Natural England Commissioned Report NECR245

Evaluation of genetic diversity and admixture in the only English population of *Nuphar pumila*

First published 12 December 2017

www.gov.uk/natural-england

NATURAL ENGLAND

Foreword

Natural England commission a range of reports from external contractors to provide evidence and advice to assist us in delivering our duties. The views in this report are those of the authors and do not necessarily represent those of Natural England.

Background

Least water-lily *Nuphar pumila* is known from over 100 sites in Scotland, but only from Cole Mere in England, although it formerly occurred in other meres in Shropshire.

Natural England commissioned this study of *Nuphar pumila,* under the Species Recovery Programme to gather information to improve its conservation status at Cole Mere SSSI, Shropshire and to develop and implement a Species Recovery Plan with the aim to maintain a self-sustaining population of N. pumila in England.

The work was commissioned in three parts:

- Commissioned report 243, to review the known ecology of *Nuphar pumila* and the population status at Cole Mere.
- Commissioned report 244 to develop a seed and rhizome propagation protocol.
- This report to assess the level of genetic variation present in English *N. pumila*; test whether English and Scottish *N. pumila* populations are genetically distinct from each other; and confirm the hybrid nature of samples identified as *N.* ×*spenneriana*.

Natural England and others will use the findings to develop a plan to conserve England's only population of least water-lily, and in particular to manage the trees and the margins around Cole Mere.

This report should be cited as: GARGIULO, R., LANSDOWN, R.V. and FAY, M.F., 2017. *Evaluation of genetic diversity and admixture in the only English population of Nuphar pumila*. Natural England Commissioned Reports, Number 245. York.

Natural England Project Officer – Mags Cousins, **mags.cousins@naturalengland.org.uk**

Contractor – Roberta Gargiulo, Royal Botanic Gardens, Kew

Key words - Nuphar pumila, Least Water-lily, Nuphar xspenneriana, Nuphar lutea, genetics, DNA, allele, introgression, genotyping, microsatellite, heterozygous, polymerase, principal coordinates analysis, genetic drift

Further information
This report can be downloaded from the Natural England Access to Evidence Catalogue: **http://publications.naturalengland.org.uk/** . For information on Natural England publications contact the Natural England Enquiry Service on 0300 060 3900.

ISBN 978-1-78354-457-8

Royal Botanic Gardens

Kew

Evaluation of genetic diversity and admixture in the only English population of *Nuphar pumila*

Report for Natural England

Roberta Gargiulo, Richard V. Lansdown and Michael F. Fay

Evaluation of genetic diversity and admixture in the only English population of *Nuphar pumila.*

Report for Natural England

Roberta Gargiulo[1], Richard V. Lansdown[2] and Michael F. Fay[1]

February 2017

1 Conservation Genetics, Royal Botanic Gardens (RBG) Kew
2 IUCN SSC Freshwater Plant Specialist Group

Background

RBG Kew's Conservation Genetics team was commissioned by Mags Cousins, Natural England to undertake this study as Part 3 of the Species Recovery Project: Develop a plan to conserve England's only population of *Nuphar pumila*, Least Water-lily (Lansdown, 2017 and Peach et al, 2017).

The objectives of this genetics study were to:

I. assess the level of genetic variation present in English *N. pumila*;
II. test whether English and Scottish *N. pumila* populations are genetically distinct from each other; and
III. confirm the hybrid nature of samples identified as *N.* ×*spenneriana.*

Genetics work was carried out by staff in RBG Kew's Conservation Genetics team, with sampling and field identifications by Richard Lansdown, Chair of the IUCN SSC Freshwater Plant Specialist Group. The project was managed by Ted Chapman, UK Native Seed Hub Coordinator.

About the UK Native Seed Hub

The UK Native Seed Hub seeks to increase the quality and diversity of native plant materials available for conservation and habitat restoration in the UK, mobilising the collections, facilities, technical knowledge and scientific expertise of RBG Kew.

We provide specialist plants and seed, technical services and research to help seed producers and conservation practitioners overcome constraints to sourcing, producing and using native plant materials in the UK.

