Evolutionary dynamics across discontinuous freshwater systems: Rapid expansions and repeated allopolyploid origins in the Palearctic white water-lilies (*Nymphaea*)

Polina A. Volkova,[1,2] Pavel Trávníček[3] & Christian Brochmann[1]

1 *National Centre for Biosystematics, Natural History Museum, University of Oslo, P.O.Box 1172 Blindern, 0318 Oslo, Norway*
2 *Moscow State University, Faculty of Biology, Department of Higher Plants, Vorobyevy Gory, Moscow, Russian Federation, 119899*
3 *Institute of Botany, Academy of Sciences of the Czech Republic, Průhonice 1, 252 43, Czech Republic; and Department of Botany, Faculty of Science, Charles University in Prague, Benátská 2, 128 01 Prague, Czech Republic*
Author for correspondence: *Polina Volkova, avolkov@orc.ru*

Abstract White water-lilies (*Nymphaea*) provide an excellent model for studying evolutionary diversification in a discontinuous habitat system, which has experienced dramatic differences in connectivity during different climatic periods. Here we explore cytogenetic and molecular variation and phylogenetic relationships in the Eurasian clade of this genus based on relative nuclear DNA contents, AFLP fingerprints, and nuclear (ITS) and cpDNA sequences. In line with some recent taxonomic treatments, our results support that the Eurasian clade consists of three species. We show that the Eurosiberian *N. candida* is an allopolyploid that originated at least twice by hybridization in the wide overlap zone between the mainly European *N. alba* lineage and the mainly Asian *N. tetragona* lineage. The *N. alba* lineage served as the chloroplast donor except for the Karelian *N. candida*, but the ITS region of the allopolyploid was homogenized in the direction of *N. alba* in all cases. The southern, widespread lineage of *N. candida* probably originated first, whereas the northwestern lineage, restricted to a small, previously glaciated area in Karelia, may have originated as late as in the current interglacial. We observed only low levels of DNA sequence variation within *N. alba* and *N. tetragona*, indicating rapid range expansions possibly preceeded by strong bottlenecks in glacial refugia. Our results suggest that populations of freshwater organisms found now in disjunct aquatic habitats across Eurasia were connected during previous climatic periods. These connections promoted rapid expansions of divergent lineages and repeated formation of evolutionary novelty via hybridization and polyploidization.

Keywords AFLP; allopolyploid speciation; cpDNA; flow cytometry; ITS; *Nymphaea*; Pleistocene glaciations

■ INTRODUCTION

The major glaciations of the Pleistocene affected large parts of Europe and western Siberia. These glaciations had a strong impact on the Eurasian paleohydrological system, causing multiple connections between presently isolated drainages, although the geographical details and exact time span of these connections are not well known (Arkhipov & al., 1995; Grosswald & Hughes, 2002). During some periods of the Pleistocene, the northern half of the West Siberian Plain was covered by the Scandinavian/Siberian ice sheet, blocking the drainage of the north-flowing rivers. This resulted in a huge system of ice-dammed lakes, overflowing via the Turgai Pass to the Aral and Caspian Seas and temporally to the Black Sea (Arkhipov & al., 1995; Grosswald & Hughes, 2002). According to the "maximum" Arctic Ice Sheet reconstruction (Grosswald & Hughes, 2002), the Siberian part of the ice sheet existed during all four last major glaciation periods (the Late Saalian, the Early, the Middle and the Late Weichselian). According to the "minimum" reconstruction (Arkhipov & al. 1995; Svendsen & al., 2004), the Siberian part of the ice sheet existed only during the Saalian and, to a lesser extent, during the Early Weichselian. In Europe, there were also large-scale drainage systems that were parallel to the ice margin during all the Pleistocene glaciations (Arkhipov & al., 1995). Additional temporary connections between different Eurasian drainages appeared due to melting water following ice retreat in the interglacial periods (Grosswald & Hughes, 2002).

The Pleistocene climatic fluctuations produced great changes in species distributions and population genetic structure (Hewitt, 2000). The genetic structure of freshwater species can be more readily interpreted in terms of historical effects than contemporary processes, as their genetic diversity is largely controlled by the island-like nature of their habitat (Ward & al., 1994). Changes in the connectivity of the presently discontinuous aquatic habitat systems during the Pleistocene climatic fluctuations must thus drastically have affected the dispersal possibilities of aquatic organisms, as it has been shown for a number of freshwater fishes (e.g., Bernatchez & Wilson, 1998; Gum & al., 2005) and freshwater crustaceans (e.g., Verovnik & al., 2005). High levels of admixture among lineages in contact zones between drainages and fast range expansions were found for these organisms, most likely resulting from glacial perturbations and ancient river connections (Gum & al., 2005; Verovnik & al., 2005).

However, there are so far no large-scale molecular studies on freshwater vascular plants to assess how changes in the paleohydrological systems may have affected their diversification,

expansion, and putative hybridization. According to Iversen (1954), aquatic plants may represent more sensitive indicators of temperature changes than terrestrial plants. Such sensitiveness of aquatic plants to climate changes has recently been substantiated by fossil pollen data. Extensive studies of nine fresh-water plant genera (e.g., *Nymphaea*, *Potamogeton*) in North America showed that most of the taxa reached the modern extent of their geographic ranges almost as rapidly as the glacial ice receded and the lakes became seasonally ice-free (Dieffenbacher-Krall & Jacobson, 2001).

We selected Eurasian water-lilies (*Nymphaea*, Nymphaeaceae) as a model group for such a study because they are morphologically polymorphic, have wide ecological and geographical distributions, and are ecologically important by dominating various fresh-water communities (Heslop-Harrisson, 1955; Borsch & al., 2008; Borsch & Soltis, 2008). The only large-scale inference of *Nymphaea* phylogeny has been based on the chloroplast *trnT-trnF* region, but although this study provided strong evidence for monophyly of the Eurasian clade of white water-lilies, it did not resolve its internal relationships (Borsch & al., 2007). This Eurasian clade is taxonomically problematic because of its high level of morphological polymorphism (Heslop-Harrison, 1955; Komarov, 1970; Volkova & Shipunov, 2008). There are both "splitting" approaches recognizing more than ten taxa, and "lumping" treatments with a single polymorphic species, *N. alba* (reviewed in Heslop-Harrison, 1955). Extensive interspecific hybridization has also been suggested by many authors (Heslop-Harrison, 1955; Uotila, 2001) but convincing evidence for this has not yet been provided (Les & Philbrick, 1993). Most modern treatments of the genus recognize three species in Eurasia, namely *N. alba* L. s.str., *N. candida* Presl (these two are sometimes treated as subspecies of *N. alba*; Heslop-Harrisson, 1955; Uotila, 2001) and *N. tetragona* Georgi (Komarov, 1970; Volkova & Shipunov, 2008). We used this treatment as an initial framework for our study (except that we tentatively also recognized *N. wenzelii*, see below) and follow it throughout this paper, in line with our final conclusions. *Nymphaea alba* is widespread in Europe and northernmost Africa, *N. candida* is Eurosiberian, and *N. tetragona* inhabits the major part of Asia as far south as India and as far east as China and Japan, and also occurs in NW Europe (Heslop-Harrisson, 1955; Komarov, 1970; Uotila, 2001; Fig. 1). The morphology of *N. candida* is highly variable but mainly intermediate between *N. alba* and *N. tetragona* (Volkova & Shipunov, 2008), what allowed Korzhinsky (1892) to suppose a hybrid origin of *N. candida*. AFLP fingerprinting has provided evidence for local hybridization between *N. alba* and *N. candida* in some sites in Germany and Sweden (Werner & Hellwig, 2006).

Among the other described taxa of Eurasian *Nymphaea*, we tentatively recognized *N. wenzelii* Maack from the Russian Southern Far East, which resembles a large-grown *N. tetragona* (Maack, 1859) and has been variably recognized as a separate species, as a subspecies of *N. tetragona*, as a subspecies of *N. candida*, or even as a *Nymphaea* × *Nuphar* intergeneric hybrid (Komarov, 1970).

Here we investigate the relationships and variation in the Eurasian white water-lilies using AFLP fingerprints, nuclear (ITS) and cpDNA sequences, and estimates of nuclear DNA content. Our study addressed the following questions: (1) Which among the available taxonomic treatments is supported by the present data? (2) Is *N. candida* a product of hybridisation between *N. alba* and *N. tetragona*? (3) How did the Pleistocene glaciations impact the genetic structure of the *Nymphaea* species in Eurasia? In particular, our intention was to explore the relative importance of diversification and hybridization in a vast and presently discontinuous habitat system, which has experienced dramatic differences in connectivity during different climatic periods. We expected that ancient connections between currently fragmented areas of suitable habitat would may have facilitated extensive admixture of lineages and interspecific hybridization. We also expected low levels of genetic variation, at least in recently colonized areas, due to rapid and extensive range changes facilitated by connections between different drainage systems in the past.

