumar TJS & Daniel P (2003) On the identity of and sterility in *Nymphaea rubra* Roxb. ex Andrews (Nymphaeaceae). *Rheedea* 13: 29–34.

Volkova PA, Trávníček P & Brochmann C (2010) Evolutionary dynamics across discontinuous freshwater systems: Rapid expansions and repeated allopolyploid origins in Palearctic white water-lilies (*Nymphaea*). *Taxon* 59: 483–494.

White TJ, Bruns T, Lee S & Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pp. 315–322 in Innis DGM, Sninsky J & White T (eds) *PCR protocols: a guide to methods and applications.* (Academic Press: San Diego, CA)

Wiersema JH (1987) A monograph of *Nymphaea* subgenus *Hydrocallis* (Nymphaeaceae). *Systematic Botany Monographs* 16: 1–112.

Wiersema JH, Novelo AR & Bonilla-Barbosa JR (2008) Taxonomy and typification of *Nymphaea ampla* (Salisb.) DC. sensu lato (Nymphaeaceae). *Taxon* 57: 967–974.

Woods K, Hilu KW, Borsch T & Wiersema JH (2005) Pattern of variation and systematics of *Nymphaea odorata*: II. Sequence information from ITS and *trnL–trnF*. *Systematic Botany* 30: 481–493.

Yoo MJ, Bell CD, Soltis PS & Soltis DE (2005) Divergence times and historical biogeography of Nymphaeales. *Systematic Botany* 30: 693–704.

Manuscript received 08 August 2010, accepted 15 November 2010

Appendix 1. Samples included in the study, respective source of material, location of voucher specimens, DNA lab codes and GenBank accession numbers of deposited sequences. Some of the sequence data were generated for this study, while others were taken from our own previous studies ([a] Borsch et al. 2003, [b] Borsch et al. 2007, [c] Löhne et al. 2007, [d] Borsch et al. 2008, [e] Löhne et al. 2008a). Additionally, ITS sequences of *N. micrantha* were taken from GenBank (Volkova et al. 2010).

Taxon	Field / Garden origin	Voucher	DNA	GenBank Numbers	
			lab code	trnT-trnF	ITS
Cabombaceae (outgroup for *trn*T–*trn*F data set)					
Brasenia schreberi J.F.Gmel.	USA, Virginia	T. Borsch & T. Wieboldt 3311 (VPI, FR)	N106	AY145329[a]	--
Cabomba caroliniana A.Gray	USA, Virginia	J.C. Ludwig s.n. (VPI)	N112	AY145328[a]	--
Nymphaeaceae (other than *Nymphaea*)					
Barclaya longifolia Wall.	Bonn Bot Gard [Aquarium plant]	C. Löhne 60 (BONN)	N376	AM422019[b]	--
Euryale ferox Salisb.	Bonn Bot Gard (14010)	T. Borsch 3830 (BONN)	N015	AM422020[b]	--
Nuphar advena (Aiton) W.T.Aiton	USA, Florida	T. Borsch & V. Wilde 3093 (FR)	N080	AY145331[a]	--
Nuphar lutea (L.) Sm.	Germany, Hesse	T. Borsch 3337 (FR)	N107	AY145330[a]	--
Nuphar advena (Aiton) W.T.Aiton subsp. *advena*	USA, Virginia	T. Borsch & T. Wieboldt 3298 (VPI, BONN)	N108	AM422021[b]	--
Nuphar japonica DC.	Bonn Bot Gard [Aquarium plant]	C. Löhne 61 (BONN)	N400	AM422022[b]	--
Victoria cruziana A.D.Orb.	Bonn Bot. Gard.	C. Löhne 55 (BONN)	N316	AM422024[b]	--
Victoria 'Longwood Hybrid'	Bonn Bot. Gard.	T. Borsch 3831 (BONN)	N378	AM422025[b]	--

