

RESEARCH

Open Access



Ensemble species distribution modeling and multilocus phylogeography provide insight into the spatial genetic patterns and distribution dynamics of a keystone forest species, *Quercus glauca*

Ying Song¹, Gang-Biao Xu¹, Ke-Xin Long¹, Chun-Cheng Wang², Ran Chen¹, He Li¹, Xiao-Long Jiang^{1*} and Min Deng^{3*}

Abstract

Background Forests are essential for maintaining species diversity, stabilizing local and global climate, and providing ecosystem services. Exploring the impact of paleogeographic events and climate change on the genetic structure and distribution dynamics of forest keystone species could help predict responses to future climate change. In this study, we combined an ensemble species distribution model (eSDM) and multilocus phylogeography to investigate the spatial genetic patterns and distribution change of *Quercus glauca* Thunb, a keystone of East Asian subtropical evergreen broad-leaved forest.

Results A total of 781 samples were collected from 77 populations, largely covering the natural distribution of *Q. glauca*. The eSDM showed that the suitable habitat experienced a significant expansion after the last glacial maximum (LGM) but will recede in the future under a general climate warming scenario. The distribution centroid will migrate toward the northeast as the climate warms. Using nuclear SSR data, two distinct lineages split between east and west were detected. Within-group genetic differentiation was higher in the West than in the East. Based on the identified 58 haplotypes, no clear phylogeographic structure was found. Populations in the Nanling Mountains, Wuyi Mountains, and the southwest region were found to have high genetic diversity.

Conclusions A significant negative correlation between habitat stability and heterozygosity might be explained by the mixing of different lineages in the expansion region after LGM and/or hybridization between *Q. glauca* and closely related species. The Nanling Mountains may be important for organisms as a dispersal corridor in the west-east direction and as a refugium during the glacial period. This study provided new insights into spatial genetic patterns and distribution dynamics of *Q. glauca*.

Keywords Oak, Dispersal corridor, Spatial genetic pattern, Subtropical evergreen broad-leaved forest, Climate change

*Correspondence:

Xiao-Long Jiang
xiaolongjiang1@gmail.com
Min Deng
dengmin@ynu.edu.cn

¹College of Forestry, The Laboratory of Forestry Genetics, Central South University of Forestry and Technology, Changsha 410004, Hunan, China

²College of Landscape Architecture, Central South University of Forestry and Technology, Changsha 410004, Hunan, China

³School of Ecology and Environmental Sciences, Yunnan University, Kunming, Yunnan 650500, China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

The geological record contains many examples of divergences in species lineages and distribution changes caused by either climate fluctuations or geologic activity [1–3]. Recent climate change, meanwhile, has affected ecosystem function and biodiversity patterns by altering forest community structure and species distribution patterns [4, 5]. Going forward, all evidence suggests rapid climate change will remain a primary threat to species and ecosystems, and these effects will intensify as the pace of change accelerates [6–8]. Since keystone species have outsized roles in ecosystem formation and community distribution dynamics [9], exploring the effects of paleogeological events and climate change on the spatial genetic patterns in keystone species can not only reveal the ancient drivers of speciation but also help to evaluate the distribution dynamics of ecosystems in the context of future climate change [10, 11].

Subtropical East Asia has experienced complex tectonic activity and climate change effects in recent geological records. These include the uplift of the Himalayan-Tibetan Plateau from about 50 Ma [12, 13], the aridification of Central Asia from about 22 Ma [14], and the formation of Asian monsoons 8–9 Ma [15]. These events significantly impacted East Asia's landforms, topography, and climate, which profoundly affected genealogical differentiation, genetic diversity, and distribution of organisms [16–18]. The diverse microhabitats in the complex topography regions will provide suitable habitats for species surviving during climate change [17, 19], thus serving as refugia for many relict taxa [20, 21].

The primary ecosystem of subtropical East Asia is an evergreen broadleaved forest (EBLF), which stretches between 24–32°N and 99–123°E. EBLFs in subtropical East Asia comprise 2,600 genera and ca.14,600 species of seed plants, of which >50% are endemic [22]. Conservation of this biodiversity in EBLFs in subtropical East Asia is the subject of much attention from researchers, especially in the face of ongoing rapid climate change.

The genus *Quercus* (oaks), which includes ca. 450 species, is widely distributed in the Northern Hemisphere. Besides providing wood and food, oaks are keystone species in different habitats [23]. Oaks can undergo frequent introgression events and are adaptable to short- and long-term environmental change [24–27]. The *Quercus robur* genome confirmed the intergenerational transmission of adaptive variants, for instance [28]. Genomic and anatomic data analysis indicated that oaks might have different defense mechanisms against pathogens [29]. Chloroplast DNA (cp.DNA) markers from *Q. acutissima*, *Q. chenii*, and *Q. variabilis* revealed haplotype sharing within section *Cerris* in East Asian EBLFs that was associated with locally stable climates and complex landscapes [30]. Based on SSR and phenotypic data in

two oak species indicated that asymmetric inter-specific selection pressures could contribute to the asymmetric trait divergence where species coexist [31]. Resequencing data indicated that the introgression between two widespread sympatric Asian oak species, *Q. acutissima* and *Q. variabilis*, confers environmental adaptation by altering the regulation of stress-related genes [32]. These factors justify using oaks as a model species for understanding the development of evolutionary patterns from ecological processes.

Quercus glauca Thunb, one of the most widespread species in the *Quercus* section *Cyclobalanopsis*, is a keystone species in subtropical evergreen forests. Population genetic patterns and local adaptation in the species have been investigated in previous studies. For example, genotyping six *Q. glauca* populations from East China using six enzymes found high genetic diversity but low genetic differentiation [33]. Based on three cp.DNA loci, southeastern Taiwan Island was suggested to be a potential glacial refugium for *Q. glauca* [34]. Genotyping 409 individuals from 42 populations from China mainland and Japan using three cp.DNA loci, it was found that genetic differentiation of *Q. glauca* began in the Miocene and it may have experienced expansion after the Last Glacial Maximum (LGM; approximately 22,000 years ago) [35]. Although the samples in Xu et al. [35] roughly cover the current *Q. glauca* distribution, they lacked representation from Taiwan Island. Furthermore, cp.DNA can only reveal seed gene flow but without information about pollen gene flow. Different markers can reveal the evolutionary history of species more objectively [36, 37]. Finally, compared with the single model (MaxEnt) previously used to predict the species range of *Q. glauca* [35], the ensemble species distribution model (eSDM) can integrate multiple models and improve prediction accuracy [38, 39].

