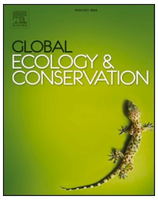
Global Ecology and Conservation 55 (2024) e03235

Contents lists available at ScienceDirect

Global Ecology and Conservation

journal homepage: www.elsevier.com/locate/gecco



Original research article

Phylogeography and population genomics of the critically endangered aquatic plant *Caldesia grandis* in China

Dong-Ying Yan [a,b], Shuai Peng [c], Samuli Lehtonen [d], Jin-Ming Chen [a,*], Zhi-Zhong Li [a,e,**]

[a] *Aquatic Plant Research Center, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China*
[b] *Research Center for Ecology and Environment of Qinghai-Tibetan Plateau, Tibet University, Lhasa 850000, China*
[c] *Plant Biodiversity Research Center, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China*
[d] *Herbarium, Biodiversity Unit, University of Turku, Turku 20014, Finland*
[e] *Collaborative Innovation Center of Recovery and Reconstruction of Degraded Ecosystem in Wanjiang Basin Co-funded By Anhui Province and Ministryof Education of the People's Republic of China, Schoolof Ecology and Environment, Anhui Normal University, Wuhu 241002, China*

ARTICLE INFO

Keywords:
Alismataceae
Conservation genetics
RAD-seq
Rare species
Refuge

ABSTRACT

Caldesia grandis, a critically endangered aquatic species, is predominantly found in the mid-low mountainous swamps of subtropical China. This study assessed the genetic diversity and population structure of the species using RAD-seq data, explored its phylogeography across the extant nine populations based on five plastid DNA (ptDNA) regions, and conducted niche modeling analysis. We found low genetic diversity ($H_E = 0.180$, $H_o = 0.222$, and $\pi = 0.197$) and genetic differentiation among populations ($Fst = 0.089$), which was likely due to genetic drift in small populations and frequent inter-population contact during the Quaternary period. Although RAD-seq analysis did not reveal a clear population structure, two distinct clades, comprising western and eastern populations, were identified using five ptDNA haplotypes. Molecular dating and niche modeling suggested that the uplift of the Luoxiao Mts may have contributed to the divergence of the eastern and western clades (ca. 1.51 Ma)during the Pleistocene, which also supports the hypothesis that the Nanling Mts acted as a refugium for *C. grandis*. Additionally, the repeated glacial periods of the Quaternary, accompanied by contraction and expansion of suitable habitats, likely facilitated gene exchange among populations, influencing the current distribution pattern in subtropical China. Our results suggested that each ptDNA haplotype should be treated as an independent unit for conservation purposes, and ex-situ efforts should be prioritized to conserve *C. grandis* in China.

1. Introduction

Caldesia Parl., a genus comprising only three aquatic species, is widely distributed across the wetlands and swamps of the eastern hemisphere (Chen et al., 2006; Wang et al., 2010). In China, two species—*C. grandis* and *C. parnassifolia*—are recognized and have been classified as Critically Endangered (CR) on the Chinese IUCN Red List of Threatened Species (https://www.mee.gov.cn/xxgk2018/xxgk/xxgk01/202305/W020230522536560832337.pdf). Notably, *C. grandis* is distinguished within the genus as the only species

---

* Corresponding author.

** Corresponding author at: Aquatic Plant Research Center, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan 430074, China.

E-mail addresses: jmchen@wbgcas.cn (J.-M. Chen), wbg_georgelee@163.com (Z.-Z. Li).

https://doi.org/10.1016/j.gecco.2024.e03235
Received 17 June 2024; Received in revised form 26 September 2024; Accepted 29 September 2024

with leaves wider than they are long and vegetatively through bulbils that occur in the inflorescences (Wang et al., 2010). It is primarily found in the mid-low mountainous swamps and is natural in limited areas ranging from 1000 to 1500 m in China and the eastern Himalayas (Chen et al., 2006; Lehtonen, 2009). Since the 20th century, excessive human activity and global climate warming have severely impacted the number and size of wild *C. grandis* populations, leading to its believed extinction in mainland China and only sole population from Taiwan region in the previous century (Wang et al., 2001; Liu et al., 2002). However, in the last decade, three natural populations were discovered in mainland, including one in Yunnan Province and two in Hunan Province, each of small size with less than 100 individuals per population (Chen et al., 2006; Zhou et al., 2014).

Recently, several studies have been conducted on *C. grandis*, focusing on cytology (Wang et al., 2001), phylogeny (Lehtonen, 2009; 2017; Mwanzia et al., 2019; Li et al., 2022), and reproductive ecology (Gituru et al., 2002; Liu et al., 2002; Chen et al., 2003; Chen et al., 2007). However, only a few researches have been dedicated to conservation genetics. Chen et al. (2006) and Yue et al. (2012) detected the genetic diversity and population structure in three populations using Random Amplified Polymorphic DNA (RAPD) and Inter-Simple Sequence Repeat (ISSR) markers. Chen et al. (2006) indicated the high levels of clonal diversity in *C. grandis,* which inferred that sexual reproduction has played a vital role during the evolutionary history. Yue et al. (2012) investigated the spatial structure and genetic variation of three populations with implications for the conservation and management of this endangered plant. However, all previous studies explicitly targeted the three initially found natural populations (BH, GH, and LPH; Table 1), and this number remained unchanged until recent years as new population records were increasingly reported in Hunan, Zhejiang, Guangdong, and Fujian Provinces (Yang et al., 2016; Xiao et al., 2016; Zheng et al., 2017). For instance, Xiao et al. (2016) discovered a new population of *C. grandis* in the tianchi swamp of Qingzhangshan Nature Reserve, located in Nanxiong County, and subsequent community research indicated that human activities (i.e., damming and grazing) have seriously impacted the reproduction and regeneration of the population. In addition, most of the extant populations were located in the Nanling Mountains, which have been proven likely to be one of the vital ice age refuges in the subtropical region of China (Shi et al., 2014; Tian et al., 2018; Yang et al., 2019). However, due to the previously limited number of wild populations, the impact of ancient climate change and geological history on the phylogeographic pattern of *C. grandis* remains insufficiently understood.

The emergence of next-generation sequencing technologies has revolutionized the study of conservation genetics (Ogden et al., 2013; Hendricks et al., 2018; Cano et al., 2022). Compared to traditional molecular markers (i.e., RAPD, ISSR, and Simple Sequence Repeat), Restriction site-associated DNA sequencing (RAD-seq) can provide sufficient resolution for genetic diversity and population structure inference with single nucleotide polymorphisms (SNPs), which has been extensively applied in the study of phylogeny, biogeography, and conservation genetics without a reference genome, such as Trichophoreae (Léveillé-Bourret et al., 2020), *Cornus* (Du et al., 2024), *Nyssa sylvatica* complex (Zhou et al., 2018), *Ipomoea batatas* (Feng et al., 2020) and *Kadsura interior* (Dong et al., 2022). In this study, we utilized RAD-seq and five plastid DNA (ptDNA) markers to assess the genetic diversity and phylogeographic patterns of *C. grandis* in China, integrating these analyses with Species Distribution Models (SDMs). We collected samples from nine populations, including five that were newly recorded in the mainland over the past decade and the sole Taiwan population, from across all known regions in China. Our findings provide valuable insights into the extent of distribution and genetic variation of *C. grandis*, aiding in the development of tailored conservation strategies for this species and other aquatic plants with similar distribution patterns in the subtropical mid-low mountain regions of China.