Taxon	Field / Garden origin	Voucher	DNA	GenBank Numbers	
			lab code	trnT-trnF	ITS
Nymphaea* subg. *Anecphya					
N. atrans S.W.L.Jacobs	Australia, Queensland, 14°38.352'S, 143°54.381'E	S.W.L. Jacobs, C.B. Hellquist & J.H. Wiersema 8212 (NASC, NSW, BRI)	NY102	AM422055[b]	FJ026554[e]
N. atrans S.W.L.Jacobs	Australia, Queensland, 14°38.352'S, 143°54.381'E	C.B. Hellquist & A. Leu 16766 (NASC, NSW, BRI)	NY432	FJ026518[d]	FJ026555[e]
N. carpentariae S.W.L.Jacobs & Hellq.	Australia, Queensland, 17°31.712'S, 141°82.696'E	S.W.L. Jacobs & C.B. Hellquist 8757 (NASC, NSW, BRI)	NY396	FJ026519[d]	FJ026556[e]
N. carpentariae S.W.L.Jacobs & Hellq.	Australia, Queensland, 18°7.362'S,140°32.249'E	S.W.L. Jacobs & C.B. Hellquist 8768 (NASC, NSW, BRI)	NY398	FJ026520[d]	FJ026557[e]
N. carpentariae S.W.L.Jacobs & Hellq.	Australia, Queensland, 17°44.774'S, 139°32.888'E	S.W.L. Jacobs & C.B. Hellquist 8770 (NASC, NSW, BRI)	NY399	FJ026521[d]	FJ026558[e]
N. elleniae S.W.L.Jacobs	Australia, Queensland, 10°53.604'S, 142°23.237'E	S.W.L. Jacobs, C.B. Hellquist, & J.H. Wiersema 8224 (NASC, NSW, BRI)	NY103	AM422056[b]	FJ026560[e]
N. elleniae S.W.L.Jacobs	Australia, Queensland, 11°9.046'S, 142°21.338'E	S.W.L. Jacobs, C.B. Hellquist, & J.H. Wiersema 8227 (NASC, NSW, BRI)	NY137	AM422057[b]	FJ026561[e]
N. elleniae S.W.L.Jacobs	Australia, Queensland, 11°9.046'S, 142°21.338'E	C.B. Hellquist & A. Leu 16757 (NASC, NSW, BRI)	NY381	AM489714[c]	FJ026562[e]
N. georginae S.W.L.Jacobs & Hellq.	Australia, Northern Territory, 20°1.517'S, 137°29.55'E	S.W.L. Jacobs & C.B. Hellquist 8868 (NASC, NSW, PERTH)	NY425	FJ026523[e]	FJ026563[e]
N. gigantea Hook.	Australia, Queensland, 26°59.08'S, 150°6.69'E	S.W.L. Jacobs & C.B. Hellquist 8357 (NASC, NSW, BRI)	NY126	AM422059[b]	FJ026564[e]
N. gigantea Hook.	Australia, Queensland, 26°40.114'S, 150°11,167'E	S.W.L. Jacobs & C.B. Hellquist 8870 (NASC, NSW, BRI)	NY426	FJ026524[e]	FJ026565[e]
N. gigantea Hook.	Australia, Queensland, 18°44.765'S, 146°8.492'E	C.B. Hellquist & A. Leu 16772 (NASC, NSW, BRI)	NY435	FJ026525[e]	FJ026566[e]
N. hastifolia Domin	Australia, Northern Territory, 12°47'S, 130°92'E	J.H. Wiersema & C.B. Hellquist s.n. (no voucher)	NY134	AM422060[b]	FJ026568[e]