In this study, we combined nuclear microsatellites (*n*SSR), cp.DNA markers, and eSDM to investigate the spatial genetic patterns and distribution dynamics of *Q. glauca*. A total of 781 individuals from 77 populations were sampled. The samples were genotyped using seven *n*SSR and three cp.DNA loci. The eSDM was used to estimate the distribution dynamics of *Q. glauca* under predicted past and future climate change. Our goals were to (i) reveal spatial genetic patterns of *Q. glauca*, (ii) estimate the effect of environmental factors on genetic diversity, and (iii) infer potential dispersal corridors. Our study provides guidelines for preserving *Q. glauca* germplasm resources and managing the biodiversity of East Asia subtropical EBLFs in the context of future climate change.

Results

Ecological niche modeling

Based on the filtered 238 distribution points and nine environmental variables (Fig. S1), nine niche models were used to predict the potential distribution of *Q. glauca* (Fig. S2). Among the nine models, RF, GLM and GBM have higher mean TSS and AUC values, and SRE has the lowest TSS and AUC values (Fig. S2 and Table S1). After excluding nine models from SRE, one model from CTA, and one model from MaxEnt with low evaluation (TSS<0.6 and AUC<0.8), the remaining 79 models were used to construct the ensemble model based on the weighted TSS (Table S1). The average TSS and AUC values for the remaining 79 models were 0.90 and 0.99, respectively. The predicted potential distribution of *Q. glauca* has an extensive suitable range in East Asia (Fig. S3). During the LGM period, the potential suitable distribution of *Q. glauca* was mainly in southern China, the southern slope of the Himalayas, the Korean Peninsula, and the East China Sea continental shelf. In the present period, southern China, Taiwan Island, and southwestern Japan show high suitability. In the future, due to climate change, southeast China and Japan will show high distribution probabilities, but the overall distribution and high suitability areas are predicted to diminish.

From the LGM to the present period, the suitable habitat of *Q. glauca* has expanded. The overall suitable habitat increased by 41%, with an area of 78.36×10^4 km² mainly concentrated in the northern portion of their distribution after a northeastern expansion (Fig. 1). In the future, the area will shrink by an estimated 33%, leaving habitat mainly concentrated in southern Korea and southwest China. According to the migration routes through the centroids, the shrinking occurs towards the northeast, similar to the expansion after the LGM.

Genetic diversity and structure based on microsatellites

A total of 620 individuals from 60 populations were genotyped by seven *n*SSR primers (Table 1). The genetic diversity indices H_o , H_e , and A_r were 0.38–0.79 (mean: 0.57), 0.43–0.77 (mean: 0.60), and 2.88–5.97 (mean: 3.74), respectively (Table 1). The highest H_o and H_e were found in populations 76 and 30, respectively, and the lowest for both statistics was found in population 17. The A_r value of population 26 was the highest, and the lowest was found in population 32. The $\text{Pr}(X|K)$ and ΔK support the presence of two major clusters of *Q. glauca* populations (Fig. S4). This result clustered 20 populations into a West group and 40 populations into an East group (Fig. 2b and c). Principal coordinate analyses (PCoA) also suggested the populations should be clustered into two similar groups. The first two PCs explained 22.9% and 14.3% of the genetic variation, respectively (Fig. 2a). The genetic structure between groups was significant

(Gp1, p -value=0.024; Gp2, p -value=0.0001). The Mantel test showed that isolation by distance was significant ($R^2=0.499$, $P=0.001$; Fig. S5).

Genetic diversity and differentiation based on cp.DNA markers

Three chloroplast loci were sequenced in 534 individuals from 57 populations (Table 1). The concatenated sequence was 2,437 bp total after alignment, and the lengths of the *psbA-trnH*, *trnT-trnL*, and *atpI-atpH* regions were 506 bp, 780 bp, and 1,151 bp, respectively. A total of 58 haplotypes with 40 variable sites were detected (Fig. 3a). Overall π and H_d of the populations were 1.22×10^{-3} and 0.96, respectively. The highest H_d was observed in populations 50 ($H_d=0.87$) and 4 ($H_d=0.80$), and the highest π was found in populations 4 ($\pi=2.98 \times 10^{-3}$) and 6 ($\pi=2.39 \times 10^{-3}$; Table 1). Spatial genetic interpolation results showed higher H_d and π in the Nanling Mountains, Wuyi Mountains, and the southern region (Fig. S6). Overall diversity (H_T) and diversity within populations (H_S) based on cp.DNA were 0.859 and 0.286, respectively. The N_{ST} value (0.68) was only slightly higher than the G_{ST} value (0.67) and the difference was not significant ($p>0.05$), suggesting the absence of a strong phylogeographic signal. Spatial distributions of haplotypes, including 18 shared and 40 private haplotypes, were arranged in a star-like structure (Fig. 3b). Haplotype 24 was the most widely distributed, found in 53 individuals from 10 populations. A total of 13 haplotypes were derived from haplotype 24. The second most frequent was haplotype 25, found in 46 individuals from 9 populations. Populations 50, 4, and 21 had the most haplotypes ($N=6, 5, 5$, respectively), and 21 populations contained one haplotype.

Based on the species distribution model and the geographical distribution of haplotypes, the least-cost path method (LCP) method was used to calculate the possible migration and diffusion routes of *Q. glauca* from the LGM to the present period (Fig. 4). The results showed that *Q. glauca* migrated across the Yunnan Guizhou Plateau and the Nanling Mountains from the LGM to the present, as east-west mountains provide the main diffusion corridors. Genetic landscape shape analysis based on *n*SSR and cp.DNA data showed that genetic distances within the West population were higher than those within the East population. Notably, different patterns of population differentiation were detected between cp.DNA and *n*SSR markers (Fig. 5). For cp.DNA, genetic distance significantly declined from the West to the East and showed a peak in southwest China. The *n*SSR genetic distance peaked around the Nanling Mountains area, consistent with the higher genetic diversity indices observed in the regions.