■ MATERIALS AND METHODS

Study group. — *Nymphaea* L. (Nymphaeaceae) contains about 45 perennial species inhabiting fresh-water reservoirs and streams of both tropical and temperate regions worldwide (Conard, 1905). The genus is divided into five subgenera: the paleotropical *N.* subg. *Lotos* (DC.) Conard (2–3 species), the neotropical subg. *Hydrocallis* (Planch.) Conard (14 species), the pantropical subg. *Brachyceras* (Gaspary) Conard (14–16 species), the Australian and New Guinean subg. *Anecphya* (Gaspary) Conard (7–10 species) and the temperate subg. *Nymphaea* L. (5–10 species) (Conard, 1905). Recently this classification was supported in a phylogenetic analysis based on several regions of cpDNA, placing the temperate *N.* subg. *Nymphaea* as sister to the other subgenera (Borsch & al., 2007; Löhne & al., 2007). This subgenus contains the southern North American *N. mexicana* Zucc., the North American *N. odorata* Ait. in addition to the Eurasian species of focus in the present study (Borsch & al., 2007; see above).

The Eurasian white water-lilies are mainly cross-pollinated by insects, but self-pollination also occurs (Heslop-Harrison, 1955). Their numerous (about 500 per fruit) and large (2–5 mm) seeds float on the water for two to three days after liberation from the fruit, promoting passive dispersal by currents. The seeds are also eaten by fish and birds, but the main dispersal vector is probably water currents. Vegetative reproduction occurs locally by branching of the rhizome and death of the older parts, and occasionally, rhizome parts break off and may be dispersed by water currents (Heslop-Harrisson, 1955).

Polyploidy and chromosomal rearrangements may have played an important role in the evolution of *Nymphaea* (Les & Philbrick, 1993). The basic chromosome number is uncertain, but usually thought to be $x = 14$ (Heslop-Harrison, 1955; Krupkina, 2001). Many chromosome numbers have been reported for the Eurasian *Nymphaea* species: $2n = 48$, 52, 56, 64, 84, 96, 105, 108, 112, and 160 for *N. alba*; $2n = 112$ and 160 for *N. candida*; and $2n = 84$, 112, and 120 for *N. tetragona* (Bolkhovskikh & al., 1969). However, these numbers should be

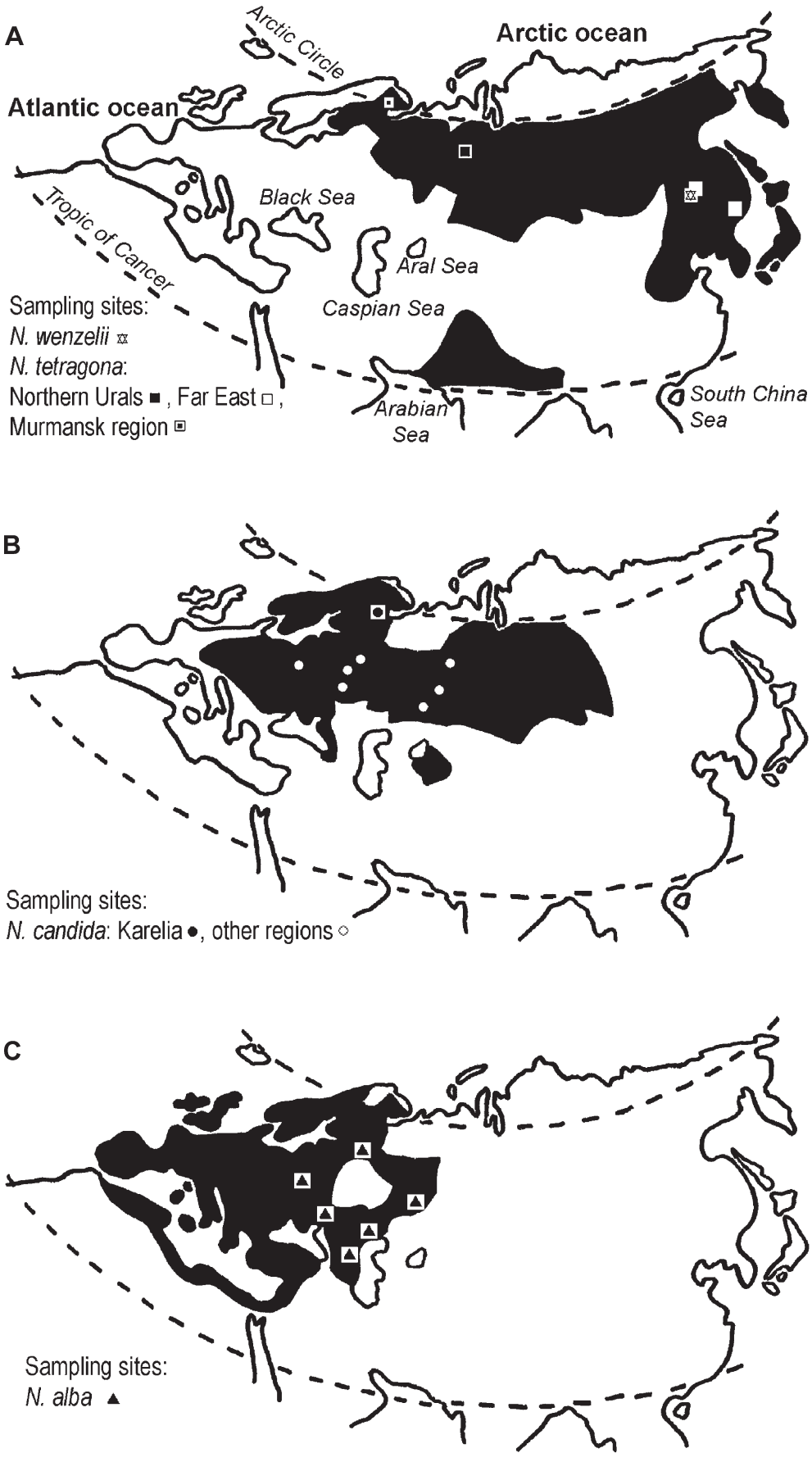

Fig. 1. Distributional range (according to Komarov, 1970; Uotila, 2001) and sampling sites of *Nymphaea* spp. (cf. Table 1): **A,** *N. tetragona* (incl. *N. wenzelii*), **B,** *N. candida*, **C,** *N. alba*. Symbols indicate species and phylogeographical subgroups.

interpreted with caution for several reasons. The high number and small size of the chromosomes make them difficult to count precisely (Heslop-Harrison, 1955). Moreover, somatically arising variation has been suggested (Gupta, 1980). Finally, some counts may be based on misidentified material and others may represent interspecific hybrids.

Sampling. — A total of 178 individuals were sampled in the field from 35 populations, ranging from one to nine (mean 5.1) individuals from each population (Fig. 1; Table 1). Because the plants are rhizomatous and capable to form large clones, they were collected at least 5 m apart to minimize resampling of clones (usually genets are not larger than 5 m across: Heslop-Harrison, 1955). The plants were selected to cover the morphological variation observed in each population. Leaf samples from all the plants, except for populations 217, 106, 119, 121 and 51 (cf. Table 1), were dried in silica gel for DNA analysis, and one plant per population was pressed as a voucher. We used the herbarium vouchers from the abovementioned populations for the sequencing part in addition to the silica-dried samples. We did not succeed in using older specimens from various herbaria for sequencing because of their low quality (leaves of water-lilies tend to dry too slowly during traditional preparation of herbarium specimens). Our field sampling included all three Eurasian species of water-lilies sensu Komarov (1970; *N. alba*, *N. candida* and *N. tetragona*) and also *N. wenzelii* from the Russian Far East. We were not able to sample across the entire range of the study group but focused on Eastern Europe, where all three putative species occur and are believed to hybridize freely (Komarov, 1970). Voucher specimens were deposited in the herbarium of Moscow state university (MW), Russia. We also included silica-dried leaf samples of the two North American species of *N.* subg. *Nymphaea* (*N. odorata* subsp. *odorata* and *N. mexicana*) and some species of the other subgenera in the sequencing analysis (Table 2).