Taxon	Field / Garden origin	Voucher	DNA	GenBank Numbers	
			lab code	trnT-trnF	ITS
N. immutabilis S.W.L.Jacobs	Australia, Queensland; 15° 18.19′ S, 144° 36.97′ E	S.W.L. Jacobs, C.B. Hellquist, & J.H. Wiersema s.n. (no voucher)	NY121	AM422061[b]	FJ026569[e]
N. immutabilis S.W.L.Jacobs	Australia, Queensland, 13° 27.052′ S, 142° 42.004′ E	C.B. Hellquist & A. Leu 16760 (NASC, NSW, BRI)	NY383	FJ026529[e]	FJ026573[e]
N. immutabilis S.W.L.Jacobs	Australia, Queensland; 17° 42.229′ S, 145° 7.749′ E	C.B. Hellquist & A. Leu 16775 (NASC, NSW, BRI)	NY503	FJ026528[e]	FJ026572[e]
N. macrosperma Merr. & L.M.Perry	Australia, Northern Territory, Kakadu Nat. Park, ~12° 45′ S, 132° 30′ E	C.B. Hellquist, J.H. Wiersema, & K. Brennan 16181 (MASS)	NY127	AM422063[b]	FJ026579[e]
N. macrosperma Merr. & L.M.Perry	Australia, Northern Territory, Kakadu Nat. Park, ~12° 54′ S, 132° 31′ E	S.W.L. Jacobs & C.B. Hellquist 8802 (NASC, NSW, DNA)	NY373	FJ026533[e]	FJ026577[e]
N. macrosperma Merr. & L.M.Perry	Australia, Northern Territory, Kakadu Nat. Park, 12° 34.25′S, 132° 13.118′ E	S.W.L. Jacobs & C.B. Hellquist 8796 (NASC, NSW, DNA, B, G)	NY391	FJ026534[e]	FJ026578[e]
N. ondinea Löhne et al.	Australia, Western Australia, 14° 15.363′ S, 126° 37.216′ E	S.W.L. Jacobs & C.B. Hellquist 8853 (NASC, NSW, PERTH)	NY377	AM422023[b]	FJ026600[e]
N. violacea Lehm.	Australia, Queensland, 12° 42.62′ S, 142° 30.82′ E	S.W.L. Jacobs, C.B. Hellquist, & J.H. Wiersema 8230 (NASC, NSW, BRI)	NY110	AM422064[b]	FJ026582[e]
N. violacea Lehm.	Australia, Northern Territory, 16° 49.841′ S, 137° 9.546′ E	S.W.L. Jacobs & C.B. Hellquist 8779 (NASC, NSW, DNA)	NY372	FJ026544[e]	FJ026590[e]
N. violacea Lehm.	Australia, Northern Territory, 14° 45.583′ S, 132° 37.166′ E	S.W.L. Jacobs & C.B. Hellquist 8863 (NASC, NSW, DNA)	NY410	FJ026549[e]	FJ026595[e]
N. violacea Lehm.	Australia, Queensland, 13° 27.305′ S, 142° 43.603′ E	C.B. Hellquist & A. Leu 16761 (NASC, NSW, BRI)	NY448	FJ026540[e]	FJ026586[e]

Taxon	Field / Garden origin	Voucher	DNA	GenBank Numbers	
			lab code	trnT-trnF	ITS
Nymphaea* subg. *Brachyceras					
N. ampla (Salisb.) DC.	Mexico, Veracruz	A. Novelo R., J.H. Wiersema, C.B. Hellquist, & C.N. Horn 1295 (MEXU)	NY100	AM422044[b]	FJ026604[e]
N. caerulea Savign.	Bonn Bot Gard 13783	T. Borsch 3834 (BONN)	NY113	AM422045[b]	--
N. caerulea Savign.	Ethiopia	M. Wondafrash 2710 (B, ETH)	NY550	FR717554	FR717585
N. cf. *colorata* Peter	Bonn Bot Gard 1073	T. Borsch 3835 (BONN)	NY122	AM422046[b]	--
N. elegans Hook	USA, Florida, Collier Co.	T. Borsch & V. Wilde 3084 (FR)	NY006	AM422047[b]	FJ026601[e]
N. elegans Hook.	USA, Louisiana, Cameron Parish	T. Borsch & K. Woods 3424 (BONN, VPI)	NY370	AM422048[b]	FJ026602[e]
N. gracilis Zucc.	Mexico, Michoacan	A. Novelo R., J.H. Wiersema, C.B. Hellquist, & C.N. Horn 1346 (MEXU)	NY025	AM422049[b]	--
N. gracilis Zucc.	Mexico, Jalisco	A. Novelo R., J.H. Wiersema, C.B. Hellquist, & C.N. Horn 1314 (MEXU)	NY429	AM422050[b]	FR717586
N. guineensis Schumach.	Senegal, Saint-Louis, 16°3'25"N 16°24'16"W	J.H. Wiersema & A.A. Camara 2387 (B, DAKAR)	NY557, NY558	FR717556	FR717588
				FR717560	--
N. guineensis Schumach.	Senegal, Saint-Louis, 16°3'25"N 16°24'16"W	J.H. Wiersema & A.A. Camara 2388 (B, DAKAR)	NY555	FR717555	FR717587
N. guineensis Schumach.	Senegal, Saint-Louis, 16°26'49"N 15°48'55"W	J.H. Wiersema & al. 2393 (B, DAKAR)	NY559, NY560, NY561	FR717559	--
				FR717557	FR717589
				FR717566	FR717590
N. guineensis Schumach.	Senegal, Saint-Louis, 14°6'40"N 16°4'5"W	J.H. Wiersema & al. 2397 (B, DAKAR)	NY564	FR717558	FR717591
N. heudelotii Planch.	Bonn Bot Gard 14244 [Rwanda]	E. Fischer s.n. (BONN)	NY066	AM422052[b]	FJ026603[e]