For further information, contact Ted Chapman at t.chapman@kew.org, 01444 894192.

Introduction

Nuphar pumila (Timm) DC. is regarded as a climate relict (Heslop-Harrison, 1955), surviving in central Europe, in the Alpine arc and in southern Europe with some scattered populations (Kozlowski and Eggenberg, 2005; Lozano *et al.*, 2008; Meusel *et al.*, 1965). Genetic diversity of the populations distributed in the Alpine mountains has been recently investigated, with special emphasis on the relationships with the lowland species *N. lutea* (L.) Sm. and their hybrids. *N. pumila* turned out to be seriously threatened by genetic swamping, due to recent and ongoing introgression with the more widespread and generalist *N. lutea* (Arrigo *et al.*, 2016). In this process, in which climate change and habitat transformation have played an important role, it is likely that *N. lutea* will expand and only minor genetic traces of *N. pumila* will persist. Conservation measures are thus strongly recommended, especially for the last pure stands of *N. pumila*.

The aim of the present report and the underlying analyses was to assess the genetics of the *N. pumila* population in Cole Mere SSSI, Shropshire, England, by comparison with some populations from Scotland and some putative samples of *N.* ×*spenneriana* Gaudin. We employed microsatellite markers which have been successful in detecting hybrid individuals and in highlighting genetic diversity and differentiation in previous studies (Arrigo *et al.*, 2016).

Methods

Sampling was carried out from six British populations; two in England, Cole Mere and Betton Mere, the remainder in Scotland. At some sites identification was not straightforward and identifications listed are those that were used in the field (Table 1). Total genomic DNA was extracted from leaves dried in silica-gel (approx. 15 mg) using the CTAB method (Doyle and Doyle, 1987) and purified using the QUIAquick PCR purification kit (QUIAGEN).

Eight nuclear SSR markers (Table 2; Ouborg *et al.*, 2000; Yokogawa *et al*., 2012) were employed for genotyping. Polymerase chain reactions were carried out in a final volume of 10 µL, with 10 ng genomic DNA, 5 pmol reverse and FAM- or JOE-labeled forward primers (Eurofins Genomics), 6 µL 2× DreamTaq PCR Master Mix (Thermo Scientific), 1 µL 0.4% (w/v) bovine serum albumin and sterile deionized water, in a GeneAmp PCR System 9700 (Applied Biosystems). The thermal profile followed Arrigo *et al.*, 2016: Amplification products (1 µL) were combined with 10 µL of HiDi™ formamide (Applied Biosystems) and 0.15 µL GeneScan 500 ROX Size Standard (Applied Biosystems). Capillary electrophoresis was conducted on an ABI3730 DNA Analyzer (Applied Biosystems). Microsatellite scoring was carried out in GeneMapper v5 (Applied Biosystems).

Presence of genotyping errors (stuttering or large allele dropout) and null alleles was assessed with MicroChecker v2.2.3 (van Oosterhout *et al*., 2004) using 1000 randomisations.

Allele frequencies, observed and unbiased expected heterozygosities (H_O, uH_E, respectively) and percentage of polymorphic loci (%*P*) were computed with GenAlEx v6.5 (Peakall and Smouse, 2006). Allelic richness (A_R) and inbreeding coefficient (F_{IS}) were computed in FSTAT v2.9.3.2 (Goudet, 2002) and GenePop v4.5.1 (Raymond and Rousset, 1995), respectively.

Principal coordinates analysis (PCoA) based on the covariance matrix as implemented in GenAlEx v6.5 was carried out in order to evaluate and represent the percentage of variation. The number of genetic clusters was inferred with the Bayesian method implemented in Structure v2.3 (Pritchard *et al*., 2000). Admixture model, correlated allele frequencies and no prior information about populations were used. Ten iterations per number of cluster (*K*) were run, with *K* ranging from one to ten (the number of populations plus three; Evanno *et al*., 2005). Each run included 50,000 burn-in steps and 200,000 Markov chain Monte Carlo (MCMC) steps. The most adequate value for the number of clusters (*K*) was determined with the method suggested by Evanno *et al*. (2005), by exploring log-likelihood values and the second-order rate of change in the values of *K* (Δ*K*), with the online tool Structure Harvester (Earl *et al*., 2012). Additional analyses were performed on reduced datasets (removing *N. lutea* accessions) to check for undetected structure.