DNA isolation and AFLP analysis. — DNA was extracted from dehydrated leaf material with the DNeasy Plant Mini Kit and DNeasy Plant 96 Kit (QIAGEN, Hilden, Germany) following the manufacturer's instructions. The DNA extracts are deposited in the DNA bank of the National Centre for Biosystematics at the Natural History Museum in Oslo, Norway. We extracted 22 randomly selected plants twice. These duplicates, as well as negative controls, were included to test for reproducibility and contamination in all stages of the AFLP analysis (cf., Bonin & al., 2004).

AFLP analysis was performed on silica-dried leaf material according to Gaudeul & al. (2000), except that the reaction volumes in the polymerase chain reaction (PCR) were reduced by 50%, that the preselective PCR-products were diluted 20 times, and that we used preselective PCR conditions following Werner & Hellwig (2006) to improve the quality of AFLP profiles (initial denaturation for 10 min at 95°C followed by 12 cycles of 30 s at 94°C, 30 s at 65°C–56°C, and 1 min at 72°C, and 24 cycles of 30 s at 94°C, 30 s at 56°C, and 1 min at 72°C, ending with 10 min at 72°C) (C. Ritz, Institute of Systematic Botany in Jena, Germany, pers. comm.). Nineteen primer pair combinations were tested on eight plants of the three species. AFLP profiles with many polymorphic markers and well separated fragments were selected. A second primer test was carried out using five primer pair combinations, chosen from the first primer test, on 32 plants of the three species from several geographical locations. Three primer pair combinations were chosen for full AFLP analysis: *Eco*RI AGT (6FAM)-*Mse*I CTA, *Eco*RI ACA (VIC)-*Mse*I CGT, and *Eco*RI ACA (NED)-*Mse*I CCA. For each individual, 2.0 µl 6-FAM, 2.0 µl VIC and 3.0 µl NED labeled selective PCR products were mixed with 11.7 µl formamide and 0.3 µl GeneScan ROX 500 size standard and run on an ABI 3100 sequencer (Applied Biosystems, Foster City, California, U.S.A.).

Raw data were analyzed using the ABI prism GeneScan v.3.7 analysis software (Applied Biosystems) and imported for scoring into Genographer (v.1.6. available at http://hordeum.oscs.montana.edu/genographer/). Fragments in the size range of 70–500 bp were scored as present or absent. Principal coordinate analysis (PCO) was used to visualize pair-wise similarities between the AFLP multilocus phenotypes, using both the simple matching and Dice similarity coefficients. The Dice coefficient is only based on similarity in presence of fragments, while the simple matching coefficient takes both presence and absence of fragments into account. PCO analyses were executed in the R 2.1.1 environment for statistical computing (R Development Core Team, 2004), using the package "arules" (Hahsler & al., 2005). Very similar results were obtained from the analyses based on the two different coefficients, and only that based on Dice is presented. Analyses of molecular variance (AMOVA; Excoffier & al., 1992) were computed with the software Arlequin v.3.0.1 (Excoffier & al., 2005) to quantify genetic differentiation at different hierarchical levels. The significance levels of the variances were estimated in a permutation test with 10,000 permutations.

The genetic structure was also examined by genetic mixture analysis using the software STRUCTURE v.2.2 (Pritchard & al., 2000), adapted for dominant markers (Falush & al., 2007) with a model-based clustering method using Markov Chain Monte Carlo estimation. This program identifies the optimal number of groups (K) by comparing the likelihood of the data estimated in different runs for different numbers of groups. Individuals are assigned (probabilistically) to one group defined by allele frequencies at each locus. Our data were analyzed at the Bioportal, University of Oslo (http://www.bioportal.uio.no), with K ranging from one to ten, ten replicate runs for each K and a burn-in period of 2×10^5 and 10^6 iterations. Both the admixture model and the no admixture model and uncorrelated allele frequencies were assumed for the analyses. We run STRUCTURE for the whole dataset and for each of three putative species separately. We calculated similarity coefficients among pairs of STRUCTURE runs to evaluate the reliability of K estimates according to Rosenberg & al. (2002) using the modified R-script AFLPdat (Ehrich, 2006).

We performed multilocus assignment tests using AFLPOP (Duchesne & Bernatches, 2002) to test the hypothesis on hybrid origin of plants referred to *N. candida*. In AFLPOP the differences in frequencies at polymorphic loci are used to assign a specimen to its most probable species or simulated F1 hybrids

Table 1. Geographic origin, average relative fluorescence intensity of 4′,6-diamidino-2-phenylindole (DAPI)-stained nuclei (with *Lycopersicon esculentum* cv. 'Stupické polní tyčkové rané' as internal standard) and GenBank accession numbers of the investigated populations of Eurasian *Nymphaea*. Tentative determinations based on morphology were made by P. Volkova prior to the molecular analysis (these were in agreement with the final determinations except that three populations (732, 737, 738) initially referred to *N. wenzelii* later were included in *N. tetragona*).

Population no.	Geographic origin (cf. Fig. 1)	N_{AFLP}	D	N_{FCM}	Average relative fluorescence intensity ± SD	GenBank accession nos.		
						petL-psbE	*rpl32-trnL*	ITS
	N. alba							
217	Astrakhan' Region, mouth of River Volga	–	–	–	–	EU428075	EU428133	EU428030
701	Kiev Region, River Kozinka	4	0.138	5	1.849 ± 0.042	EU428077	EU428135	EU428032
703	Kiev Region, River Dnepr	5	0.187	4	1.827 ± 0.046	EU428079	EU428137	EU428033
704	Kiev Region, River Dnepr	4	0.236	4	1.807 ± 0.039	EU428080	EU428138	EU428034
705	Kiev Region, River Dnepr	3	0.216	3	1.773 ± 0.083	EU428081	EU428139	EU428035
710	Tver' Region, River Volchna	5	0.188	5	1.836 ± 0.027	EU428084	EU428142	EU428037
718	Chelyabink Region, Miass District, Lake Maloje Miassovoe	4	0.272	5	1.817 ± 0.020	–	–	EU428041
742	Voronezh Region, Riv. Usmanka	5	0.227	–	–	EU428108	EU428172	EU428059
805	Dagestan Republic, Lake Shajtak-Kazak	1	–	1	1.762	EU428109	EU428173	EU428060
	N. candida							
106	Tver' Region, Lake Golovets	–	–	–	–	EU428071	EU428119	EU428021
119	Moscow Region, unnamed pond	–	–	–	–	EU428072	EU428122	EU428023
121	Chelyabink Region, Bredy District, River Karaganka	–	–	–	–	EU428073	EU428123	EU428025
702	Kiev city, River Dnepr	5	0.279	4	2.571 ± 0.110	EU428078	EU428136	–
708	Tver' Region, River Kagra	2	0.179	2	2.448 ± 0.014	EU428082	EU428140	–
709	Tver' Region, Lake Glukhoe	4	0.342	5	2.460 ± 0.052	EU428083	EU428141	EU428036
712	Khanty-Mansijskij AO, Lake Aran-Tur	7	0.341	6	2.529 ± 0.089	EU428085	EU428144	–
714	Khanty-Mansijskij AO, Lake Alas	9	0.351	8	2.470 ± 0.055	EU428087	EU428146	–
715	Khanty-Mansijskij AO, Lake Lopukhovoe	5	0.331	4	2.569 ± 0.061	EU428088	EU428147	EU428040
716	Chelyabink Region, Miass District, Lake Bol'shoje Miassovoe	3	0.331	4	2.542 ± 0.083	–	–	–
717	Chelyabink Region, Miass District, Lake Bol'shoje Miassovoe	4	0.307	5	2.542 ± 0.049	EU428089	EU428149	–
719	Chelyabink Region, Miass District, Lake Bol'shoje Miassovoe	3	0.358	5	2.562 ± 0.031	–	–	–
720	Chelyabink Region, Miass District, Lake Bol'shoje Miassovoe	4	0.310	4	2.585 ± 0.024	EU428090	EU428152	EU428042
721	Chelyabink Region, Miass District, Lake Argayash	4	0.242	5	2.505 ± 0.027	EU428091	EU428153	–
723	Karelia Republic, Kostomuksha District, Lake Kamennoe	4	0.272	5	2.580 ± 0.049	EU428096	EU428158	–
724	Karelia Republic, Kostomuksha District, unnamed lake-1	5	0.132	4	2.578 ± 0.021	EU428097	EU428159	EU428047
725	Karelia Republic, Kostomuksha District, unnamed lake-2	5	0.230	4	2.543 ± 0.055	EU428098	EU428160	–
741	Voronezh Region, bog near village Maklok	5	0.234	–	–	–	–	EU428058
	N. tetragona							
51	Murmansk Region, River Kuzreka	–	–	–	–	EU428076	EU428134	EU428031
711	Khanty-Mansijskij AO, Lake Aran-Tur	7	0.119	9	0.803 ± 0.009	–	–	EU428038
713	Khanty-Mansijskij AO, River Akh	9	0.190	9	0.802 ± 0.013	EU428086	EU428145	EU428039
730	Amur Region, unnamed lake near town Arkhara	5	0.131	5	0.824 ± 0.024	EU428099	EU428161	EU428048
732	Amur Region, Lake Krivoe	4	0.129	5	0.799 ± 0.010	EU428100	EU428162	EU428049
733	Amur Region, Lake Glubokoje	5	0.118	5	0.821 ± 0.013	EU428101	EU428163	EU428050
734	Amur Region, Lake Bol'shoje Pereshjejechnoje	5	0.122	4	0.813 ± 0.012	EU428102	EU428164	EU428051
735	Amur Region, Lake Dolgoe	4	0.115	4	0.816 ± 0.018	EU428103	EU428165	EU428052
736	Amur Region, Lake Kljoshinkoje	5	0.085	5	0.811 ± 0.009	EU428104	EU428166	EU428053
737	Amur Region, Lake Krivoe	5	0.070	5	0.796 ± 0.012	–	–	EU428055
738	Amur Region, Lake Krivoe	4	0.109	5	0.803 ± 0.017	EU428105	EU428168	EU428054
739	Amur Region, Lake Krivoe	5	0.097	5	0.809 ± 0.014	EU428106	EU428169	EU428056
740	Primorskij Kraj, Lake Mramornoje	7	0.096	7	0.810 ± 0.004	EU428107	EU428170	EU428057