Taxon	Field / Garden origin	Voucher	DNA	GenBank Numbers	
			lab code	trnT-trnF	ITS
N. micrantha Perr. & Guill.	Bonn Bot Gard 5830 [Zimbabwe]	M. Koehnen s.n. (BONN)	NY007	AM422051[b]	--
N. micrantha Perr. & Guill.	Senegal, Saint-Louis, 14°33'12"N 12°15'35"W	J.H. Wiersema & al. 2394 (B, DAKAR)	NY562	FR717561	FR717592
N. micrantha Perr. & Guill.	Senegal, Saint-Louis, 12°40'20"N 12°16'42"W	J.H. Wiersema & al. 2396 (B, DAKAR)	NY563	FR717562	FR717593
N. micrantha Perr. & Guill.	--	Volkova et al. (2010)	--	--	EU428065
N. minuta K.C.Landon et al.	Berlin Bot Garden 235-01-09-30 [ex RBG Kew]	no voucher	NY566	FR717564	FR717594
N. pulchella DC.	Cuba, Guantanamo, Baracoa	T. Borsch 4296 (B, HAJB)	NY567	FR717563	FR717596
N. thermarum Eb.Fisch.	Bonn Bot Gard 12088 [Rwanda]	E. Fischer s.n. (BONN)	NY065	AM422054[b]	--
Nymphaea sp.	Angola	W. Lobin s.n. (BONN, B)	NY551	FR717565	FR717595
Nymphaea* subg. *Hydrocallis					
N. amazonum Mart. & Zucc.	Mexico, Veracruz	A.Novelo R., J.H.Wiersema, C.B. Hellquist & C.N.Horn 1281(MEXU)	NY428	AM422026[b]	FR717598
N. belophylla Trickett	Colombia, Meta	U.Schmidt-Mumm 942 (no voucher)	NY027	AM422027[b]	--
N. conardii Wiersema	Mexico, Veracruz	A. Novelo R., J.H.Wiersema, C.B. Hellquist & C.N. Horn 1306 (MEXU)	NY022	AM422028[b]	--
N. gardneriana Planch.	Guyana, Upper Takutu-Upper Essequibo Distr.	C.N.Horn & J.H.Wiersema 10084 (US, BRG, NBYC)	NY026	AM422029[b]	--
N. glandulifera Rodschied	Guyana, Pomeroon Distr.	C.N.Horn & J.H.Wiersema 4523 (US, BRG, NBYC)	NY390	AM422030[b]	--
N. lingulata Wiersema	Guyana, Upper Takutu-Upper Essequibo Dist.	C.N.Horn & J.H.Wiersema 11000 (US, BRG, NBYC)	NY029	AM422031[b]	--