Table 1 Details of the population sampling

Site	Code	Field identification	Number of samples	Notes
Betton Mere	**BM**	*N.* ×*spenneriana*	6	Samples N27-N31
Cole Mere	**CML**	*N. lutea*	6	Samples N17-N21
Cole Mere	**CMP**	*N. pumila*	5	Samples N12-N16
Kirrieroch	**K**	*N. lutea*	6	Samples N6-N11 Possibly one sample of *N.* ×*spenneriana* (N10) and one indet. sample (N11)
Little Rogart Loch	**LRL**	*N. pumila*	5	Samples N32-N36
Loch Uvie	**LU**	*N.* ×*spenneriana*	5	Samples N1-N5 Some samples might be *N. pumila*
Rannoch Moor	**RM**	*N. pumila*	5	Samples N22-N26

Table 2 Nuclear SSR markers for *Nuphar* ("NL" from Ouborg *et al.*, 2000; "Nsub" from Yokogawa *et al.*, 2012).

Primer	Sequence (5'-3')
NLGA2	(FAM) F: CTTTAGGAGGGTCTTTAGCC
	R: CCAATCTCTAGTAGGAGGAGC
NLGA3	(JOE) F: GTTGTAACGTAAATGCCGTCC
	R: CTTGCCGATGAAACCCAT
NLGA5	(FAM) F: CCCGCCATATCTGATGAC
	R: AAGTGGAGGGGACGAAAG
NLGA7	(FAM) F: ATTTATTCCCAGCACTTTGG
	R: CTTGACATGATTTCTCTGAACC
NLCA1	(JOE) F: CTCAGAAACGAGGCTCTATG
	R: TTTGGTTGGAAGACAAGAAG
NLTG/GA1	(JOE) F: AAGCAGCAGCAAAATTTGTA
	R: TGTGCAAGTTACCTGTTTCC
Nsub033	(FAM) F: ACACACACACACTCTCTCTCTC
	R: ACTTGCAAAGATCCTCTCAGAT
Nsub176	(JOE) F: AGAGAGAGAGAGACACACACAC
	R: GGCAACAGGTCTATTAATCTCA

Results

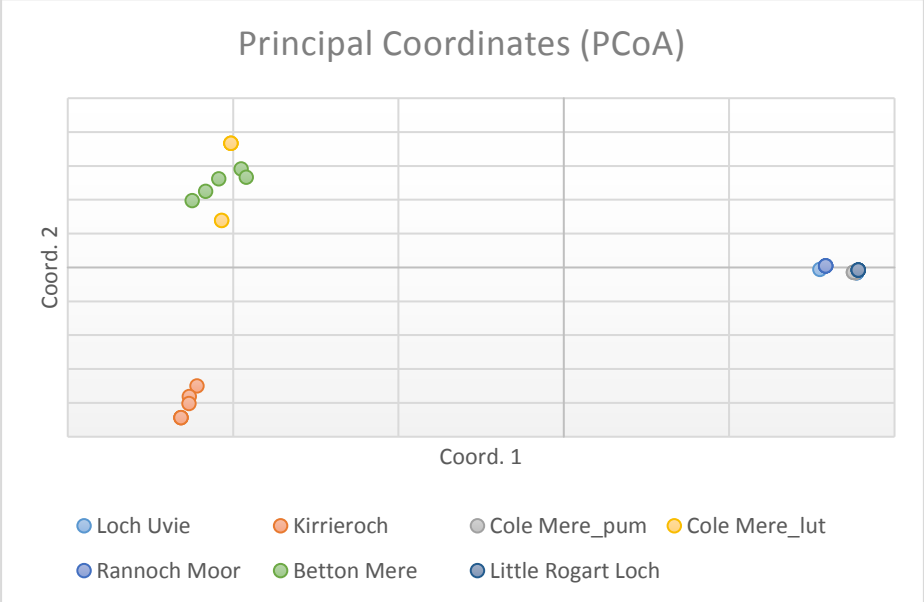
No evidence of null alleles or genotypic errors was detected in Microchecker. Average values of genetic diversity indices for each population are summarised in Table 3. The *N. ×spenneriana* population in Betton Mere and the *N. lutea* population in Cole Mere exhibit the highest values of observed heterozygosity, percentage of polymorphic loci and allelic richness. In general, *N. pumila* populations exhibit low or even null observed heterozygosity and loci completely fixed for one allele (Table 3).