2. Material and methods

2.1. Sample collection

From 2007–2018, all currently known nine extant populations of *C. grandis* in China were visited (Fig. 1), and six to nine individuals were sampled randomly from each population by immediately drying fresh leaves with silica gel for subsequent DNA experiments (Table 1). However, due to the extremely small size of the TR population, all the six individuals found were sampled.

Table 1
Summary of genetic statistics for *Caldesia grandis* in China.

Pop code	Location	N	Geography coordinates	ptDNA haplotypes	RAD-seq				
					AR	HE	HO	π	Fis
GD	Guidong, Hunan	6/7	E 113.98; N 25.98	H1 (6)	1.320	0.194	0.224	0.213	−0.001
NX	Nanxiong, Guangdong	6/8	E 114.44; N 25.04	H2 (6)	1.286	0.177	0.241	0.191	−0.077
GH	Guaihu, Hunan	6/7	E 112.85; N 24.99	H3 (6)	1.348	0.211	0.251	0.232	−0.020
LPH	Langpanhu, Hunan	6/8	E 112.90; N 24.92	H3 (6)	1.287	0.177	0.146	0.191	0.143
QT	Qingtian, Zhejiang	6/9	E 120.32; N 27.99	H1 (6)	1.253	0.158	0.212	0.169	−0.052
TR	Tuorong, Fujian	6/3	E 119.94; N 27.20	H1 (6)	1.268	0.149	0.206	0.179	−0.043
LW	Linwu, Hunan	6/8	E 112.41; N 25.15	H3 (6)	1.267	0.165	0.227	0.178	−0.068
BH	Beihai, Yunnan	6/9	E 098.57; N 25.13	H4 (6)	1.299	0.186	0.240	0.199	−0.049
YL	Yilan, Taiwan	6/8	E 121.57; N 24.61	H1 (5), H5(1)	1.331	0.204	0.248	0.220	−0.021
Mean					1.295	0.180	0.222	0.197	−0.021

N, sample size of ptDNA/RAD-seq; H_E, expected heterozygosity; H_o, observed heterozygosity; π, nucleotide diversity; Fis, inbreeding coefficient

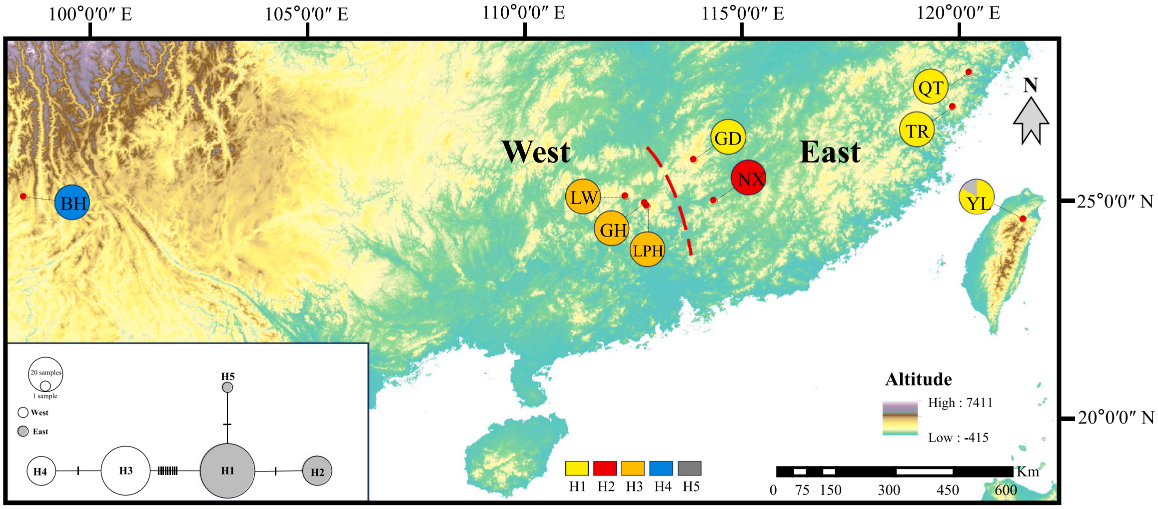


Fig. 1. The dated phylogenetic tree of *Caldesia grandis* and *C. parnassifolia*. Sharps indicates the data downloaded from Li et al. (2022). Numbers under the tree branches represented mean divergent ages and 95 % confidence interval of each node. Red rhombus indicated the second calibrating points.

2.2. DNA extraction, PCR amplification, and sequencing

Genomic DNA was isolated from all samples using the modified CTAB protocol (Doyle and Doyle, 1987), and the quality of the DNA was assessed using a NanoDrop spectrophotometer. Five cpDNA intergenic regions (*trnV*$^{(UAC)}$-*accD*, *rps20-clpP*, *ycf1-trnN*$^{(GUU)}$, *ndhF-rpl32*, *ndhD-psaC*) were amplified from six individuals per population using the specific primers (Table S1), developed on the basis of the published plastome of *C. grandis* (Mwanzia et al., 2019). The PCR amplification and DNA sequencing were performed following Li et al. (2020).

2.3. RAD library construction, sequencing, and SNP calling

According to the population size, six to nine samples in each population were selected for RAD sequencing. The library was prepared following the traditional single-digest RAD protocol with *Eco*RI (Baird et al., 2008). We conducted 150 bp paired-end sequencing on the Illumina Hiseq 2500 platform at Novogene Bioinformatics Technology Co. Ltd. (Beijing, China). After sequencing, three individuals from the TR population were removed due to the low sequencing quality. All RAD-seq data (ca. 0.2 G - 2 G per sample; Table S2) of the remaining 67 samples were first filtered using program process_radtags v2.66 (Rochette et al., 2019). Then, we utilized the denovo_map.pl v2.66 program to implement the Stacks pipeline, including *ustacks*, *cstacks*, *sstacks*, *tsv2bam*, *gstacks*, and *populations* in the Stacks components, with default settings (Rochette et al., 2019). We implemented additional filtering criteria, selecting only loci present in at least five populations (–min-populations 5) and in a minimum of 70 % of samples within each population (–min-samples-per-pop 0.7). Furthermore, only SNPs with a minor allele frequency (MAF) of at least 5 % were included (–min-maf 0.05). Then, potential plastid fragments in *de novo* assembled loci from RAD-seq were searched using the BlastN program (e-value cutoff of 1e-5), and the complete plastome of *C. grandis* (MH262300.2) was retrieved from the Genbank and applied as the reference. The filtered high-quality nuclear SNPs were applied for the subsequent analysis.