N_{AFLP}, number of individuals analyzed for AFLPs; D, Nei's unbiased diversity measure; N_{FCM}, number of individuals analyzed with flow cytometry.

Table 2. GenBank accession numbers for the investigated non-Eurasian *Nymphaea* species

Species	Origin	GenBank accesion nos.		
		petL-psbE	*rpl32-trnL*	ITS
Nymphaea* subg. *Anecphya				
N. immutabilis S.W.L. Jacobs	Personal collection of C.B. Hellquist (Massachusetts, U.S.A.)	EU428112	EU428176	EU428062
N. violacea Lehm.	Personal collection of C.B. Hellquist (Massachusetts, U.S.A.)	EU428118	EU428185	EU428070
Nymphaea* subg. *Hydrocallis				
N. rudgeana G. Mey	Personal collection of C.B. Hellquist (Massachusetts, U.S.A.)	EU428117	EU428182	EU428068
Nymphaea* subg. *Brachyceras				
N. caerulea Savign.	Personal collection of C.B. Hellquist (Massachusetts, U.S.A.)	EU428110	EU428174	EU428061
N. micrantha Guill. & Perr.	Personal collection of C.B. Hellquist (Massachusetts, U.S.A.)	EU428114	EU428179	EU428065
N. minuta Landon, Edwards & Nozaic	Personal collection of C.B. Hellquist (Massachusetts, U.S.A.)	EU428115	EU428180	EU428066
Nymphaea* subg. *Nymphaea				
N. mexicana Zuccarini	Bergius Botanic Garden, Stockholm, Sweden	EU428113	EU428178	EU428064
N. odorata Aiton	Sarasota County, Florida, U.S.A., sampled in the field on 2 Apr. 2007	EU428116	EU428181	EU428067

(see Kapralov & al., 2006 for details). Marker frequencies of zero were replaced by 1 / (sample size + 1).

Conversion of the output of Genographer to an AFLP data matrix for PCO, generation of the input files for Arlequin and STRUCTURE and calculations of Nei's unbiased intrapopulation gene diversity measure were done with the R-script AFLPdat (Ehrich, 2006).

cpDNA and ITS sequencing. — Seven noncoding and potentially highly variable cpDNA regions were initially amplified and sequenced for one sample of each of five species (*N. mexicana*, *N. odorata*, *N. alba*, *N. candida*, *N. tetragona*): *petL-psbE*, *atpI-atpH*, *rpl32-trnL*, *trnS-trnG-trnG*, *psbJ-petA*, *3'rps16–5'trnK* (Shaw & al., 2007) and *trnD*(GUC)-*trnT*(GGU) (Shaw & al., 2005). We found variation in all these regions, but selected the intergenic spacers *petL-psbE* and *rpl32-trnL* for the final analysis as they were most variable. The complete ITS region (ITS1, 5.8S, ITS2) was amplified using the primers ITS5P (Möller & Cronk, 1997) and ITS4 (White & al., 1990).

Only one plant was analyzed from each population, as none or only little variation between different populations was found within species (Tables 1, 3). We were not able to amplify ITS from all samples, probably due to poor DNA quality. We also sequenced samples of the additional species listed in Table 2. Sequences of *Illicium oligandrum* Merr. & Chun (Illiciaceae) from GenBank were used as an outgroup according to Borsch & al. (2007) in the cpDNA (EF380354) and ITS (EF138795) datasets.

PCR was performed according to Schönswetter & al. (2006), except that the annealing temperature was 50°C. Negative control samples (milliQ-H_2O) were included in all amplification runs. PCR products were checked on agarose gels and cleaned with ExoSAP-IT (USB Corp.). Sequencing was performed in both directions using BigDye v.1.1 (Applied Biosystems) according to the manufacturer's manual except for using 10 µl reaction volumes. The samples were cleaned with Sephadex G-50 gel (GE Healthcare, Little Chalfont, U.K.) and run on an ABI 3100 sequence analyzer (Applied Biosystems).

The program Sequencher v.4.2.1 (Gene Codes Corp., Ann Arbor, Michigan, U.S.A.) was used to compile and edit contiguous sequences from electropherograms generated on the automated sequencer. Polymorphic sites in the ITS sequences were IUPAC coded. The sequences were manually aligned using GeneTool v.2.0 (BioTools Inc.. Edmonton, Canada). Alignment gaps were treated as missing values. Maximum parsimony (MP) analyses were carried out using PAUP* (v.4.0b10) for Macintosh (Swofford, 2002) using a heuristic search strategy starting with 100 random trees, the MULTREES option on, and DELTRAN optimization. Bootstrap analysis was performed with full heuristics, 1000 pseudoreplicates, TBR branch

Table 3. Analyses of molecular variance (AMOVA) of the AFLP data (151 markers) for the three studied species of *Nymphaea* (the populations that were tentatively referred to *N. wenzelii* were included in *N. tetragona* according to molecular data).

	Source of variation	d.f.	% of total variance
Three species	Among species	2	45.4
	Among populations within species	30	7.8
	Within populations	127	46.8
N. alba	Among populations within species	6	8.6
	Within populations	23	91.4
N. candida	Among two geographical subgroups	1	8.2
	Among populations within subgroups	13	9.1
	Within populations	54	82.8
N. tetragona	Among two geographical subgroups	1	17.1
	Among populations within subgroups	10	10.4
	Within populations	53	72.5

P-values were <0.001 in all cases (estimated with 10,000 permutations).

swapping, the MULTREES option off, and random addition of sequences with four replicates including both informative and uninformative characters.

Flow cytometry. — Because chromosome numbers were not available for our samples and because of the limited reliability of the literature data on these species (cf. above), we could not use relative DNA contents for absolute DNA ploidy inference. We measured fluorescence intensity according to the modified two-step Otto procedure (Otto, 1990) on silica-dried leaves. Leaf tissue of *Nymphaea* and of an internal standard (*Lycopersicon esculentum* cv. 'Stupické polní tyčkové rané', 2C = 1.96 pg; Doležel & al., 1992) was chopped with a razor blade in 0.5 ml ice-cold Otto I buffer (0.1 M citric acid, 0.5% Tween 20), filtered through a nylon mesh (42 µm), and incubated at room temperature for 15 min. The staining solution consisted of 1 ml Otto II buffer (0.4 M $Na_2HPO_4 \cdot 12\ H_2O$) with 4′,6-diamidino-2-phenylindole (DAPI; 4 µg/ml) and 2-mercaptoethanol (2 µl/ml). After incubation for 5 min at room temperature, relative fluorescence intensity of at least 5000 particles was recorded using a Partec PA II flow cytometer (Partec GmbH, Münster, Germany) equipped with an HBO mercury arc lamp. The flow cytometry (FCM) results were expressed as fluorescence intensity relative to internal standard (i.e., the intensity of *Lycopersicon esculentum* was set as unit value).