The scatter plot from the PCoA analysis as implemented in GenAlEx is shown in Figure 1; *N. pumila* populations (including *N. ×spenneriana* from Loch Uvie) appear well separated from *N. lutea* populations. *N. lutea* from Kirrieroch is genetically differentiated from the cluster composed of *N. lutea* from Cole Mere and *N. ×spenneriana* from Betton Mere. Similarly, the Structure Harvester tool revealed two genetic clusters, corresponding to the *N. lutea* accessions (including *N. ×spenneriana* in Betton Mere) and the *N. pumila* accessions (including *N. ×spenneriana* from Loch Uvie) (Fig. 2A). The smaller scale analysis conducted on the latter cluster, detected three genetic groups (Fig. 2B) However, all genetic diversity indices and differentiation statistics should be taken with caution, as they are affected by the small sample size.

Table 3 Estimates of genetic diversity parameters averaged over loci for *Nuphar* populations.

Population (number of individuals)	Number of Genotypes	N_A (SE)	H_O (SE)	uH_E (SE)	*%P*	A_R	F_{IS}	Q range (*K*=3)
BM (6)	5	1.62 (0.26)	0.27 (0.11)	0.24 (0.10)	50.0	1.62	-0.143	
CML (6)	2	1.50 (0.19)	0.150 (0.098)	0.25 (0.09)	50.0	1.5	0.429	
CMP (5)	2	1.12 (0.12)	0.05 (0.05)	0.04 (0.04)	12.5	1.12	-0.143	
K (6)	4	1.37 (0.18)	0.08 (0.04)	0.08 (0.04)	37.5	1.33	-0.053	
LRL (5)	1	1.00	0	0	0	1.0	-	
LU (5)	2	1.25 (0.16)	0	0.09 (0.06)	25.0	1.25	1.0	0.007-0.841
RM (5)	2	1.00	0	0	0	1.0	-	

N_A = no. of different alleles; H_O = Observed Heterozygosity; uH_E = Unbiased Expected Heterozygosity; %P = percentage of polymorphic loci as implemented in GenAlEx. A_R = average allelic richness based on the minimum sample size (5), as implemented in FSTAT; F_{IS} inbreeding coefficient per sample over loci as implemented in GenePop (Weir and Cockerham, 1984); Q range resulting from the Structure analysis of *N. pumila* populations is shown only when highest q < 0.95 (Fig. 2B).



Percentage of variation explained by the first 3 axes

Axis	**1**	**2**	**3**
%	56.47	12.13	10.45
Cum %	56.47	68.61	79.05

Fig. 1 PCoA analysis and partition of the genetic variation.