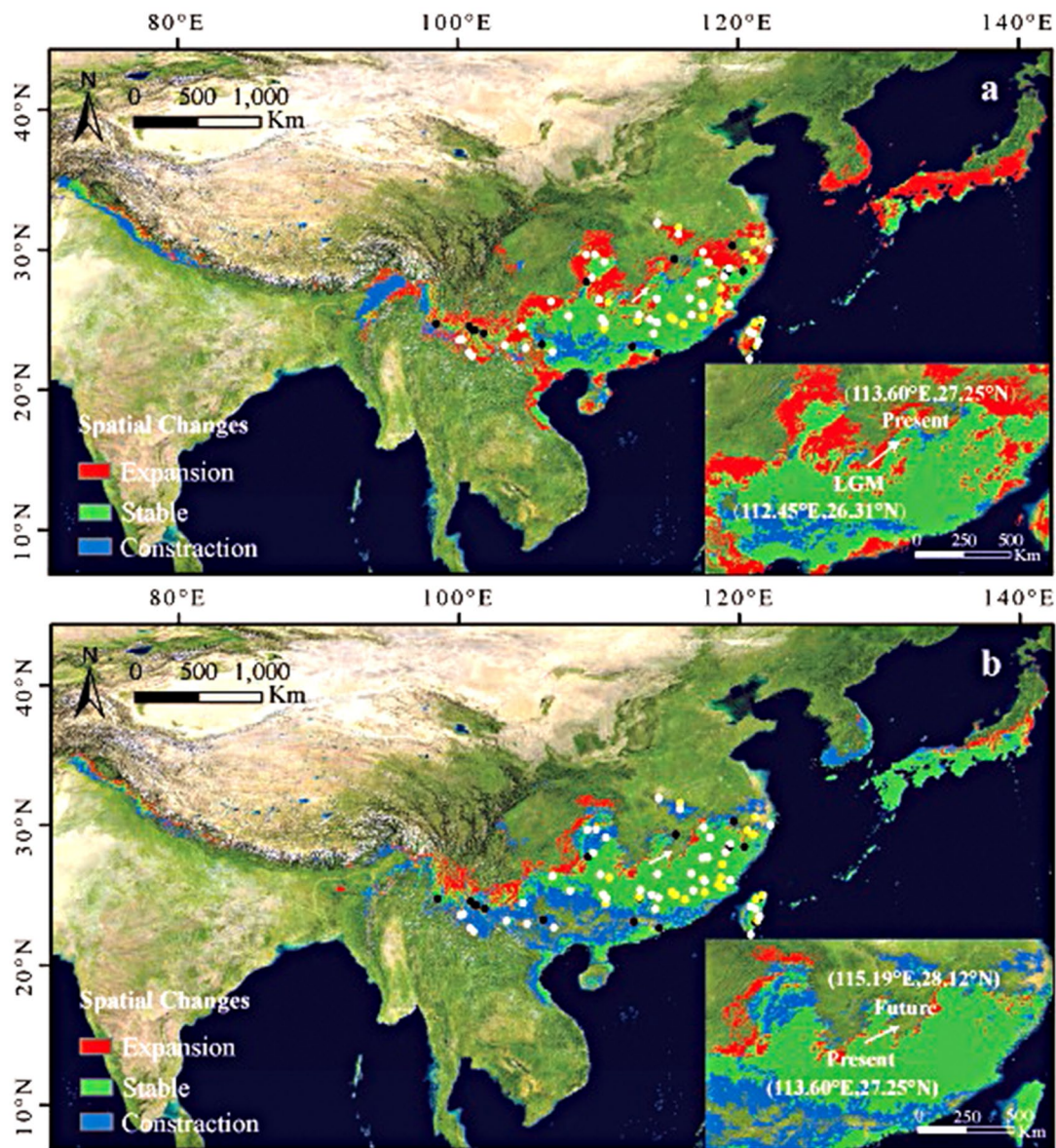


Fig. 1 The distribution dynamic of *Quercus glauca* under climate change is based on the ensemble model and 238 distribution points. **a** the LGM (CCSM4) to present period (1970–2000s); **b** present to future period (Mean of four BCC scenarios, SSP126, SSP245, SSP370, and SSP585, in the 2081–2100s). The blue, green, and red regions represent the areas of loss, stability and gain in response to climate change. The yellow, black, and white dots represent populations for *n*SSR analysis only (20 populations), for cp.DNA analysis only (17 populations), and for both *n*SSR and cp.DNA analysis (40 populations), respectively. The bottom right of each figure represents the change in the centroid of the species distribution in response to climate change

Correlation between genetic diversity and climate factors
 GLM analysis indicated that the present period suitable habitat (E_{Pre}) was significantly correlated with A_r and H_e when other factors were controlled (Table 2). Habitat stability (E_{Stab}) was significantly correlated with C_a , but it negatively correlated with H_e . The model that best explained C_a included E_{Stab} , longitude, and latitude, and the correlation with longitude was the strongest result. Longitude’s effect on genetic differentiation was greater than that of E_{Stab} . For H_e , E_{Pre} and E_{Stab} were included

in the final model. Only E_{Pre} was significantly correlated with A_r .

Discussion
Species spatial genetic patterns

Genetic diversity is essential for a species’ survival in the face of environmental changes and, as a result, for maintaining its role in ecosystem functions [40]. High genetic diversity ($H_T=0.859$ and $A_T=2.88–5.97$ in 620 individuals from 60 populations) of *Q. glauca* detected from *n*SSR data is similar to other oaks that are distributed

Table 1 Sampling information, nSSR and cpDNA genetic diversity, locality habitat suitability, stability obtained from SDMs for the *Quercus glauca* populations