Table 2
Weir & Cockerham Fst (lower triangle) and gene flow (Nm; upper triangle) between different populations.

Pop code	GD	NX	GH	LPH	QT	TR	LW	BH	YL
GD	-	3.434	1.975	4.049	1.188	18.115	2.058	1.906	1.824
NX	0.068	-	2.505	3.839	1.929	8.594	5.023	6.703	2.696
GH	0.112	0.091	-	4.302	1.544	8.723	3.290	2.789	1.201
LPH	0.058	0.061	0.055	-	2.428	53.107	5.376	3.967	1.455
QT	0.174	0.115	0.139	0.093	-	5.019	1.188	2.287	1.007
TR	0.014	0.028	0.028	0.005	0.047	-	12.885	43.395	3.637
LW	0.108	0.047	0.071	0.044	0.174	0.019	-	5.080	0.901
BH	0.116	0.036	0.082	0.059	0.099	0.006	0.047	-	1.054
YL	0.121	0.085	0.172	0.147	0.199	0.064	0.217	0.192	-

2.4. RAD-seq data analysis

The basic statistics of genetic diversity for each population, including expected heterozygosity (H_E), observed heterozygosity (H_o), nucleotide diversity (π), and inbreeding coefficient (Fis), were obtained from the *populations* module. Also, Allelic richness (A_R) was calculated using hierfstat v0.5–11 (Goudet, 2005) with the minimum of three samples. For the inference of population structure, principal component analysis (PCA) and maximum-likelihood approach were performed using Plink v1.90b6.21 (Purcell et al., 2007) and Admiture v1.3.0 (Alexander et al., 2009), respectively. The 20-fold cross-validation (CV) was run with *K* values set from 1 to 9, and the optimal *K* was determined according to the lowest CV error. To understand the genetic difference among populations, we calculated the Weir & Cockerham Fst and Nei's genetic distance (Da; Nei et al., 1983) between populations using hierfstat v0.5–11 (Goudet, 2005) and gene flow (Nm) was estimated based on Fst value, Nm=(1 - Fst)/4Fst (Slatkin and Barton, 1989). Also, the relationship between geographical and genetic distances was tested using the vegan v2.6–6.1 package (Oksanen et al., 2007) with 999 permutations.

2.5. ptDNA sequence analysis

All newly generated ptDNA sequences of five regions were aligned using MAFFT v7.310 (Kuraku et al., 2013) with default parameters, respectively. The haplotypes were identified from the concentrated ptDNA matrix in DnaSP v6.0 (Rozas et al., 2017), and the parsimony network was assessed by PopArt v1.7 (Leigh et al., 2015). The phylogenetic relationship among the identified haplotypes was inferred using maximum-likelihood analyses conducted in IQTREE v1.6.12 (Nguyen et al., 2015) with 1000 bootstrap support (BS) replicates, and *C. parnassifolia* (MH262301.2) was used as an outgroup.

To estimate the divergence time among ptDNA haplotypes, uncorrelated log-normal Bayesian analysis with a strict clock and Yule process prior was conducted in BEAST v1.10.4 (Bouckaert and Rambaut, 2014) under the GTR substitute model, which was selected by Modeltest v3.7 (Posada and Crandall, 1998). Here, two second calibration points were chosen based on our previous study (Li et al., 2022): 1) the crown age of *C. grandis* and sister species *C. parnassifolia* at 5.46 Ma; 2) the divergence between Eurasian and African populations of *C. parnassifolia* from each other was set at 1.06 Ma. The time range of each point was set according to the 95 % highest posterior densities in Li et al. (2022). 1×10^8 MCMC generations were run and sampled every 2000 generations, and the first 20 % of trees were discarded as burn-in. The program Tracer v1.7.1 (Rambaut et al., 2018) was utilized to check the convergence according to the ESS (> 200), and TreeAnnotator v1.8.4 in BEAST package was applied to summarize the Bayesian trees.

2.6. Niche modeling analysis

The MaxEnt v3.4.4 software (Phillips et al., 2006) was used to predict the historical and current potential distribution of *C. grandis*.

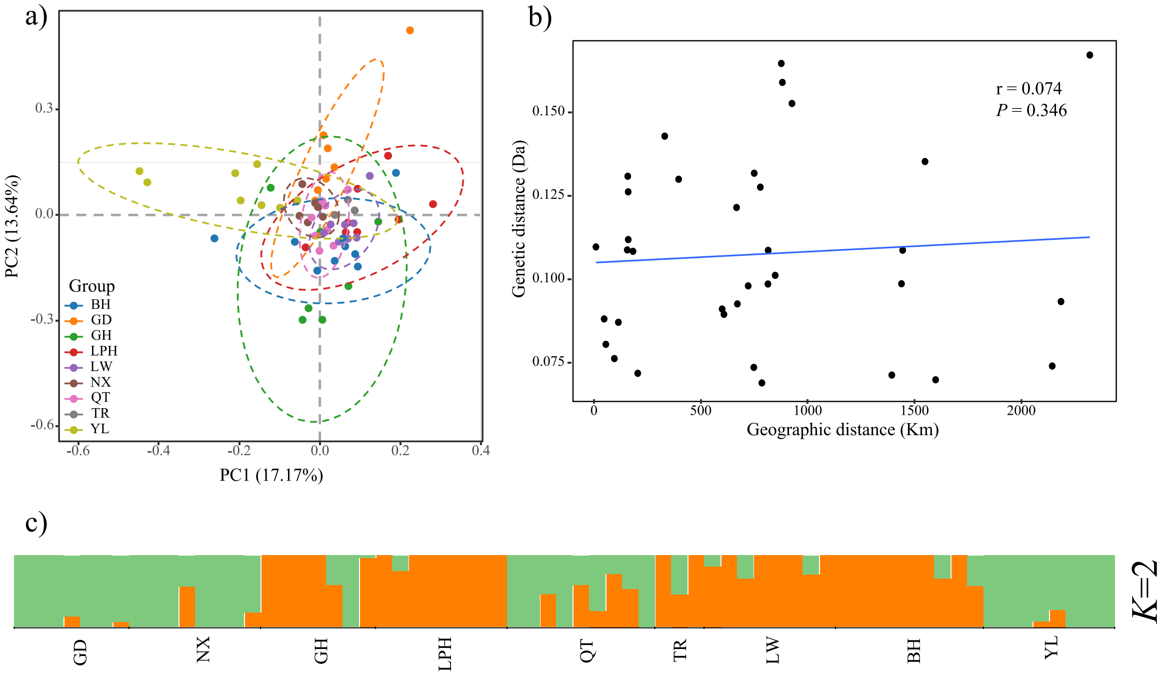


Fig. 2. Genetic stucture and mantel test between genetic and geographic distance for nine extant populations of *Caldesia grandis* using RAD-seq data. a) Principle coordinates analysis in all 67 individuals from nine populations. b) Mantel test of between Da and geographic distance. c) Proportional membership of nine extant populations to *K*=2 genetic clusters.