Taxon	Field / Garden origin	Voucher	DNA	GenBank Numbers	
			lab code	trnT-trnF	ITS
N. jamesoniana Planch.	USA, Florida	T. Borsch & B. Summers 3220 (BONN, MO)	NY071	AM422032[b]	--
N. jamesoniana Planch.	Ecuador	M. Schwerdtfeger (BONN, GOET)	NY098	AM422033[b]	FR717597
N. novogranatensis Wiersema	Mexico, Oaxaca	A.Novelo R. & J.H.Wiersema 1187 (MEXU)	NY021	AM422034[b]	--
N. oxypetala Planch.	Bolivia, Santa Cruz	N. Ritter, G. E. Crow, M. Garvizu, & C. Crow 4491 (NHA)	NY387	AM422035[b]	--
N. potamophila Wiersema	Guyana, Upper Takutu-Upper Essequibo Distr.	C.N.Horn & J.H.Wiersema 11090 (US, BRG, NBYC).	NY389	AM422036[b]	--
N. rudgeana G.Mey.	Guyana, Mahaica-Berbice Distr.	C.N. Horn, S. Hill, & D. Gopaul 10045 (US, BRG, NBYC)	NY032	AM422037[b]	--
N. rudgeana G.Mey.	BG Bonn 1088 [Guyana]	Koehnen s.n.(BONN)	NY124	AM422038[b]	--
N. tenerinervia Casp.	Guyana, Upper Takutu-Upper Essequibo Distr.	C.N.Horn & J.H.Wiersema 11086 (US, BRG, NBYC)	NY140	AM422039[b]	--
Nymphaea* subg. *Lotos					
N. lotus L. var. *thermalis* (DC) Tuzson	Bonn Bot Gard 11547-11 [Romania]	T. Borsch 3832 (BONN)	NY003	AM422040[b]	--
N. lotus L. var. *thermalis* (DC) Tuzson	Bonn Bot Gard 05553 [Romania]	T. Borsch 3833 (BONN)	NY105	AM422041[b]	--
N. lotus L. var. *lotus*	Ivory Coast	S. Porembski s.n. (no voucher)	NY078	AM422042[b]	--
N. petersiana Klotzsch	Malawi	Ch. Chawanje s.n. (BONN, FR)	NY058	AM422053[b]	FM242156[d]
N. pubescens Willd.	Australia, Northern Territory	S.W.L. Jacobs 8798 (NSW)	NY406	AM422043[b]	--

Taxon	Field / Garden origin	Voucher	DNA	GenBank Numbers	
			lab code	trnT-trnF	ITS
Nymphaea* subg. *Nymphaea					
N. alba L.	Germany, Bavaria, Luttensee	T. Borsch 3339 (BONN)	NY056	AM422066[b]	--
N. alba L.	Finland, Nylandia, Porvoo	T. Borsch 3151 (BONN, H)	NY061	AM422067[b]	--
N. candida C.Presl	Finland, Tavastia australis, Katloisteryärvi	T. Borsch 3154 (BONN, H)	NY062	AM422068[b]	--
N. candida C.Presl	Finland, Tavastia australis, Maaranyärvi	T. Borsch 3152 (BONN, H)	NY063	AM422069[b]	--
N. candida C.Presl	Russia, Siberia	C.B. Hellquist s.n. (MASS)	NY109	AM422070[b]	--
N. mexicana Zucc.	USA, Florida	T. Borsch & B. Summers 3226 (BONN, VPI)	NY069	AM422071[b]	--
N. mexicana Zucc.	USA, Texas	K. Woods & T. Borsch 701 (BONN, VPI)	KN008	AM422072[b]	--
N. odorata Aiton subsp. *odorata*	USA, Georgia, Okefenokee Swamp	T. Borsch & V. Wilde 3132 (BONN, VPI)	NY012	AY145333[a]	--
N. odorata Aiton subsp. *tuberosa* (Paine) Wiersema & Hellq.	Canada, Manitoba	T. Borsch, J.H. Wiersema, C.B. Hellquist 3389 (BONN, NASC)	NY269	AM422073[b]	--
N. tetragona Georgi	Finland, Tavastia australis, Kanajärvi	T. Borsch 3155 (BONN, H)	NY060	AM422074[b]	--