Population ID	Longitude /Latitude	nSSR		cp.DNA				SDM						
		N	A _r	H _e	H _o	C _A	N	H _d	π (×10 ⁻³)	Haplotype	F _{Pre}	F _{LGM}	E _{Srab}	
West group														
1	98.48/24.75	-	-	-	-	-	8	0	0	0	0	0.72	0.67	0.95
2	100.12/23.57	14	3.45	0.56	0.49	0.43	10	0	0	0	0	0.74	0.40	0.66
3	100.32/23.66	13	3.04	0.51	0.42	0.23	11	0.33	0.28	0.28	0.28	0.78	0.61	0.84
4	100.79/22.68	7	3.51	0.50	0.45	0.29	10	0.80	2.98	2.98	0.28	0.59	0.52	0.93
5	100.83/24.54	-	-	-	-	-	8	0.25	0.22	0.22	0.22	0.74	0.76	0.97
6	101.04/22.45	10	3.46	0.59	0.54	0.00	10	0.64	2.39	2.39	0.59	0.48	0.88	
7	101.10/24.28	-	-	-	-	-	10	0.20	0.17	0.17	0.84	0.76	0.93	
8	101.26/24.27	-	-	-	-	-	10	0	0	0	0.34	0.36	0.99	
9	101.85/24.03	-	-	-	-	-	5	0	0	0	0.86	0.72	0.86	
10	103.38/23.20	10	3.45	0.62	0.58	0.10	14	0.26	0.57	0.57	0.86	0.50	0.63	
11	104.57/24.46	9	3.50	0.61	0.66	0.78	10	0	0	0	0.87	0.57	0.69	
12	104.85/22.98	10	3.69	0.62	0.60	0.10	10	0.20	0.17	0.17	0.64	0.40	0.76	
13	106.01/23.25	-	-	-	-	-	7	0	0	0	0.80	0.70	0.90	
14	106.67/26.33	5	3.43	0.51	0.54	0.60	10	0	0	0	0.76	0.46	0.70	
15	106.78/22.72	13	3.40	0.60	0.61	0.00	10	0.20	0.35	0.35	0.88	0.74	0.86	
16	107.94/25.29	11	4.26	0.71	0.63	0.18	10	0.78	1.93	1.93	0.97	0.81	0.84	
17	109.17/29.67	6	2.99	0.43	0.38	0.50	8	0	0	0	0.92	0.44	0.52	
18	109.20/27.74	-	-	-	-	-	5	0	0	0	0.72	0.58	0.86	
19	109.59/28.03	10	3.28	0.60	0.51	0.50	10	0.71	0.61	0.61	0.86	0.76	0.90	
20	109.79/29.68	26	4.17	0.69	0.65	0.04	11	0.62	0.60	0.60	0.88	0.17	0.30	
21	110.11/26.47	13	4.39	0.73	0.56	0.31	10	0.60	0.49	0.49	0.94	0.43	0.49	
22	110.31/25.08	10	4.27	0.71	0.61	0.10	10	0	0	0	0.96	0.91	0.96	
23	110.47/24.39	6	3.26	0.59	0.61	0.00	-	-	-	-	0.98	0.90	0.93	
24	110.48/29.13	10	4.00	0.63	0.59	0.00	10	0	0	0	0.82	0.69	0.87	
25	110.5/24.78	10	3.67	0.63	0.72	0.00	10	0	0	0	0.96	0.90	0.94	
26	110.75/26.28	8	5.97	0.54	0.54	0.50	-	-	-	-	0.94	0.71	0.77	
East group														
27	112.48/23.11	-	-	-	-	-	8	0	0	0	0.88	0.73	0.85	
28	112.87/24.81	-	-	-	-	-	12	0.30	0.13	0.13	0.97	0.76	0.79	
29	112.88/24.85	8	3.89	0.61	0.64	0.75	-	-	-	-	0.97	0.57	0.60	
30	112.95/25.38	24	4.91	0.77	0.58	0.04	10	0.69	0.63	0.63	0.93	0.80	0.87	
31	113.75/25.04	5	3.57	0.52	0.49	1.00	8	0.25	0.11	0.11	0.94	0.85	0.91	
32	113.99/24.03	10	2.88	0.52	0.53	1.00	10	0	0	0	0.98	0.77	0.79	
33	114.20/26.52	10	3.74	0.57	0.57	0.80	10	0.20	0.09	0.09	0.74	0.48	0.74	
34	114.24/24.82	9	3.87	0.61	0.62	0.89	9	0	0	0	0.97	0.81	0.84	
35	114.26/31.93	10	2.93	0.48	0.47	1.00	10	0.20	0.35	0.35	0.58	0.12	0.54	
36	114.28/22.65	-	-	-	-	-	9	0	0	0	0.97	0.71	0.74	

Table 1 (continued)