The location information was retrieved from the Global Biodiversity Information Facility (GBIF; https://www.gbif.org) and our study. The bioclimatic factors of current and five historical periods, including the Marine Isotope Stage Ma (M2; ca. 3.3 Ma), mid-Pliocene warm period (mPWP; ca. 3.205 Ma), Marine Isotope Stage 19 (MIS19; ca. 787 ka), Last Interglacial (LIG; ca. 130 ka) and Last Glacial Maximum (LGM; ca. 21 ka), were obtained from PaleoClim (Brown et al., 2018). To avoid spatial autocorrelation and sampling bias, only one site was kept in a grid of 10×10 km cells using spThin package v1.0.0 (Aiello-Lammens et al., 2015). Also, we excluded the highly correlated bioclimatic factors with the Pearson correlation coefficient values of $> |0.8|$. After filtering, the mean temperature of the warmest quarter (bio10), mean temperature of the coldest quarter (bio11), annual precipitation (bio12), precipitation seasonality (bio15), and precipitation of the warmest quarter (bio18) were further utilized to predict the suitable distribution. To improve the accuracy of the model, regularization multiplier (rm) and feature classes (fc) were determined using ENMeval v2.0.4 package (Kass et al., 2021) with five fcs (L, LQ, H, LQH, LQHP, LQHPT) and the rm was set from 0.5 to 4 at increment of 0.5. Other parameters followed Nzei et al. (2022) and the area under the curve (AUC) was applied to evaluate the model's accuracy.

3. Results

3.1. Genetic diversity and population structure

A total of ca. 35.67 G of high-quality data was generated, with 294,171,791 clean reads, with an average of 0.55 G for each individual (Table S2). After the *denovo* pipeline of Stacks and strict SNP filtering, a total of 360 high-quality SNPs from 1814 loci remained, and the average sequencing depth of individuals and mean genotyped site per locus were 5.8 x and 118 bp, respectively (Table S2). In our study, five genetic indices (A_R, H_E, H_o, π and Fis) were detected among populations (Table 1). Overall, the average H_o and H_E ranged from 0.146 (LPH) to 0.251 (GH) and 0.149 (TR) to 0.211 (TR), respectively. The GH population showed the highest nucleotide diversity ($\pi = 0.232$) and allelic richness ($A_R = 1.348$), and the lowest values ($\pi = 0.169$; $A_R = 1.253$) were detected in the QT population. For the inbreeding coefficient, except LPH (Fis = 0.143), the values of the remaining populations were less than 0 and ranged from −0.001 to −0.077.

Pairwise Fst values were detected that ranged from 0.005 (between TR and LPH populations) to 0.217 (between YL and LW populations), and the Nm ranged from 0.901 to 53.107. The CV-error values were increased with each *K* and did not show the valley (Fig. S1), which indicated that there was no clear genetic structure consistent with the PCA result (Fig. 2). To investigate the potential presence of genetic structures consistent with ptDNA, two genetic components were assumed across nine populations. Most individuals of western populations were assigned to Cluster I and Cluster II, including the remaining samples, which are predominantly from eastern populations (Fig. 2). Additionally, we did not find a significant correlation between geographical and genetic distances across populations using the Mantel test ($r = 0.074$, $P = 0.346$; Fig. 2).

3.2. ptDNA phylogeography

The aligned ptDNA matrice from nine populations was 3391 bp in length, and was concentrated into five haplotypes (H1-H5). Except for YL which has two haplotypes (H1 and H5), each of the other populations has only one haplotype (Fig. 1). Among these, H2

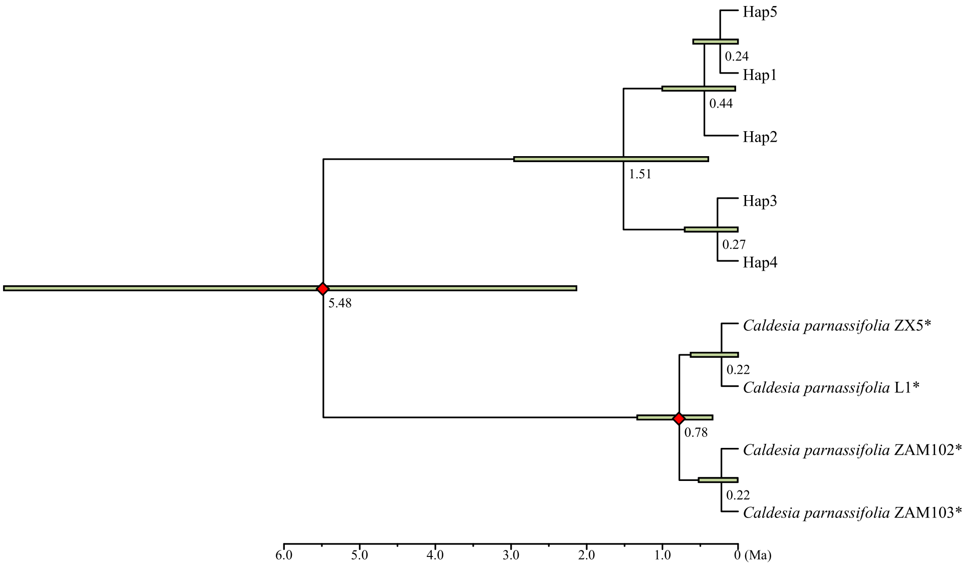


Fig. 3. Geographical distribution and network of five ptDNA haplotypes identified in *Caldesia grandis*.

and H4 were exclusive to NX and BH populations, respectively (Fig. 1; Table 1). Based on the phylogenetic analysis, two clades, CladeI (H1, H2, and H5) and Clade II (H3 and H4), were revealed with high supports (BS = 100; Fig. S2), which was consistent with the parsimony network (Fig. 1). Our analysis dates the crown node of *C. grandis* to ca. 1.51 Ma (95 % HPD: 0.39–2.96 Ma) and the CladeI and II emerged at ca. 0.44 Ma (95 % HPD: 0.04–1 Ma) and 0.27 Ma (95 % HPD: 0–0.7 Ma), respectively (Fig. 3).