■ RESULTS

AFLPs. — The AFLP analysis provided 151 polymorphic markers for 165 analyzed plants. The average reproducibility, calculated as the average proportion of correctly reproduced bands over all replicates (Bonin & al., 2004), was 98.3% for the three primer pair combinations. We did not observe any private AFLP markers for individual species. However, *N. candida* uniquely shared 23 markers with *N. alba* (i.e., bands which were absent in *N. tetragona*) and 14 with *N. tetragona*. There were no uniquely shared bands between *N. alba* and *N. tetragona*. The average intrapopulation gene diversity (mean Nei's average gene diversity over all populations ± SD) was highest in *N. candida* (0.29 ± 0.043) and *N. alba* (0.21 ± 0.043) and much lower in *N. tetragona* including *N. wenzelii* (0.12 ± 0.030; differences among species significant at $P < 0.05$ in Tukey's honestly significant difference test).

In the PCO analysis (Fig. 2), the plants were divided into three groups along axis 1 (50% of the variation) corresponding to the tentative species. The *N. candida* group was placed in the middle, but closest to the *N. alba* group. The plants tentatively referred to *N. wenzelii* were placed within the *N. tetragona* group, indistinguishable from other plants from the Far East. The second axis (5%) reflected variation within each species. In *N. tetragona*, most plants from the Northern Urals were separated from those from the Far East along axis 2. In *N. candida*, axis 2 partly separated the Karelian plants from the plants from other geographic areas. There was no geographical structuring in *N. alba*.

The STRUCTURE analysis based on the "no admixture" model also revealed three groups corresponding to the three species (Fig. S1 in the Electronic Supplement to this article). The graph of the log likelihood of the AFLP data estimated from three runs showed a point of inflection for three groups (Fig. S1A). Up to three groups, the outputs from the program were identical (Fig. S1B). Separate STRUCTURE analyses of each species resulted in subdivision of *N. tetragona* into one Northern Ural subgroup and one Far East subgroup, but no clear subdivision in the other two species (not shown). The STRUCTURE analyses based on the "admixture" model gave very similar results (not shown).

In the AMOVA analysis of the entire AFLP dataset, the largest proportions of the variation were found among species (45%) and within populations (47%; Table 3). In the AMOVA analyses of each species, 73%–91% of the variation was found within populations, and 17% and 8% of the variation was found between the geographic subgroups inferred from the PCO/STRUCTURE analyses in *N. tetragona* and *N. candida*, respectively (Table 3).

In the assignment tests, performed to test a hypothesis of hybrid origin of *N. candida*, an AFLP phenotype was assigned to a group if the likelihood was 100 times higher for belonging to this group than to any other group. *Nymphaea tetragona* was divided into two subgroups (Northern Urals and Far East). Most (61) plants of *N. candida* were assigned to the simulated F1 hybrid *N. alba* × Northern Urals *tetragona*; the remaining eight *N. candida* plants were not assigned.

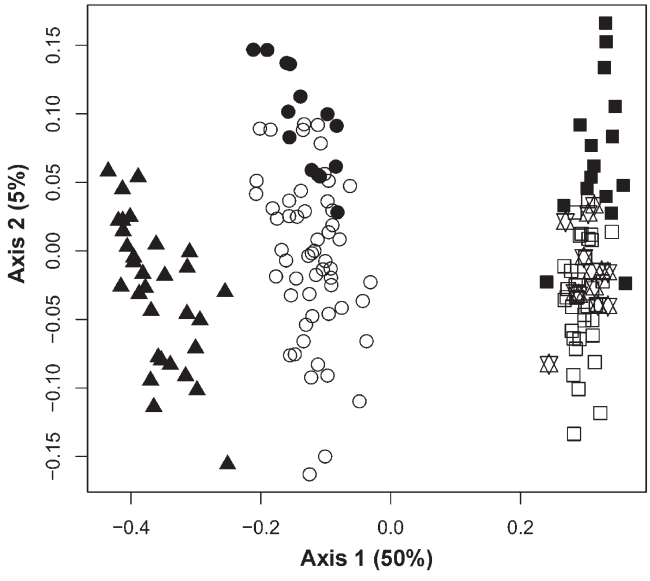

Fig. 2. Principal coordinate analysis (PCO) of individual AFLP multilocus phenotypes detected in *Nymphaea* based on Dice similarity (165 plants). Symbols indicate species and phylogeographical subgroups (AFLP data were not available for *N. tetragona* from the Murmansk Region).

cpDNA sequences. — The concatenated *petL-psbE* and *rpl32-trnL* matrix consisted of 1992 aligned positions for 43 ingroup terminals and one outgroup species, and included 769 parsimony informative characters (17 in *N.* subg. *Nymphaea*). The parsimony analysis resulted in a single most parsimonious tree of 822 steps with consistency index (CI) = 0.93 and retention index (RI) = 0.98 excluding uninformative characters (Fig. 4). Monophyly of *N.* subg. *Nymphaea* was strongly supported (100%). The Eurasian species formed a moderately (73%) supported clade with two subclades, the Tetragona subclade (100%) and the Alba subclade (89%). The Tetragona subclade contained all plants of *N. tetragona* (including *N. wenzelii*) and the Karelian *N. candida*. The Alba subclade contained all plants of *N. alba* and the remaining plants of *N. candida*. There was no sequence variation within the subclades, except that *N. tetragona* from the Japan Sea coast (pop. 740) differed by one substitution.

ITS sequences. — The ITS alignment was 732 bp long and included 38 ingroup terminals and one outgroup species, and contained 381 parsimony informative characters (79 in *N.* subg. *Nymphaea*). Polymorphic sites were observed in a few cases (1–5 sites in seven plants: populations 217, 713, 720, 732, 737, 738, 805; cf. Table 1). The parsimony analysis resulted in six most parsimonious trees with congruent topologies of 664 steps each with CI = 0.83 and RI = 0.95 excluding uninformative characters (Fig. 5). Monophyly of *N.* subg. *Nymphaea* was strongly supported (100%). The relationships between the Eurasian and the American members of the subgenus were poorly resolved, but the Eurasian members formed two strongly supported (100%) subclades. The Tetragona subclade consisted of all plants of *N. tetragona* (including *N. wenzelii*), and the Alba subclade consisted of all plants of *N. alba* and *N. candida*. There was little variation within the subclades. The haplotypes of the two *N. alba* populations (217, 805) from the Caspian Sea coast differed by one substitution from the common *N. alba* haplotype. One plant of *N. candida* (population 720) had a polymorphic site consisting of one nucleotide from the Alba subclade and one from the Tetragona subclade.

Nuclear DNA contents. — A total of 160 plants were subjected to FCM estimation (Table 1). The mean relative fluorescence intensities ± SD and CV (ranges of variation are given in parentheses) were as follows: *Nymphaea tetragona* (N = 68) 0.811 ± 0.015, 11.0% (0.776–0.862); *N. alba* (N = 27) 1.819 ± 0.046, 12.5% (1.693–1.914) and *N. candida* (N = 65) 2.533 ± 0.068, 9.7% (2.407–2.641). The taxa were thus clearly separated (differences among species significant at $P < 0.05$ in Tukey's honestly significant difference test), and the mean value for *N. candida* was almost identical to the sum of the values for *N. alba* and *N. tetragona*. The plants tentatively referred to *N. wenzelii* ([N = 15] 0.800 ± 0.013, 5.8% [0.776–0.821]) fell within the range of *N. tetragona*.