Population ID	Longitude /Latitude	nSSR		cpDNA						SDM			
		N	A _r	H _e	H _o	C _A	N	H _d	π (×10 ⁻³)	Haplotype	E _{Pre}	E _{LGM}	E _{Stab}
37	115.21/25.21	10	3.56	0.61	0.53	1.00	-	-	-	-	0.97	0.77	0.80
38	115.45/25.02	9	3.93	0.63	0.53	0.89	-	-	-	-	0.97	0.75	0.78
39	115.46/29.34	-	-	-	-	-	8	0.25	0.32	H29(1), H42(7)	0.43	0.40	0.97
40	115.76/31.60	8	3.61	0.55	0.54	1.00	-	-	-	-	0.69	0.18	0.49
41	115.78/31.13	9	3.84	0.65	0.62	1.00	10	0	0	H31(10)	0.93	0.37	0.44
42	116.13/24.70	10	3.34	0.54	0.57	0.90	-	-	-	-	0.93	0.81	0.88
43	116.57/26.55	9	3.88	0.59	0.57	0.89	9	0.56	0.24	H4(5), H6(4)	0.93	0.54	0.61
44	116.72/25.36	10	3.76	0.60	0.55	0.90	11	0.69	0.60	H14(5), H24(2), H31(4)	0.96	0.76	0.80
45	117.36/27.65	10	3.64	0.61	0.54	1.00	10	0.20	0.09	H35(9), H54(1)	0.93	0.77	0.84
46	117.46/26.18	10	3.89	0.64	0.61	1.00	-	-	-	-	0.80	0.81	0.99
47	117.49/29.84	10	3.83	0.57	0.56	1.00	10	0.53	0.69	H25(6), H29(4)	0.79	0.22	0.43
48	117.53/25.99	17	4.04	0.70	0.69	0.00	10	0.64	1.02	H14(4), H29(4), H31(1), H33(1)	0.89	0.38	0.49
49	117.54/25.21	10	3.77	0.59	0.53	0.90	-	-	-	-	0.96	0.74	0.78
50	117.67/27.73	10	3.34	0.55	0.51	0.90	10	0.87	1.64	H25(2), H26(1), H27(2), H29(1), H34(2), H35(2)	0.73	0.43	0.70
51	117.75/27.68	10	3.77	0.57	0.57	0.90	10	0.60	0.75	H14(1), H25(6), H31(3)	0.84	0.54	0.70
52	117.91/29.13	10	4.13	0.64	0.64	1.00	10	0	0	H27(10)	0.87	0.55	0.69
53	118.53/25.75	10	3.97	0.67	0.69	0.90	-	-	-	-	0.97	0.55	0.58
54	118.71/26.28	10	3.38	0.54	0.51	1.00	-	-	-	-	0.94	0.44	0.50
55	118.82/25.54	9	3.86	0.64	0.58	1.00	-	-	-	-	0.96	0.49	0.53
56	118.82/27.23	10	3.70	0.59	0.43	0.90	-	-	-	-	0.94	0.67	0.74
57	119.03/28.09	10	3.93	0.58	0.60	1.00	10	0.78	0.75	H25(3), H31(4), H40(2)	0.89	0.90	0.99
58	119.26/28.44	-	-	-	-	-	10	0.47	0.61	H25(6), H46(1), H47(3)	0.96	0.67	0.71
59	119.37/26.06	10	3.13	0.54	0.49	1.00	10	0.20	0.17	H24(9), H32(1)	0.93	0.64	0.71
60	119.37/28.68	10	3.94	0.59	0.59	0.90	9	0	0	H29(9)	0.93	0.74	0.80
61	119.61/30.30	-	-	-	-	-	10	0	0	H49(10)	0.60	0.17	0.57
62	120.37/28.48	-	-	-	-	-	10	0.51	0.41	H25(2), H30(7), H50(1)	0.96	0.82	0.86
63	120.61/29.63	10	3.55	0.55	0.44	1.00	-	-	-	-	0.83	0.24	0.41
68	121.06/29.28	10	3.81	0.62	0.62	0.90	-	-	-	-	0.94	0.46	0.53
69	121.14/30.61	10	4.03	0.65	0.70	0.90	-	-	-	-	0.62	0.10	0.48
77	122.21/30.02	16	3.45	0.60	0.56	1.00	10	0	0	H25(10)	0.72	0.31	0.59
Taiwan group													
64	120.72/22.18	-	-	-	-	-	8	0	0	H19(8)	0.90	0.63	0.73
65	120.79/24.17	10	3.63	0.63	0.58	0.70	6	0.60	0.61	H1(4), H4(1), H6(1)	0.97	0.49	0.52
66	120.80/22.21	13	3.75	0.62	0.59	0.46	7	0.48	1.03	H14(2), H19(4), H2(1)	0.96	0.80	0.83
67	120.84/24.18	8	3.30	0.58	0.71	0.13	-	-	-	-	0.98	0.55	0.57
70	121.14/23.03	-	-	-	-	-	5	0.70	0.52	H4(1), H14(3), H15(1)	0.98	0.85	0.87
71	121.14/24.03	9	3.80	0.65	0.65	1.00	10	0.64	0.40	H6(5), H12(4), H13(1)	0.98	0.63	0.65
72	121.21/24.76	10	4.39	0.68	0.64	0.60	-	-	-	-	0.98	0.91	0.93

Table 1 (continued)

Population ID	Longitude /Latitude	nSSR		cp.DNA					SDM				
		N	A _r	H _e	H _o	C _A	N	H _d	π (×10 ⁻³)	Haplotype	E _{Pre}	E _{LGM}	E _{Stab}
73	121.30/23.14	7	3.40	0.56	0.51	0.71	-	-	-	-	0.98	0.73	0.75
74	121.37/24.81	-	-	-	-	-	7	0	0	H11(7)	0.98	0.91	0.93
75	121.47/23.48	11	4.08	0.63	0.65	0.55	11	0.35	0.24	H1(5), H16(1), H17(1), H18(4)	0.98	0.92	0.94
76	121.68/24.99	8	4.16	0.65	0.79	0.50	-	-	-	-	0.98	0.84	0.86

Notes N: number of sampled individuals, A_r: standardized allelic richness, H_e: expected heterozygosity, H_o: observed heterozygosity, C_A: cluster A, H_d: haplotype diversity, π: nucleotide diversity, N_{Pre}: present habitat suitability, N_{LGM}: LGM habitat suitability, N_{Stab}: habitat stability since LGM, -: data missing

in subtropical China, such as the *Q. franchetii* complex ($H_T=0.982$ and $A_r=3.18-4.34$ in 303 individuals from 33 populations) [41], *Q. delavayi* ($H_T=0.907$ and $A_r=3.75-5.24$ in 493 individuals from 33 populations) [42], and *Q. schottkyana* ($H_T=0.828$ and $A_r=4.83-7.78$ in 380 individuals from 29 populations) [11]. For cp.DNA, the number of haplotypes in this study was higher than in other oaks, such as *Q. aliena* with 33 haplotypes in 52 populations [43], *Q. fabri* with 21 haplotypes in 31 populations [44], *Q. acutissima* with 19 haplotypes in 30 populations [45]. Different genetic patterns in nSSR and cp.DNA might be related to sampling variance. Although the same cp.DNA loci were sequenced here, the number of haplotypes was higher in the previous study from 409 *Q. glauca* individuals of 42 populations [35]. The additional 125 samples in this study added 25 new haplotypes, mainly from the western region (10 haplotypes) and Taiwan Island (9 haplotypes). Most of the additional haplotypes were private. Long-term geographical or environmental isolation could drive in-situ diversification and maintain private haplotypes as species adapt to local environments [46–48].

No significant phylogeographic structure was evident based on cp.DNA ($N_{ST}>G_{ST}$, $p>0.05$), a result consistent with other subtropical evergreen species such as *Q. variabilis* ($N_{ST}=0.855>G_{ST}=0.852$, $p>0.05$) [49], *Q. acutissima* ($N_{ST}=0.689>G_{ST}=0.630$, $p>0.05$) [45], and *Q. chenii* ($N_{ST}=0.888>G_{ST}=0.863$, $p>0.05$) [50]. In contrast, the genetic structure in nSSR data revealed a distinct pattern: 60 *Q. glauca* populations were split between two clusters representing the western and eastern portions of the sample area. The different genetic patterns detected in cp.DNA and nSSR makers are not unexpected. Because of their modes of inheritance, for instance, the markers have different dispersal patterns. Chloroplast DNA disperses through seeds and the nuclear genomes are transmitted via seeds and pollen. Oaks are prolific pollen producers and depend on wind for dispersal [51]. High levels of recent and ongoing gene flow should be evident in genetic patterns detected in nuclear markers, especially highly polymorphic SSRs [52]. To this latter point, more information may be available from nSSR than cp.DNA as the mutation rate in the former increases testable variation [53]. Meanwhile, the chloroplast genome is haploid with no recombination and has a smaller effective population size, which means it should be more sensitive to genetic drift and differentiation [54]. Therefore, differences in heredity between the nuclear and chloroplast genomes account for the different phylogeographic structure patterns we observed in the present study.