3.3. Current and historical ecological niches

Based on the five selected bioclimatic factors, we predicted both the current and five historical distributions of *C. grandis* (Fig. 4) with high AUC values (> 0.9). Our prediction for current distribution closely aligns with the actual distribution of *C. grandis* in Southeast Asia, albeit with a notable expansion in Japan, where the species is not known to occur. During the M2 period (ca. 3.3 Ma), *C. grandis* enjoyed a broad, suitable range across East Asia and the Himalayas. However, during the mPWP (ca. 3.205 Ma), as temperatures rose and the Himalayas underwent significant orogeny, the suitable areas in the Himalayas contracted, while more favorable conditions emerged in warm regions such as Central and East China, Japan, and the Korean Peninsula. Later, during MIS 19 (ca. 787

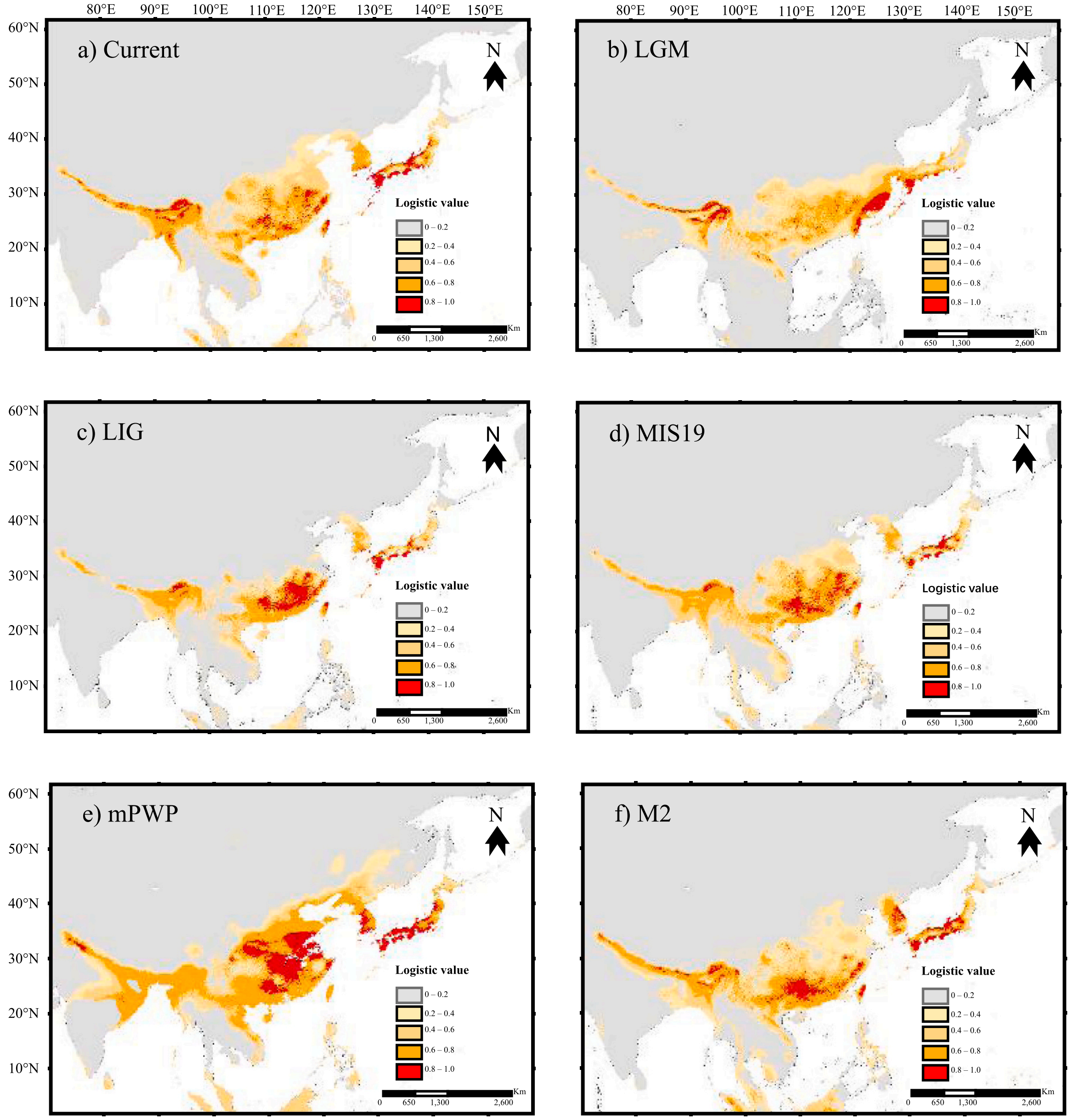


Fig. 4. Potential suitable distribution of *Caldesia grandis* inferred by MaxEnt modeling. a) The Current. b) The LGM (ca. 21 ka). c) The LIG (ca. 130 ka). d) The MIS19 (ca. 787 ka). e) The mPWP (ca. 3.205 Ma). f)The M2 (ca. 3.3 Ma).

ka) and the Last Interglacial (LIG, ca. 130 ka), the distribution range shifted southeastward in response to rising temperatures. By the Last Glacial Maximum (LGM, ca. 21 ka), the potential distribution of *C. grandis* contracted in southwest China and migrated southeastward, adapting to the decreasing global temperatures.

4. Discussion

Previous research on *C. grandis* was constrained by the number of populations and the molecular markers employed, focusing mainly on clonal diversity and spatial genetic structure through traditional markers (Chen et al., 2006; Yue et al., 2012). In our study, we provided a genomic perspective revealing that the extant nine populations display a lower genetic diversity ($H_E = 0.180$, ranging from 0.149 to 0.211) than some other endangered aquatic plants assessed using SSR markers, such as *Brasenia schreberi* ($H_E = 0.256$; Li et al., 2018), *Ottelia jingxiensis* ($H_E = 0.441$; Li et al., 2019), *Isoetes yunguiensis* ($H_E = 0.463$; Zheng et al., 2020) and the closely related species *Sagittaria natans* ($H_E = 0.340$; Yue et al., 2011). We cannot reject the possibility that the choice of molecular markers has contributed to the observed lower genetic diversity in *C. grandis*. Generally, SSR markers are considered more productive than SNPs, as demonstrated in various species (e.g., *Zea mays*, Van Inghelandt et al., 2010; *Sesamum indicum*, Basak et al., 2019). However, when compared to recent studies using reduced-representation genome sequencing in endangered aquatic plants, such as wild *Nelumbo nucifera* ($H_E = 0.29$; Liu et al., 2021) and *B. schreberi* ($H_E = 0.426$; Zhu et al., 2020) in China, *C. grandis* still exhibits relatively low genetic diversity. The level of genetic diversity might be associated with the range of species and mode of reproduction. With narrow distributions and clonal reproduction, species generally tend to exhibit lower levels of genetic variation (Zheng et al., 2020). *Caldesia grandis* is typically found sporadically distributed across mid-low mountainous swamps (Chen et al., 2006). Both previous studies and our field investigations have shown that the existing populations predominantly reproduce asexually using bulbils, with only a few seedlings observed in the wild (Gituru et al., 2002). In contrast, a lower inbreeding coefficient (mean Fis = -0.021, ranging from -0.077–0.143) has been detected across these populations. Also, high levels of clonal diversity were detected in three early populations (BH, GH, and LPH populations; Chen et al., 2006). This phenomenon can be attributed to founder effects and the genetic drift associated with the small size of all known *C. grandis* populations. However, it should be mentioned that sexual reproduction may play an important role at some time in the history of *C. grandis* populations (Chen et al., 2006).