■ DISCUSSION

Our results clearly show the importance of the relatively recent connections between the presently discontinuous aquatic habitats across Eurasia for the evolution of aquatic plants. As discussed further below, these connections promoted vast and rapid expansions of divergent lineages of the Eurasian clade of

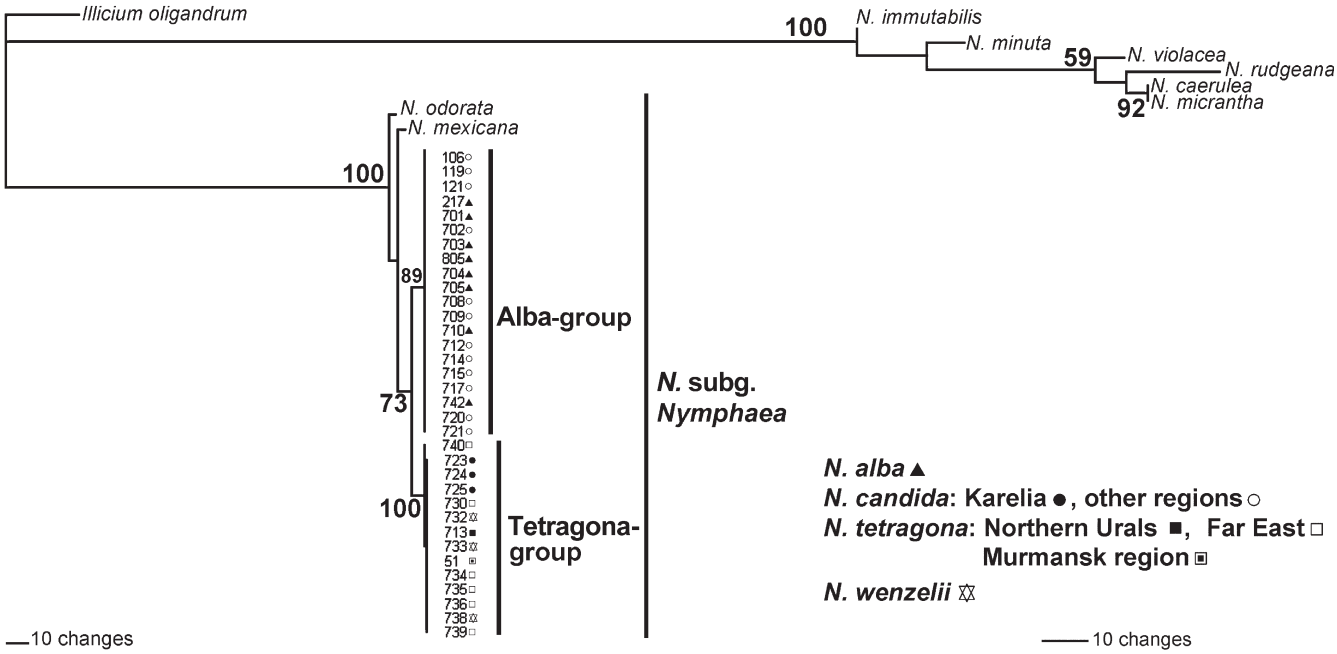

Fig. 4. The single most parsimonious tree recovered in the analysis of the concatenated *rpl32-trnL* and *petL-psbE* cpDNA sequences. Numbers at nodes indicate bootstrap values. Branch lengths are proportional to number of changes. Symbols indicate species and phylogeographical subgroups (cf. Fig. 1).

white water-lilies, low diversity within lineages, and repeated formation of evolutionary novelty via hybridization and polyploidization. Similar phylogeographic patterns, namely fast range expansions and abundant secondary contacts between different lineages during the Pleistocene, have also been shown for a number of freshwater fish species (reviewed in Gum & al., 2005) and freshwater crustaceans (e.g., Verovnik & al., 2005). Thus, the changes in the paleohydrological system connectivity probably had a strong universal impact on the evolutionary dynamics of different types of freshwater organisms.

Taxonomical implications. — In line with taxonomic treatments based on morphology (Komarov, 1970; Volkova & Shipunov, 2008), our results support that the Eurasian clade of white water-lilies consists of three species, namely *N. alba*, *N. candida* and *N. tetragona*. The high proportion of AFLP variation we observed within populations in all three species is consistent with a predominant outcrossing mating system, as suggested by Heslop-Harrison (1955). The multivariate analysis of the AFLP data, the marker-sharing patterns and the assignment test suggest that *N. candida* originated by hybridization in the wide overlap zone between the mainly European *N. alba* lineage and the mainly Asian *N. tetragona* lineage. The intermediate and highly variable morphology of *N. candida* (Volkova & Shipunov, 2008; P. Volkova, unpub. data) and the present range of *N. candida* (Komarov, 1970; see also Fig. 1) are in agreement with this conclusion. Interestingly, our results thus support the old and almost forgotten hypothesis of hybrid origin of *N. candida* suggested by Korzhinsky (1892). In addition, our estimates of relative nuclear DNA contents strongly suggest that the hybridization was associated with polyploidization, although we cannot provide definite evidence for this without chromosome counts.

According to the molecular as well as flow cytometric data, the plants tentatively referred to "*N. wenzelii*" clearly belong to the *N. tetragona* lineage, only differing by one unique polymorphic site in the ITS sequences. Thus, our data support that these plants should be included in *N. tetragona*, possibly treated as a separate variety characterized by being conspicuously large-grown (*N. tetragona* Georgi var. *wenzelii* (Maack) Vorosch; cf. Skvortsov, 1985).

Interspecific hybridization. — According to the AFLP data, the *N. candida* plants are genetically more similar to *N. alba* than to *N. tetragona*. This can be explained by considerable difference in the DNA content of the two progenitor lineages. The relative DNA content (expressed as relative fluorescence intensity) of *N. alba* is more than twice that of *N. tetragona*.

The cpDNA data suggest that the *N. alba* lineage served as the chloroplast donor for all the examined *N. candida* plants but the Karelian ones. Thus, the Karelian populations, which inhabit a previously glaciated area, probably originated from an independent *N. alba* × *tetragona* polyploidization event, in which the *N. tetragona* lineage served as the chloroplast donor. This finding adds to the many examples accumulated over the last two decades, showing that multiple origins of polyploids are a common phenomenon in plants (Soltis & Soltis, 1990; Brochmann & al., 2004). Alternatively, the Karelian plants might have been interpreted as *N. candida* × *tetragona* backcrosses, but this is not likely because they have nuclear DNA contents similar to those of the other *N. candida* plants (Table 1). The fact that the Karelian *N. candida* could be separated from the other populations of this species based on AFLP markers, which are mainly nuclear, also supports the hypothesis of independent polyploid origins rather than a single origin

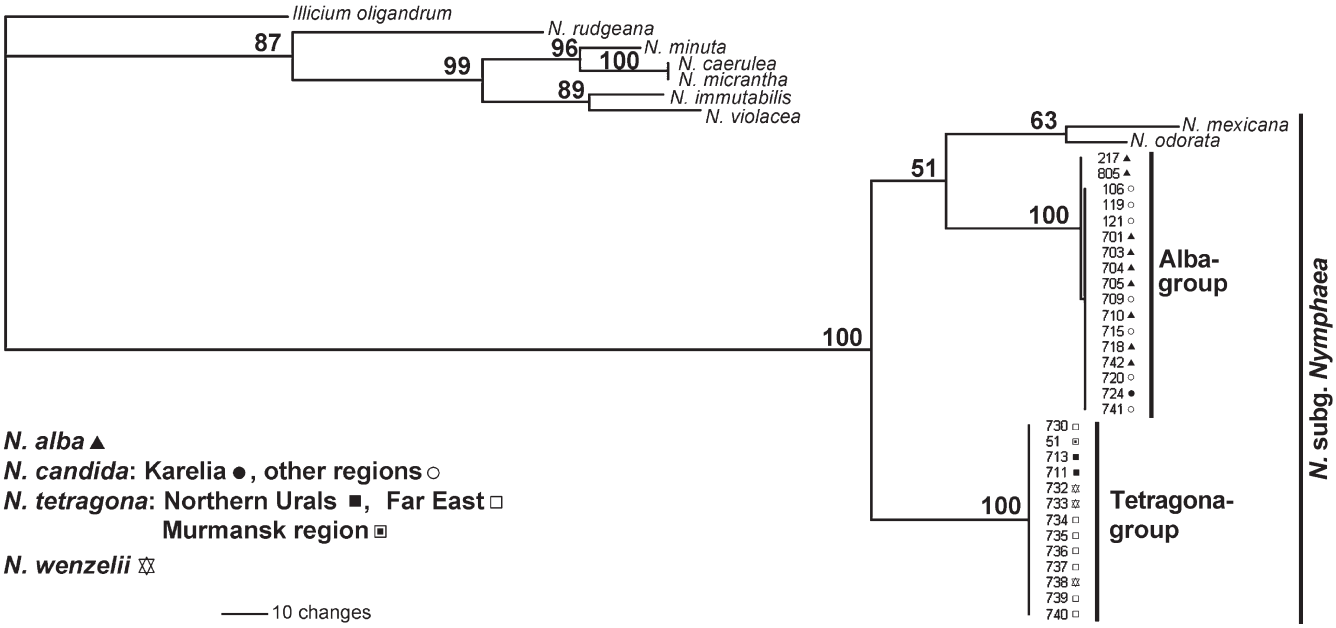

Fig. 5. One of the six most parsimonious trees recovered in the analysis of the ITS sequences. Numbers at nodes indicate bootstrap values. Branch lengths are proportional to number of changes. Symbols indicate species and phylogeographical subgroups (cf. Fig. 1).

followed by later hybridization and chloroplast capture. We therefore conclude that it is most likely that *N. candida* originated first, and at least once, in the east, from where it expanded westwards over vast areas, and that it later originated once more in the previously glaciated areas in the northwest, from where it not yet has expanded (see also below). In general, our data do not support existence of backcrosses between *N. candida* and the parent species. Some evidence of *N. alba* × *candida* backcrosses is available (Werner & Hellwig, 2006), although the taxonomical status of these plants is needed to be evaluated as that study did not include any *N. tetragona* plants.