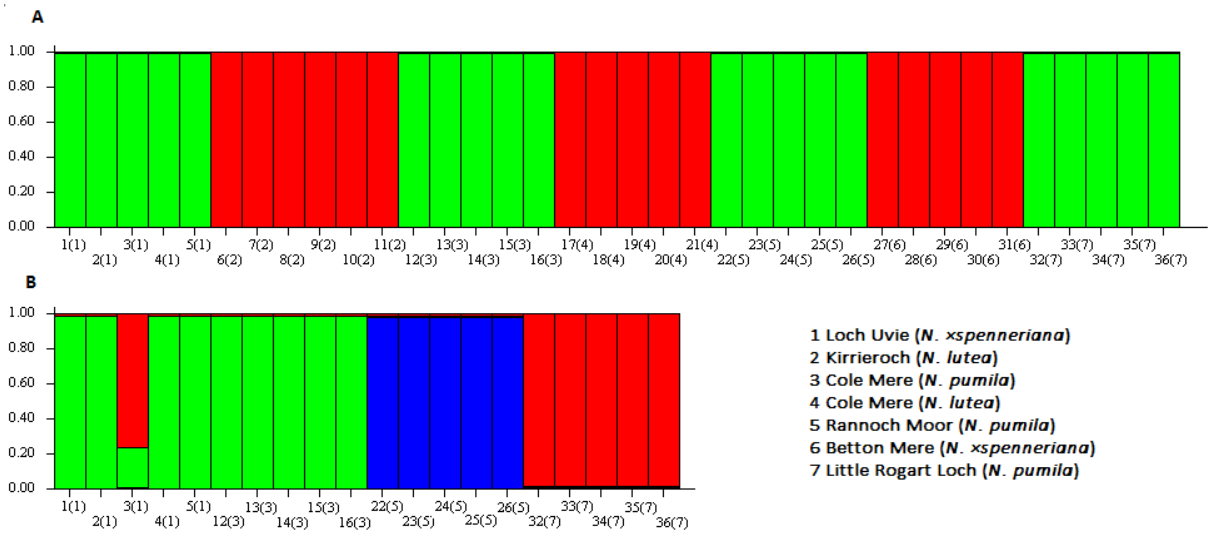


Fig. 2 Structure Bar Plots of (A) total populations with K = 2, (B) *N. pumila and N. x spenneriana* population from Loch Uvie with K = 3.

On the basis of the molecular analysis, we have updated the field identifications. The final names are presented in Table 4.

Table 4 Comparison of field identification and molecular identification.

Site	Code	Field identification	Molecular identification
Betton Mere	**BM**	*N. ×spenneriana*	*N. lutea*
Cole Mere	**CML**	*N. lutea*	*N. lutea*
Cole Mere	**CMP**	*N. pumila*	*N. pumila*
Kirrieroch	**K**	*N. lutea*	*N. lutea*
Little Rogart Loch	**LRL**	*N. pumila*	*N. pumila*
Loch Uvie	**LU**	*N. ×spenneriana*	*N. pumila*
Rannoch Moor	**RM**	*N. pumila*	*N. pumila*

Discussion and Conclusions

The present assessment provides a general overview of the genetic variation within the British populations of *Nuphar*. No first or second generation hybrids were detected with the employed markers, which have been previously used to assess genetic admixture in the European populations of *Nuphar* (Arrigo *et al.*, 2016) The putative hybrid individuals sampled in Betton Mere are identified genetically as *N. lutea*, whereas the putative hybrid population in Loch Uvie is identified genetically as *N. pumila*. The population from Betton Mere shows a fixed private (i.e., not shared) allele at one locus; this was not found in any of the populations of the putative parental species; moreover, the population is comparatively more variable (in terms of number of genotypes, heterozygosity and allele diversity). However, it was not possible to detect whether the source of this allele is actually the result of ancestral hybridisation between *N. lutea* and *N. pumila*, as it was hypothesised (Heslop-Harrison, 1953) or whether it represents variation in *N. lutea* not sampled from elsewhere in this study.

Concerning the genetic status of *N. pumila*, the Cole Mere population is not very differentiated from the other populations, except for a private allele at one locus. In general, heterozygosity and inbreeding coefficient indicate that genetic drift and allelic variant fixation is ongoing and almost complete (e.g., on Rannoch Moor). This is very likely due to complete absence of gene flow from other populations and/or to an exclusively vegetative reproduction (e.g., by fragmentation). However, all the results refer to few individuals; it would be advisable to extend the sampling not only by including all the (possible) individuals in each site but also all the British populations. In rare and threatened species, such as *N. pumila*, a total evidence approach is usually the best choice, in order to capture the existing diversity and to make effective strategic plans. This might be also useful in order to better characterise potential hybrid individuals (Levin *et al.*, 1996). Moreover, it would be interesting to compare the British diversity with continental diversity.