Genetic variation and environmental heterogeneity

Species with long generation times and large effective population sizes, such as oaks, are likely delayed in their responses to environmental change [55]. Consequently,

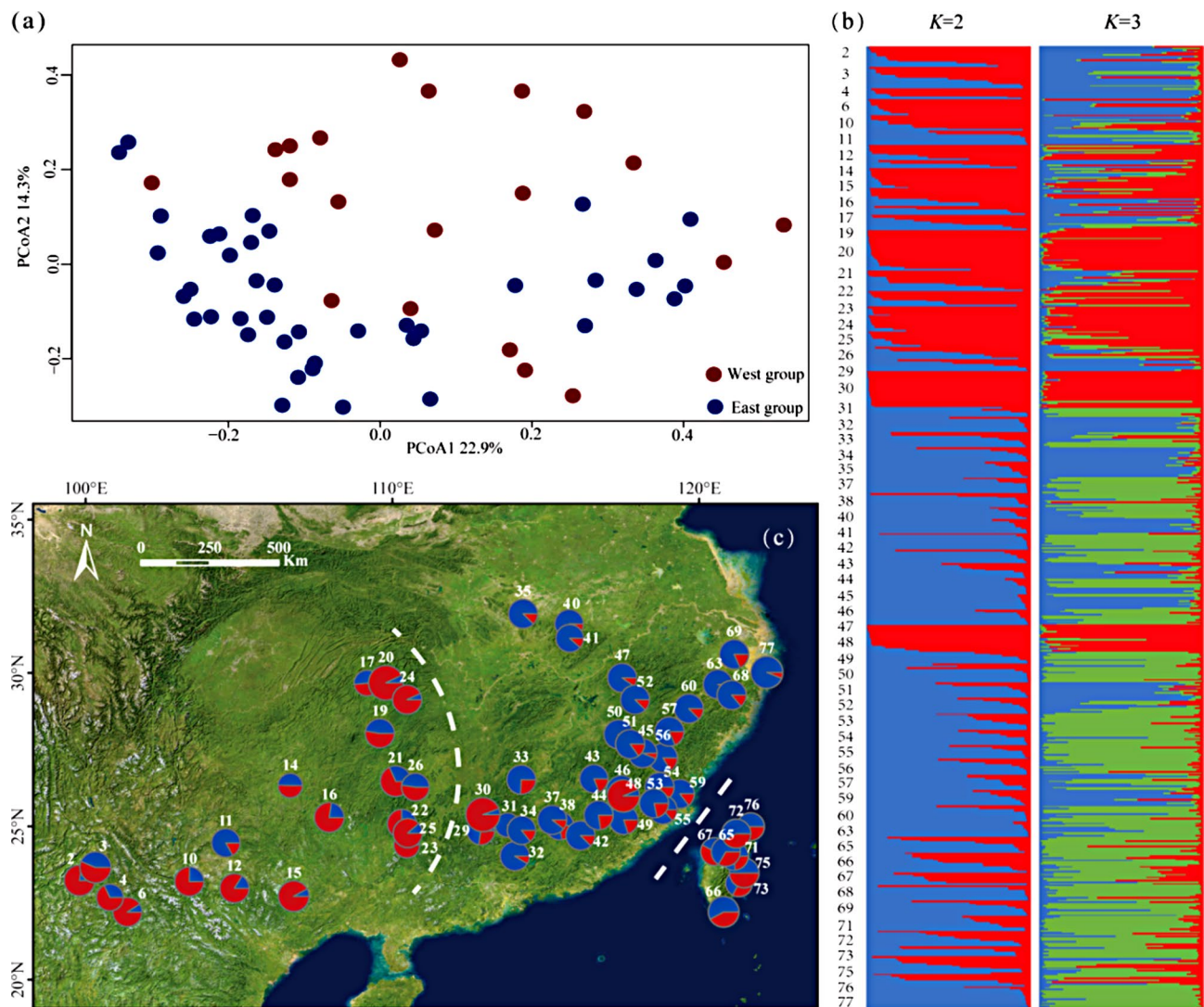


Fig. 2 Structure analysis of 60 *Quercus glauca* populations. **(a)** Principal coordinates analysis (PCoA) of the 60 populations of the *Quercus glauca* based on genetic distance using *n*SSR data. **(b)** Bayesian clustering plots for 60 populations of *Quercus glauca* based 7 *n*SSR loci, $K=2$ and $K=3$ was presented **(c)** Geographical distribution of the genetic clusters and genetic cluster composition in each population. The color of the pie chart represents different groups, and the white dotted line represents the geographical obstacles based on barrier analysis

contemporary genetic patterns are often explained by past environmental conditions [56, 57]. In *Q. glauca*, however, habitat stability since the LGM (E_{Stab}) was negatively correlated with genetic diversity (H_e , $p < 0.05$), a surprising result considering the expectation that more climatically stable regions will retain higher genetic diversity [58, 59]. In addition to accumulating mutations over time, interspecies hybridization and admixture between populations can also increase genetic diversity [60]. Hybridization and introgression are common in oaks [32], and since East Asia is the main distribution range of the *Quercus* section *Cyclobalanopsis*, many closely related oaks have a sympatric distribution with *Q. glauca* [60, 61]. The distribution overlap of *Q. glauca* and closely related species affords the possibility

of introgression. Range changes by *Q. glauca* since the LGM may have contributed to admixture among populations from different lineages. These factors could result in a negative correlation between genetic diversity and habitat stability. Similar results were also found in other oaks, such as *Q. kerrii* [62] and *Q. schottkyana* [11], in which the migration and mixing of genetically differentiated populations in unstable habitat areas led to increased genetic diversity.