The ITS sequences of both Karelian and non-Karelian *N. candida* were virtually identical to those of *N. alba*, and highly divergent from those of *N. tetragona*. This seems to provide one more example of concerted evolution in nuclear rDNA in polyploids in the direction of one progenitor lineage (Brochmann & al., 1996; Alvarez & Wendel, 2003). The position of *N. odorata* and *N. mexicana* within the temperate *Nymphaea* species on the ITS tree (Fig. 5) is not congruent with our (Fig. 4) and previously published cpDNA phylogenies of the genus (Borsch & al., 2007), but the position of these species in our ITS tree was poorly supported. This incongruence can be explained by a number of molecular genetic processes impacting ITS sequences, especially in polyploids. These phenomena create a network of paralogous ITS sequence relationships that confound accurate phylogenetic reconstruction (Alvarez & Wendel, 2003).

Phylogeography. — Although the two progenitor lineages of *N. candida*, *N. alba* and *N. tetragona*, occupy vast geographic ranges and are highly divergent in both cpDNA sequences (ten sites in concatenated *petL-psbE* and *rpl32-trnL*) and in ITS sequences (60 sites), they showed a striking lack of intralineage sequence variation (Figs. 4–5). This finding is especially intriguing since the earliest fossils of *N.* subg. *Nymphaea* in Europe (Romania) are dated to the Oligocene and referred to *N. alba* (Muller, 1981), and molecular dating suggests that this species appeared not later than the Middle Miocene (Löhne & al., 2008). Thus, enough time should have elapsed for accumulating some mutations in nuclear and chloroplast DNA. Two subgroups of *N. tetragona* (Northern Urals and the Russian Far East) were identified in the AFLP analysis, but even these subgroups have identical ITS and chloroplast sequences. We suggest that this pattern results from bottlenecks during the Pleistocene climatic shifts, followed by rapid and extensive range expansion.

During the Pleistocene glaciations the vast areas of unglaciated Eurasia were covered with tundra and cold steppe (Hewitt, 1996), unsuitable (too cold and/or too dry) habitats for the white water-lilies. Thus, the Eurasian *Nymphaea* species probably survived during the Ice Ages under milder and more humid climates in the Mediterranean and Southern Asia, rapidly expanding to their present range after the warming. The existence of two AFLP subgroups in *N. tetragona* may reflect that this species survived in at least two separate refugia during the last glaciation, which may have been located to the west and to the east of the glaciated Himalayas (cf., Hewitt, 2000). The lack of geographic structuring of the AFLP variation within *N. alba* may suggest that the populations we analyzed of this species (Fig. 1) are derived from a single refugium, but additional refugial lineages may be identified in other parts of its distribution area. Our data nevertheless point to rapid expansion as well as multiple hybridization events between the *N. alba* and the *N. tetragona* lineages, facilitated by the extensive earlier connections between the presently isolated drainages. These connections were caused by latitudinal systems of periglacial meltwaters, huge ice-dammed lakes, and diversion of the north-flowing rivers across the entire unglaciated territory of Northern Eurasia (Arkhipov & al., 1995; Grosswald & Hughes, 2002). Moreover, the last glaciations consisted of more than one ice advance, suggesting repeated changes of drainage conditions during each glaciation (Arkhipov & al., 1995).

The lack of unique AFLP markers in *N. candida*, its sharing of many markers with either *N. alba* or *N. tetragona*, and the assignment of most plants to the simulated F1 hybrid, suggest that too little time has elapsed after the origin of this hybrid species to accumulate new mutations at AFLP loci. Thus, it probably originated very recently, for example in the Early Weichselian. In that period the huge Trans-Siberian drainage system still existed, even according to the "minimum glaciation" reconstruction of the Siberian part of the Arctic Ice Sheet (Arkhipov & al., 1995; Svendsen & al., 2004), thus facilitating contact between the mainly European *N. alba* lineage and the mainly Asian *N. tetragona* lineage. However, the independent polyploidization event leading to the Karelian *N. candida* probably occurred even more recently, possibly after the Last Glacial Maximum, following expansion of the progenitor lineages into the previously glaciated northwestern areas.

■ ACKNOWLEDGEMENTS

We thank G. Larsson, J.A. Dragon, M. Kozhin, R.A. Murtozaliev, and the International Waterlily and Water Gardening Society (in particular C.B. Hellquist) for providing material. Part of the material was sampled in the Russian nature reserves "Kostomukshskij" (special thanks to S.V. Tarkhov), "Astrakhanskij" (special thanks to A.K. Gorbunov), "Il'menskij" (special thanks to E.I. Vejsberg), "Khinganskij" (special thanks to S.G. Kudrin), "Dal'nevostochnyj Morskoj" (special thanks to A.N. Malyutin), in the national park "Kondinskije Ozjora" (special thanks to A.L. Vasina), in the nature-historical reserve "Arkaim" (special thanks to E. Chibiljov), and at the biological station "Lake Moldino" of Moscow South-West High School. We are grateful to V. Mirré for training PV in the molecular laboratory, to C. Ritz for sharing her knowledge about optimization of the AFLP protocol, to G. Gusarova for advice on the parsimony analyses, to J. Rauchová for help with FCM analyses and to J. Suda and G. Gusarova for comments on earlier drafts of the manuscript. This study was supported by a grant to PV from the Research Council of Norway under the Norwegian Government Scholarships programme. The laboratory costs were covered by the guest researcher programme at the National Centre for Biosystematics and by grants from the Academy of Sciences of the Czech Republic (AV0Z60050516), the Czech Science Foundation (206/08/H049) and the Ministry of Education, Youth and Sport of the Czech Republic (MSM 0021620828).

■ LITERATURE CITED

Alvarez, I. & Wendel, J.F. 2003. Ribosomal ITS sequences and plant phylogenetic inference. *Molec. Phylog. Evol.* 29: 417–434.

Arkhipov, S.A., Ehlers, J., Johnson, R.G. & Wright, H.E. 1995. Glacial drainages towards the Mediterranean during middle age and late Pleistocene. *Boreas* 24: 196–206.

Bernatchez, L. & Wilson, C.C. 1998. Comparative phylogeography of Nearctic and Palearctic fishes. *Molec. Ecol.* 7: 431–452.

Brochmann, C., Brysting, A.K., Alsos, I.G., Borgen, L., Grundt, H.H., Scheen A.-C. & Elven, R. 2004. Polyploidy in arctic plants. *Biol. J. Linn. Soc.* 82: 521–536.

Brochmann, C., Nilsson, T. & Gabrielsen, T.M. 1996. A classic example of postglacial allopolyploid speciation re-examined using RAPD markers and nucleotide sequences: *Saxifraga osloensis* (Saxifragaceae). *Symb. Bot. Upsal.* 31: 75–89.

Bolkhovskikh, Z., Grif, V., Matvejeva, T. & Zakhazyeva, O. 1969. [*Khromosomnye chisla tsvetkovyh rastenij*]. Leningrad: Academy of Sciences of the U.S.S.R. [In Russian.]

Bonin, A., Bellemain, E., Eidesen, P.B., Pompanon, F., Brochmann, C. & Taberlet, P. 2004. How to track and assess genotyping errors in population genetics studies. *Molec. Ecol.* 13: 3261–3273.

Borsch, T., Hilu, K.W., Wiersema, J.H., Löhne, C., Barthlott, W. & Wilde, V. 2007. Phylogeny of *Nymphaea* (Nymphaeaceae): Evidence from substitutions and microstructural changes in the chloroplast *trnT-trnF* region. *Int. J. Pl. Sci.* 168: 639–671.

Borsch, T., Löhne, C. & Wiersema, J.H. 2008. Phylogeny and evolutionary patterns in Nymphaeales: Integrating genes, genomes and morphology. *Taxon* 57: 1052–1081.

Borsch, T. & Soltis, P. 2008. Nymphaeales: The first globally diverse clade? *Taxon* 57: 1051.

Conard, H.S. 1905. The water-lilies: A monograph of the genus *Nymphaea*. *Publ. Carnegie Inst. Wash.* 4: 1–279.