These results suggest that some interventions could be made in order to avoid the complete erosion of genetic diversity. Some re-introduction (and subsequent artificial crossing, if necessary) might be carried out in Cole Mere, by using individuals from Scotland. However, this should be treated with caution in case there is local adaptation separating the English and Scottish populations. To minimise the risk of disrupting any local adaptation, plants from such experimental crosses could be maintained in *ex situ* collections or introduced to bodies of water where *N. pumila* does not occur. These interventions should follow further assessments related to habitat suitability and to the effective limiting factors which have caused the decline in the last decades.

In conclusion:

- There is no evidence of hybridisation between *N. lutea* and *N. pumila* in England or Scotland, on the basis of the samples included here.
- As expected, *N. lutea* proved to be more variable genetically than *N. pumila*.

References

Arrigo N, Bétrisey S, Graf L, Bilat J, Gerber E, Kozlowski G. 2016. Hybridization as a threat in climate relict *Nuphar pumila* (Nymphaeaceae). *Biodiversity and Conservation* **25:**1863–1877.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation method for small quantities of fresh tissues. *Phytochemical Bulletin, Botanical Society of America* **19:**11–15.

Earl DA, von Holdt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4:**359–361.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology.* **14:**2611–2620.

Goudet J. 2001. FSTAT, version 2.9.3, a program to estimate and test gene diversities and fixation indices. Lausanne University, Lausanne, Switzerland.

Heslop-Harrison Y. 1955. *Nuphar* Sm. *Journal of Ecology* **43:**342–364.

Heslop-Harrison Y. 1953. *Nuphar intermedia* Ledeb., a presumed relict hybrid in Britain. *Watsonia* **3:**7-25.

Kozlowski G, Eggenberg S. 2005. Vorkommen der Kleinen Teichrose *Nuphar pumila* und des hybrids *N. x intermedia* in der Schweiz. *Botanica Helvetica* **115:**125–136.

Lansdown, RV. 2017. Development of a conservation plan for Least Water-lily (*Nuphar pumila*) in England. Natural England Commissioned Report, Number 243. York.

Levin DA, Francisco-Ortega J, Jansen RK. 1996. Hybridization and the extinction of rare plant species. *Conservation Biology*. **10:**10–16.

Lozano FD, Moreno Saiz JC, Sainz Ollero H, Herbada DG, Rivero LM. 2008. Lista roja de la flora vascular espanola. Conservation Vegetal, Madrid.

Meusel H, Mühlberg H, Fuchs HP. 1965. Nymphaeaceae. In: Hegi G, ed, Illustrierte Flora von Mitteleuropa. Band III. Teil 3, München edn. Carl Hanser, Germany.

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* **4:** 535–538.

Ouborg NJ, Goodall-Copestake WP, Epplen JT. 2000. Novel polymorphic microsatellite loci isolated from the yellow waterlily, *Nuphar lutea*. *Molecular ecology* **9.4:**497–498.

Peach, J, Davies, R, Walmisley, J, Chapman, T. 2017. An assessment of seed viability, germination and vegetative propagation requirements for *Nuphar pumila*. Natural England Commissioned Report, Number 244. York

Peakall R, Smouse PE. 2006. GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6:**288–295.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics*. **155:**945–959.

Raymond M, Rousset F. 1995. GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. *Heredity* **86:**248–249.

Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure.

Evolution **38:**1358–1370.

Yokogawa M, Shiga T, Kaneko S, Isagi Y. 2012. Development of nuclear microsatellite markers for the critically endangered freshwater macrophyte, *Nuphar submersa* (Nymphaeaceae), and cross-species amplification in six additional *Nuphar* taxa. *Conservation Genetics Resources*. **4:**295–298.