Two distinct geography-related genetic lineages, West and East, were detected in *Q. glauca* populations from *n*SSR data. Genetic differentiation among populations in the western region is higher than in the eastern region. More complex topography in the west than in the east could provide higher stability for longer persistence and

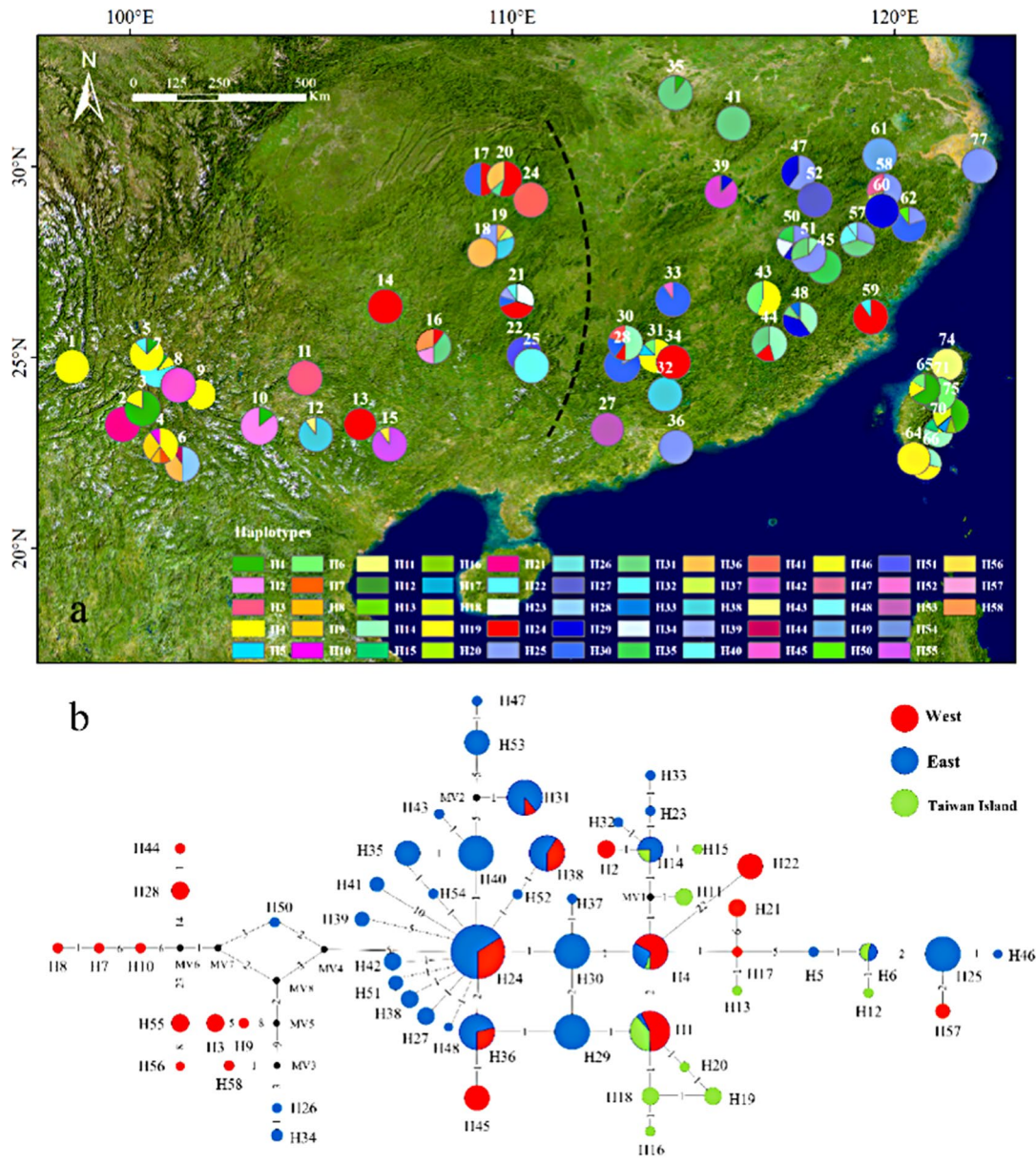


Fig. 3 Distribution of the cp.DNA haplotypes detected in *Quercus glauca*. **a** geographic distribution of 58 cp.DNA haplotypes; **b** haplotypes of the cp.DNA network of *Q. glauca*. The haplotype network was colored according to the three regions delineated on the map (West, East, and Taiwan island). The size of points representing populations on maps and haplotypes in networks are proportional to the number of individuals. Numbers on the branches indicate the number of substitutions. Black spots indicate unsampled or extinct ancestral haplotypes

preservation of genetic variation, which may have facilitated the divergence of lineages through glacial/interglacial cycles [63–65]. Meanwhile, long-term environmental heterogeneity along altitudinal or latitudinal gradients also contributes to species genetic differentiation [62]. Distinct east-west phylogeographic differentiation has also been observed in other plants, such as *Dyosma versipellis*, *Cephalotaxus oliveri* and *Liriodendron chinense* [66–68]. Similar genetic patterns from different species indicated that topography and environmental

heterogeneity could be the main drivers for intra-species differentiation in East Asia.

Species distribution dynamics and dispersal route

Range shifts are a primary response by species to environmental change [69, 70]. The fossil record reflects numerous examples in both pollen and fossilized remains of species distribution shifts [71]. Combining species distribution models with spatial genetic patterns is useful for reconstructing distribution dynamics and dispersal routes [72, 73]. Although the eSDM inferred the suitable