Dieffenbacher-Krall, A.C. & Jacobson, G.L. 2001. Post-glacial changes in the geographic ranges of certain aquatic vascular plants in North America. *Biol. Environm.* 101: 79–84.

Doležel, J., Sgorbati, S. & Lucretti, S. 1992. Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. *Physiol. Pl.* 85: 625–631.

Duchesne, P. & Bernatches, L. 2002. AFLPOP: A computer program for simulated and real population allocation using AFLP markers. *Molec. Ecol. Notes* 2: 380–383.

Ehrich, D. 2006. AFLPdat: A collection of R functions for convenient handling of AFLP data. *Molec. Ecol. Notes* 6: 603–604.

Excoffier L., Laval G. & Schneider S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol. Bioinformatics Online* 1: 47–50.

Excoffier, L., Smose, P.E. & Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* 131: 479–491.

Falush, D., Stephens, M. & Pritchard, J. 2007. Inference of population structure using multilocus genotype data: Dominant markers and null alleles. *Molec. Ecol. Notes* 7: 574–578.

Gaudeul, M., Taberlet, P. & Till-Bottraud, I. 2000. Genetic diversity in an endangered alpine plant, *Eryngium alpinum* L. (Apiaceae), inferred from amplified length polymorphism markers. *Molec. Ecol.* 9: 1625–1637.

Grosswald, M.G. & Hughes, T.J. 2002. The Russian component of an Arctic Ice Sheet during the Last Glacial Maximum. *Quatern. Sci. Rev.* 21: 121–146.

Gum, B., Gross, R. & Kuehn, R. 2005. Mitochondrial and nuclear DNA phylogeography of European grayling (*Thymallus thymallus*): Evidence for secondary contact zones in Central Europe. *Molec. Ecol.* 14: 1707–1725.

Gupta, P.P. 1980. Cytogenetics of aquatic ornamentals: VI, Evolutionary trends and relationships in the genus *Nymphaea*. *Cytologia* 45: 307–314.

Hahsler, M., Gruen, B. & Hornik, K. 2005. Arules: A computation environment for mining association rules and frequent item sets. *J. Statist. Software* 14(15). Available at http://www.jstatsoft.org/v14/i15/.

Heslop-Harrisson, J. 1955. *Nymphaea*. *J. Ecol.* 43: 719–734.

Hewitt, G.M. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. *Biol. J. Linn. Soc.* 58: 247–276.

Hewitt, G.M. 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405: 907–913.

Iversen, J. 1954. The late-glacial flora of Denmark and its relation to climate and soil. *Danmarks Geol. Undersøg.* 80: 87–119.

Kapralov, M.V., Gabrielsen, T.M., Sarapultsev, I.E. & Brochmann, C. 2006. Genetic enrichment of the arctic clonal plant *Saxifraga cernua* at its southern periphery via the alpine sexual *Saxifraga sibirica*. *Molec. Ecol.* 15: 3401–3411.

Komarov, V.L. 1970. *Nymphaea* L. Pp. 8–12 in: Komarov, V.L. (ed.), *Flora of the USSR*, vol. 7. Jerusalem: Israel program for Scientific Translations.

Korzhinsky, S.I. 1892. [*Flora vostoka evropejskoj Rossii v ejo sistematicheskom I geograficheskom aspekte*]. Tomsk: Tipolitografiya V.V. Mikhajlova i P.I. Makushkina. [In Russian.]

Krupkina, L.I. 2001. Nymphaeceae Salisb. Pp. 25–30 in: Tsvelev, N.N. (ed.), [*Flora vostochnoj Evropy*], vol. 10. Saint Petersburg: Nauka. [In Russian.]

Les, D.H. & Philbrick, C.T. 1993. Studies of hybridization and chromosome number variation in aquatic angiosperms: Evolutionary implications. *Aquatic Bot.* 44: 181–228.

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

Löhne, C., Yoo, M.-J., Borsch, T., Wiersema, J., Wilde, V., Bell, C.D., Barthlott, W., Soltis, D.E. & Soltis, P.S. 2008. Biogeography of Nymphaeales: Extant patterns and historical events. *Taxon* 57: 1123–1146.

Maack, R.K. 1859. [*Puteshestvie na Amur*]. Saint Petersburg: Izdatelstvo S.F. Solovjova. [In Russian]

Möller, M. & Cronk, Q.C.B. 1997. Origin and relationships of *Saintpaulia* (Gesneriaceae) based on ribosomal DNA internal transcribed spacer (ITS) sequences. *Amer. J. Bot.* 84: 956–965.

Muller, J. 1981. Fossil pollen records of extant angiosperms. *Bot. Rev.* 47: 1–142.

Otto, F. 1990. DAPI staining of fixed cells for high-resolution flow cytometry of nuclear DNA. Pp. 105–110 in: Crissman, H.A. & Darzynkiewicz, Z. (eds.), *Methods in cell biology*, vol. 33. New York: Academic Press.

Pritchard, J.K., Stephens, M. & Donnely, P.J. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

R Development Core Team. 2004. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Rosenberg, N.A., Pritchard, J.K., Weber, J.L., Cann, H.M., Kidd, K.K., Zhivotovsky, L.A. & Feldman, M.V. 2002. Genetic structure of human populations. *Science* 298: 2381–2385.

Schönswetter, P., Suda, J., Popp, M., Weiss-Schneeweiss, H. & Brochmann, C. 2007. Circumpolar phylogeography of *Juncus biglumis* (Juncaceae) inferred from AFLP fingerprints, cpDNA sequences, nuclear DNA content and chromosome numbers. *Molec. Phylog. Evol.* 42: 92–103.

Shaw, J., Lickey, E., Beck, J.T., Farmer, S.B., Liu, W., Miller, J., Siripun, K.C., Winder, C.T., Schilling, E.E. & Small, R.L. 2005. The tortoise and the hare II: Relative utility of 21 noncoding chloroplast sequences for phylogenetic analysis. *Amer. J. Bot.* 92: 142–166.

Shaw, J., Lickey, E.B., Schilling, E.E. & Small, R.L. 2007.

Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III. *Amer. J. Bot.* 94: 275–288.

Skvortsov, A.K. 1985. *Floristicheskie issledovaniya v razlichnyh regionakh SSSR*. Moscow: Mysl'. [In Russian.]

Soltis, D.E. & Soltis, P.S. 1990. Chloroplast DNA and nuclear rDNA variation: Insights into autopolyploid and allopolyploid evolution. Pp. 97–117 in: Kawano, S. (ed.), *Biological approaches and evolutionary trends in plants*. London: Academic Press.

Svendsen, J.I., Alexanderson, H., Astakhov, V.I., Demidov, I., Dowdeswell, J.A., Funder, S., Gataulin, V., Henriksen, M., Hjort, C., Houmark-Nielsen, M., Hubberten, H., Ingolfsson, O., Jakobsson, M., Kjær, K.H., Larsen, E., Lokrantz, H., Lunkka, J.P., Lysså, A., Mangerud, J., Matiouchkov, A., Murray, E., Möller, P., Niessen, F., Nikolskaya, O., Polyak, L., Saarnisto, M., Siegert, C., Siegert, M.J., Spielhagen, R.F. & Stein, R. 2004. Late Quaternary ice sheet history of northern Eurasia. *Quatern. Sci. Rev.* 23: 1229–1271.

Swofford, D.L. 2002. *PAUP*: Phylogenetic analysis using parsimony (*and other methods)*, version 4.10. Sunderland, Massachusetts: Sinauer.

Uotila, P. 2001. *Nymphaea* L. Pp. 216–221 in: Jonsell, B. (ed.), *Flora Nordica*, vol. 2. Stockholm: Bergius Foundation.

Verovnik, R., Sket, B. & Trontelj, P. 2005. The colonization of Europe by the freshwater crustacean *Asellus aquaticus* (Crustacea: Isopoda) proceeded from ancient refugia and was directed by habitat connectivity. *Molec. Ecol.* 14: 4355–4369.

Volkova, P.A. & Shipunov, A.B. 2008. Morphological variation of *Nymphaea* (Nymphaeaceae Juss.) in European Russia. *Nord. J. Bot.* 25: 329–338.

Ward, R.D., Woodwark, M. & Skibinski, D.O.F. 1994. A comparison of genetic diversity levels in marine, freshwater, and anadromous fishes. *J. Fish Biol.* 44: 213–232.

Werner, K. & Hellwig, F. 2006. Hybridization between *Nymphaea alba* and *Nymphaea candida* investigated by AFLP fingerprinting and morphological data. P. 227 in: *Abstracts of 17th Symposium Biodiversity and Evolutionary Biology*, 24–28 September 2006, Bonn.

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