The population structure of endangered plants is typically influenced by multiple factors, including habitat fragmentation, geographical isolation, breeding systems, and gene flow (Li et al., 2019). Our RAD-seq analysis revealed no significant population structure and low genetic differentiation (mean Fst = 0.089) across nine natural populations, consistent with a previous study showing that only a small proportion of genetic variation exists among populations (Chen et al., 2006). Also, our results indicated high gene flow (Nm > 1) and lack of isolation-by-distance ($r = 0.074$, $P = 0.346$), which may be attributed to frequent inter-population contacts facilitated by the expansion of the paleo-distribution area during the Quaternary period (Fig. 4). All extant populations of *C. grandis* in the current study were located in subtropical China, which is rarely affected by the repeated glacial periods of the Quaternary due to the complex topography. Moreover, it has given rise to multiple geographically isolated refugia for many plants (i.e., *Castanopsis eyrei*, Shi et al., 2014; *Eomecon chionantha*, Tian et al., 2018; *Liriodendron chinense*, Yang et al., 2019). The Nanling Mountains have been well studied and regarded as one of the important refugia in this region (Qiu et al., 2011). Compared to nuclear DNA, plastid datasets can offer insights into historical changes and genetic structure in species distribution that occur when new habitats are colonized through seed dispersal (Petit et al., 2003). Here, our phylogenetic analysis of five ptDNA haplotypes indicates that the western and eastern populations of *C. grandis* form two distinct clades across subtropical China. Notably, two haplotypes (H2 and H3) are located in the Nanling Mts and the east-west oriented Luoxiao Mts likely played a significant role in shaping the current phylogeographic patterns of *C. grandis*. Additionally, the genetic diversity held by the populations in the Nanling Mountains (NX, GH, LPH, and LW) is higher than the mean H_E of the species (0.183 vs 0.180), supporting the hypothesis that the Nanling Mountains serve as a refugium (Yang et al., 2019).

Estimates of divergence times based on ptDNA suggest that the split between the western and eastern clades of *C. grandis* occurred at ca. 1.51 Ma, possibly due to the uplift of Luoxiao Mts as the western-eastern barrier in the subtropical regions of China during the Pleistocene. Further diversification within the clade appeared ca. 0.2–0.4 Ma, which coincides with the Lushan glacial period (ca. 0.2–0.3 Ma; Wanti et al., 1980). Similar to other plants in this region—such as *Cercis chuniana* (Liu et al., 2021), *L. chinense* (Yang et al., 2019), and *Euptelea* species (Cao et al., 2016)—sharply reduced temperature forced further retreat into refuges, exacerbating differentiation among populations. The BH population from the Yunnan-Guizhou Plateau (YGP), characterized by the unique haplotype (H4), diverged from the populations of the Nanling Mountains around 0.27 Ma. This divergence coincided with the rapid and dramatic uplift of the YGP during the Middle Pleistocene (Tang et al., 1994). With the slowing of tectonic movements and the rise in temperature during the interglacial period (MIS19 and LIG), the suitable distribution area for *C. grandis* began to expand, which allowed previously isolated glacial populations to come into contact, leading to dispersal and the admixed hereditary components revealed by nuclear RAD-seq results (Fig. 2). Furthermore, the populations (QT and TR) from the Mts of Zhejiang and Fujian provinces are likely the result of the eastward expansion of GD population located in the Luoxiao Mts, as they share the same haplotype (H1). A similar pattern was previous observed in aquatic species *B. schreberi* (Li et al., 2018). During the Pleistocene, as sea levels dropped, land bridges between Taiwan and the Asian mainland emerged and may have facilitated the long-distance dispersal of the Taiwan population (YL) from the adjacent continent.

Recently, there has been growing attention directed towards the protection of aquatic plants in alpine wetlands (Li et al., 2006; Zhou et al., 2022; Qiu et al., 2024). However, the conservation efforts for aquatic plants in mid-low mountainous regions remain inadequate. Compared to alpine areas, wetlands in mid-low mountains are more vulnerable to climate change and human activities, posing significant threats to the survival of hydrophytes (i.e., *B. schreberi*, *S. lichuanensis* and *C. grandis*) in these regions (Liao et al., 2016; Li et al., 2018). In our field investigations over the last two decades and recent studies on *C. grandis* (Xiao et al., 2016; Yang et al., 2016), we have consistently observed severe declines in populations, with local habitats destroyed and even local extinctions occurring (i.e., LPH and TR populations), almost without exception. These negative impacts are primarily due to human agroforestry activities such as grazing, dam construction, and pesticide use. This study provides foundational information for the management and conservation of this endangered aquatic plant species. We identified low genetic diversity and two distinct clades of *C. grandis* within the existing populations. Combined with the geographical distribution of ptDNA haplotypes, we propose that each haplotype be considered an independent unit for conservation purposes. Moreover, ex-situ conservation efforts for *C. grandis* have been successfully implemented at the Wuhan Botanical Garden, CAS, and South China National Botanical Garden. However, caution is necessary to prevent the mixing of individuals or seeds from different populations—or at least from different haplotypes—to avoid potential outbreeding depression (Chen et al., 2008).

Ethical statement

This study did not involve any human participants, animal subjects, or environmental impact that would require ethical approval. As such, no ethical approval was necessary for the conduct of this research.

CRediT authorship contribution statement

Dong-Ying Yan: Writing - original draft, Writing -review & editing, Data curation, Formal analysis. **Shuai Peng**: Investigation, Methodology, Data curation. **Samuli Lehtonen**: Methodology, Writing -review & editing. **Jin-Ming Chen**: Supervision, Validation, Resources. **Zhi-Zhong Li**: Writing-review & editing, Funding acquisition, Project administration, Resources, Supervision, Validation.

Declaration of Competing Interest

The authors declare that there are no conflicts of interest.

Data Availability

I have released the data in the public database

Acknowledgments

This study was supported by grants from the National Natural Science Foundation of China (Nos. 32100186 and 32070253) and the Special Research Assistant Project, Chinese Academy of Sciences (No. E2291M01).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.gecco.2024.e03235.

References

Aiello-Lammens, M.E., Boria, R.A., Radosavljevic, A., Vilela, B., Anderson, R.P., 2015. spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. Ecography 38 (5), 541–545.

Alexander, D.H., Novembre, J., Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Genome Res 19 (9), 1655–1664.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A., Johnson, E.A., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PloS One 3 (10), e3376.

Basak, M., Uzun, B., Yol, E., 2019. Genetic diversity and population structure of the Mediterranean sesame core collection with use of genome-wide SNPs developed by double digest RAD-Seq. PLoS One 14 (10), e0223757.

Cano, M.J., Twyford, A.D., Hollingsworth, P.M., 2022. Population and conservation genetics using RAD sequencing in four endemic conifers from South America. Biodivers. Conserv. 31 (13), 3093–3112.

Cao, Y.N., Comes, H.P., Sakaguchi, S., Chen, L.Y., Qiu, Y.X., 2016. Evolution of East Asia's Arcto-tertiary relict *Euptelea* (Eupteleaceae) shaped by late Neogene vicariance and quaternary climate change. BMC Evol. Biol. 16, 1–17.

Chen, D., Chen, J.M., Wang, Y., Wang, Q.F., 2003. Floral organogenesis of *Caldesia grandis* Samuel. (Alismataceae). Acta Phytotax. Sin. 41 (3), 229–234.

Chen, J., Gituru, W.R., Wang, Q., 2007. Quantitative estimates of outcrossing rates in a natural population of *Caldesia grandis* (Alismataceae). Wuhan. Univ. J. Nat. Sci. 12 (2), 337–342.

Chen, J.M., Gituru, W.R., Wang, Y.H., Wang, Q.F., 2006. The extent of clonality and genetic diversity in the rare *Caldesia grandis* (Alismataceae): comparative results for RAPD and ISSR markers. Aquat. Bot. 84 (4), 301–307.

Chen, Y.Y., Li, X.L., Yin, L.Y., Li, W., 2008. Genetic diversity of the threatened aquatic plant *Ottelia alismoides* in the Yangtze River. Aquat. Bot. 88 (1), 10–16.

Dong, Y., Wei, X., Qiang, T., Liu, J., Che, P., Qi, Y., Zhang, B., Liu, H., 2022. RAD-seq and ecological niche reveal genetic diversity, phylogeny, and geographic distribution of *Kadsura interior* and its closely related species. Front. Plant Sci. 13, 857016.

Doyle, J.J., Doyle, J.L., 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull.

Du, Z.Y., Cheng, J., Xiang, Q.Y., 2024. RAD-seq data provide new insights into biogeography, diversity anomaly, and species delimitation in eastern Asian–North American disjunct clade Benthamidia of *Cornus* (Cornaceae). J. Syst. Evol. 62 (1), 1–19.

Feng, J., Zhao, S., Li, M., Zhang, C., Qu, H., Li, Q., Li, J., Lin, Y., Pu, Z., 2020. Genome-wide genetic diversity detection and population structure analysis in sweetpotato (*Ipomoea batatas*) using RAD-seq. Genomics 112 (2), 1978–1987.

Gituru, W.R., Wang, Q.F., Wang, Y., Guo, Y.H., 2002. Pollination ecology, breeding system, and conservation of *Caldesia grandis* (Alismataceae), an endangered marsh plant in China. Bot. Bull. Acad. Sin. 43.

Goudet, J., 2005. Hierfstat, a package for R to compute and test hierarchical F-statistics. Mol. Ecol. Notes 5 (1), 184–186.

Hendricks, S., Anderson, E.C., Antao, T., Bernatchez, L., Forester, B.R., Garner, B., Hand, B.K., Hohenlohe, P.A., Kardos, M., Koop, B., Sethuraman, A., Waples, R.S., Luikart, G., 2018. Recent advances in conservation and population genomics data analysis. Evol. Appl. 11 (8), 1197–1211.

Kass, J.M., Muscarella, R., Galante, P.J., Bohl, C.L., Pinilla-Buitrago, G.E., Boria, R.A., Soley-Guardia, M., Anderson, R.P., 2021. ENMeval 2.0: Redesigned for customizable and reproducible modeling of species' niches and distributions. Methods Ecol. Evol. 12 (9), 1602–1608.

Kuraku, S., Zmasek, C.M., Nishimura, O., Katoh, K., 2013. aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. Nucleic Acids Res. 41 (W1), W22–W28.

Lehtonen, S., 2009. Systematics of the Alismataceae—a morphological evaluation. Aquat. Bot. 91 (4), 279–290.

Lehtonen, S., 2017. Splitting Caldesia in favour of *Albidella* (Alismataceae). Aust. Syst. Bot. 30 (1), 64–69.

Leigh, J.W., Bryant, D., Nakagawa, S., 2015. POPART: full-feature software for haplotype network construction. Methods Ecol. Evol. 6 (9).

Léveillé-Bourret, É., Chen, B.H., Garon-Labrecque, M.È., Ford, B.A., Starr, J.R., 2020. RAD sequencing resolves the phylogeny, taxonomy and biogeography of Trichophoreae despite a recent rapid radiation (Cyperaceae). Mol. Phylogenet. Evol. 145, 106727.

Li, Z., Yu, D., Xiong, W., Wang, D., Tu, M., 2006. Aquatic plants diversity in arid zones of Northwest China: patterns, threats and conservation. Biodivers. Conserv. 15, 3417–3444.

Li, Z.Z., Gichira, A.W., Wang, Q.F., Chen, J.M., 2018. Genetic diversity and population structure of the endangered basal angiosperm *Brasenia schreberi* (Cabombaceae) in China. PeerJ 6, e5296.

Li, Z.Z., Lu, M.X., Gichira, A.W., Islam, M.R., Wang, Q.F., Chen, J.M., 2019. Genetic diversity and population structure of *Ottelia acuminata* var. *jingxiensis*, an endangered endemic aquatic plant from southwest China. Aquat. Bot. 152, 20–26.

Li, Z.Z., Ngarega, B.K., Lehtonen, S., Gichira, A.W., Karichu, M.J., Wang, Q.F., Chen, J.M., 2020. Cryptic diversity within the African aquatic plant *Ottelia ulvifolia* (Hydrocharitaceae) revealed by population genetic and phylogenetic analyses. J. Plant Res. 133, 373–381.

Li, Z.Z., Lehtonen, S., Martins, K., Wang, Q.F., Chen, J.M., 2022. Complete genus-level plastid phylogenomics of Alismataceae with revisited historical biogeography. Mol. Phylogenet. Evol. 166, 107334.

Liao, Y.Y., Gichira, A.W., Wang, Q.F., Chen, J.M., 2016. Molecular phylogeography of four endemic *Sagittaria* species (Alismataceae) in the Sino-Japanese floristic region of East Asia. Bot. J. Linn. Soc. 180 (1), 6–20.

Liu, K.M., Lei, I.G., Hu, G.W., 2002. Developmental study on the inflorescence and flower of *Caldesia grandis* Samuel (Alismataceae). Bot. J. Linn. Soc. 140 (1), 39–47.

Liu, W., Xie, J., Zhou, H., Kong, H., Hao, G., Fritsch, P.W., Gong, W., 2021. Population dynamics linked to glacial cycles in *Cercis chuniana* FP Metcalf (Fabaceae) endemic to the montane regions of subtropical China. Evol. Appl. 14 (11), 2647–2663.

Liu, Z., Xiao, B., Zhu, H., Kuang, J., Ji, Q., Peng, J., Li, F., Sun, Y., Ke, W., 2021. Genetic diversity and population structure of wild lotus (*Nelumbo nucifera* Gaertn.) in China. Plant Sci. J. 39 (03), 278–287.

Mwanzia, V.M., Nzei, J.M., Yan, D.Y., Kamau, P.W., Chen, J.M., Li, Z.Z., 2019. The complete chloroplast genomes of two species in threatened monocot genus *Caldesia* in China. Genetica 147, 381–390.

Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32 (1), 268–274.

Nzei, J.M., Ngarega, B.K., Mwanzia, V.M., Kurauka, J.K., Wang, Q.F., Chen, J.M., Li, Z.Z., Pan, C., 2022. Assessment of climate change and land use effects on Water Lily (*Nymphaea* L.) habitat suitability in South America. Diversity 14 (10), 830.

Ogden, R., Gharbi, K., Mugue, N., Martinsohn, J., Senn, H., Davey, J.W., Pourkazemi, M., McEwing, R., Eland, C., Vidotto, M., Sergeev, A., Congiu, L., 2013. Sturgeon conservation genomics: SNP discovery and validation using RAD sequencing. Mol. Ecol. 22 (11), 3112–3123.

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M.H.H., Oksanen, M.J., Suggests, M.A.S.S., 2007. The vegan package. Community Ecol. Package 10 (631-637), 719.

Petit, R., Aguinagalde, I., de Beaulieu, J.L., Bittkau, C., Brewer, S., Cheddadi, R., Ennos, R., Fineschi, S., Grivet, D., Lascoux, M., Mohanty, A., Müller-Starck, G., Demesure-Musch, B., Palmé, A., Martín, J.P., Rendell, S., Vendramin, G.G., 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. Science 300 (5625), 1563–1565.

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecol. Model. 190 (3-4), 231–259.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14 (9), 817–818.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81 (3), 559–575.

Qiu, D., Zhang, H., Ren, Y., Zhu, Y., 2024. The lost biodiversity and degraded alpine wetlands caused by strong earthquake on the Qinghai–Tibet Plateau did not self-restore in the short term. Glob. Ecol. Conserv., e02830

Qiu, Y.X., Fu, C.X., Comes, H.P., 2011. Plant molecular phylogeography in China and adjacent regions: tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. Mol. Phylogenet. Evol. 59 (1), 225–244.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A., 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67 (5), 901–904.

Rochette, N.C., Rivera-Colón, A.G., Catchen, J.M., 2019. Stacks 2: analytical methods for paired-end sequencing improve RADseq-based population genomics. Mol. Ecol. 28 (21), 4737–4754.

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., Sánchez-Gracia, A., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol. Biol. Evol. 34 (12), 3299–3302.

Shi, M.M., Michalski, S.G., Welk, E., Chen, X.Y., Durka, W., 2014. Phylogeography of a widespread Asian subtropical tree: genetic east–west differentiation and climate envelope modelling suggest multiple glacial refugia. J. Biogeogr. 41 (9), 1710–1720.

Slatkin, M., Barton, N.H., 1989. A comparison of three indirect methods for estimating average levels of gene flow. Evolution 43 (7), 1349–1368.

Tang, G.B., Zhang, J.P., Yang, X.D., Luo, B.X., Wang, Y.Z., Chen, P.Y., 1994. Late Cenozoic palynoflora and environment changes in Yunnan-Guizhou Plateau. Mar. Geol. Quat. Geol. 14, 91–104.

Tian, S., Kou, Y., Zhang, Z., Yuan, L., Li, D., López-Pujol, J., Fan, D., Zhang, Z., 2018. Phylogeography of Eomecon chionantha in subtropical China: the dual roles of the Nanling Mountains as a glacial refugium and a dispersal corridor. BMC Evol. Biol. 18, 1–12.

Van Inghelandt, D., Melchinger, A.E., Lebreton, C., Stich, B., 2010. Population structure and genetic diversity in a commercial maize breeding program assessed with SSR and SNP markers. Theor. Appl. Genet. 120, 1289–1299.

Wang, Q.F., Haynes, R.R., Hellquist, C.B., Alismataceae, M., Wu, Z.Y., Hong, D.Y., Raven, P.H., 2010. Flora of China, 23. Missouri Botanical Garden Press, Beijing, pp. 84–89 (Science Press & St. Louis).

Wang, Y., Wang, Q.F., Gituru, W.R., Guo, Y.H., 2001. Karyotypical studies on *Caldesia grandis* Samuel., a rare and endangered marsh plant in China. Plant Sci. J. 19 (3), 187–190.

Wanti, D., Quingyu, P., Xihao, W., 1980. Climatic variations in China during the Quaternary. GeoJournal 4, 515–524.

Xiao, J.S., Zhong, P.S., Hu, Y.H., 2016. Characteristic of community structure studies on *Caldesia grandis* at Tianchi swamp on Qingzhangshan nature reserve on Nanxiong County of Guangdong Province. For. Environ. Sci. 32 (02), 48–51.

Yang, A., Zhong, Y., Liu, S., Liu, L., Liu, T., Li, Y., Yu, F., 2019. New insight into the phylogeographic pattern of *Liriodendron chinense* (Magnoliaceae) revealed by chloroplast DNA: east–west lineage split and genetic mixture within western subtropical China. PeerJ 7, e6355.

Yang, W., Zeng, X., Li, M., 2016. Analysis on the interspecific association of *Caldesia grandis* community in Linwu Pingxidong mountain wetland of Hunan Province. Hunan For. Sci. Technol. 43 (03), 56–61.

Yue, X.L., Chen, J.M., Guo, Y.H., Wang, Q.F., 2011. Population genetic structure of *Sagittaria natans* (Alismataceae), an endangered species in China, revealed by nuclear SSR loci analyses. Biochem. Syst. Ecol. 39 (4-6), 412–418.

Yue, X.L., Chen, J.M., Guo, Y.H., Wang, Q.F., 2012. Fine-scale spatial genetic structure of an endangered marsh herb, *Caldesia grandis* (Alismataceae). Genet. Mol. Res. 11 (3), 2412–2421.

Zheng, L.X., Miao, X.T., An, C., Wang, P.C., Ma, J.Q., Bai, Z.Y., Lin, Q.Q., Lin, Y., Huang, Z.H., 2017. New data of seed plants from Fujian. Fujian For. Sci. Technol. 44 (02), 109–112.

Zheng, T., He, X., Ye, H., Fu, W., Peng, M., Gou, G., 2020. Phylogeography of the rare and endangered lycophyte *Isoetes yunguiensis*. PeerJ 8, e8270.

Zhou, W., Ji, X., Obata, S., Pais, A., Dong, Y., Peet, R., Xiang, Q.Y.J., 2018. Resolving relationships and phylogeographic history of the *Nyssa sylvatica* complex using data from RAD-seq and species distribution modeling. Mol. Phylogenet. Evol. 126, 1–16.

Zhou, Y.D., Xiao, K.Y., Chen, S.C., Liu, X., Wang, Q.F., Yan, X., 2022. Altitudinal diversity of aquatic plants in the Qinghai-Tibet Plateau. Freshw. Biol. 67 (4), 709–719.

Zhou, Z., Onevilay, S., Deng, T., Tan, Y., Sun, H., 2014. *Caldesia* Parl., A new record genus of Alismataceae to Laos. J. Trop. Subtrop. Bot. 22 (5), 447–449.

Zhu, H., Du, J., Liu, Z., Sun, Y., Li, M., Peng, J., Zhou, K., Ke, W., 2020. Investigation and genetic diversity of wild water shield (Brasenia schreberi) in China. J. Plant Genet. Resour. 21 (06), 1586–1595.