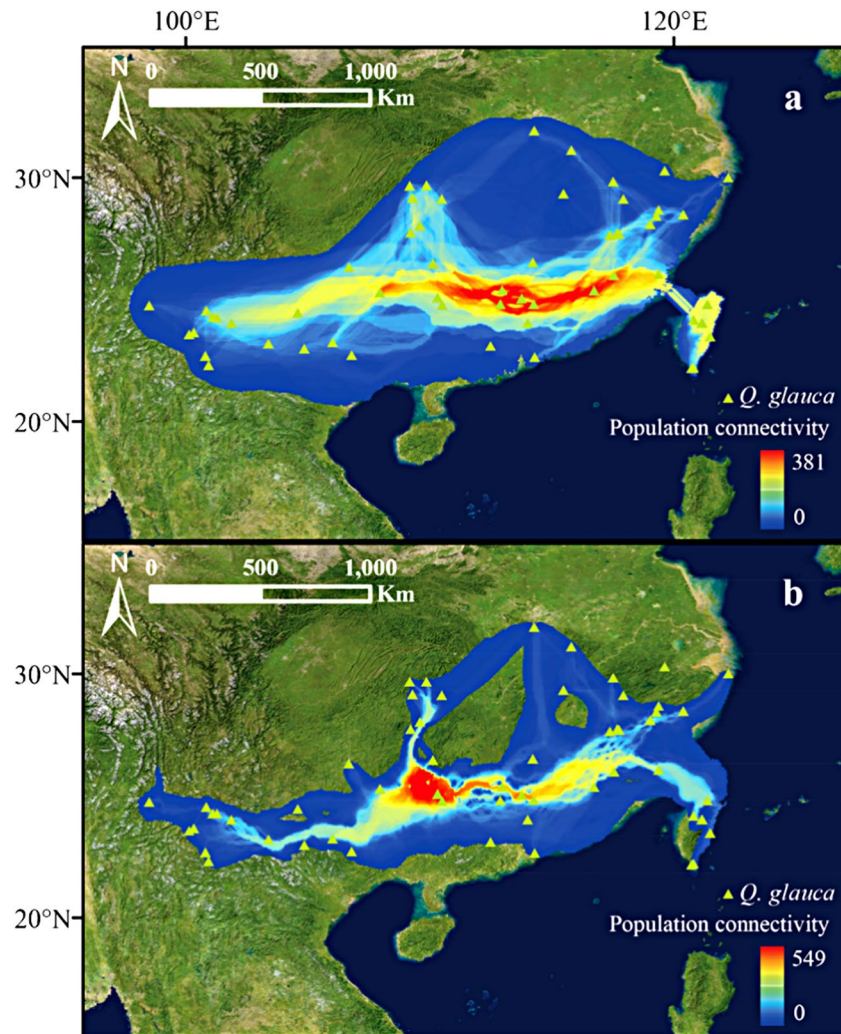


Fig. 4 Potential dispersal corridors of the *Quercus glauca* during (a) the present and (b) the LGM period. Colors from blue to red represent the probability of species dispersal corridors from low to high

habitat of *Q. glauca* has shifted northeastward since the LGM, the distribution centroids analysis revealed that this species underwent significant expansion in the southwestern region. This suggests a pattern of northward expansion and southward retreat experienced by the species during the glacial/interglacial period [74]. The potential distribution of *Q. glauca* in the future under a broad but realistic climate scenario is predicted to shrink and move northeast. According to Ni [75], average temperature increases in China by 0.6–6.3°C by the end of the 21st century will result in northward movements by forests and the formation of tropical rainforest climate in low latitudes. These predictions are supported by studies of other oaks, such as *Q. fabri* [76] and *Q. acutissima* [77].

The Nanling Mountains is a latitudinally arranged mountain range [78] that is commonly found to be a main route for species dispersal during glacial periods

[79]. This is also the case for *Q. glauca*, which dispersed between southwest and southeast China between differing climate periods. Southwest China has been proposed as a potential glacial refugium and is regarded as the “cradle” of East Asian flora [80]. This region is also the biodiversity center of section *Cyclobanopsis* [81], including through evidence from the fossil record such as with *Q. preglauca* (found in Sichuan, western China from the Pliocene and early Pleistocene, 4.5–2.3 Ma) [82] and *Q. preschottkyana* (found in Yunnan, southwest China from the late Miocene, about 11.5 Ma) [83]. Therefore, the direction of *Q. glauca* expansion from the LGM was from southwest to southeast China. Taiwan, meanwhile, is the largest subtropical mountainous island at the Tropic of Cancer in the monsoonal western Pacific region, and it emerged during the late Miocene and maintained a connection to the mainland at the latest in the Tertiary Pliocene [84–86]. The flora of Taiwan Island

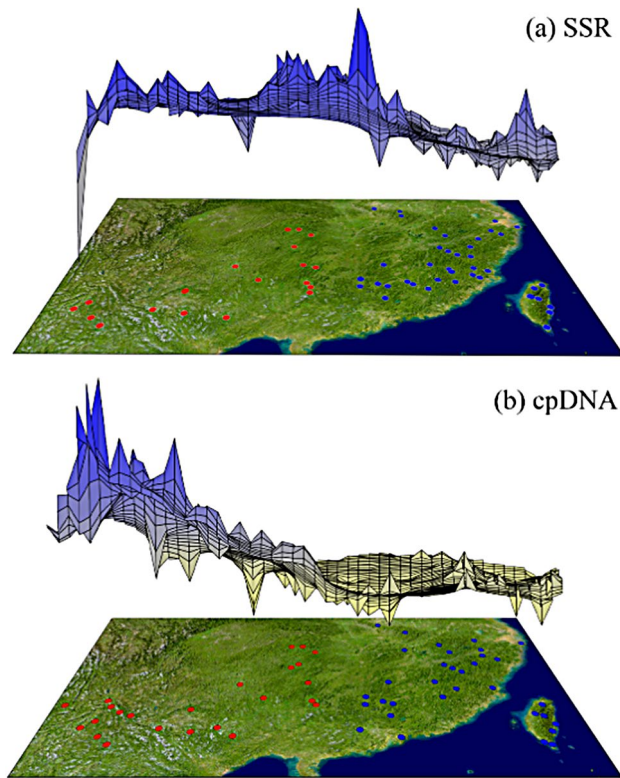


Fig. 5 Genetic landscape shape analysis of *Quercus glauca* population based on (a) *n*SSR and (b) cpDNA. The x-axis and y-axis represent the geographical location, and a Delaunay triangulation network is established between the sampled populations. Land surface height reflects the genetic distance between populations

is most similar to the Southeast China [87]. The island now has 181 species of angiosperms in 20 families and 63 genera, of which 19 families, 62 genera, and 83 species are common to mainland China [88]. The Taiwan Strait (TS), which is ca. 130 km wide at its narrowest point and has a depth ranging between ca. 50–160 m, is the main migration route between Taiwan and southern China. The most extensive land bridges in the TS, as well as the East and South China Sea, would have formed during the LGM period when sea level was ca. 130 m lower [89]. For *Q. glauca*, the TS was the conduit for dispersal and gene flow between the island and mainland during the LGM. Moreover, phylogeographic studies on *Quercus championii* [90], *Dyosma versipellis-pleiantha* [91], *Juglans cathayensis* [92] have revealed high genetic similarity on both sides of the TS.

Conclusions

This study sheds light on spatial genetic patterns and historical distribution dynamics of *Q. glauca*. The cpDNA and *n*SSR data revealed that the *Q. glauca* has high genetic diversity and can be divided into two distinct groups. Environmental heterogeneity contributed to the spatial genetic patterns observed. Interglacial warming

Table 2 Correlation between genetic diversity, habitat and geographical factors of *Q. glauca*

	A_r			H_e			C_a					
	Est	SE ^a	t	p	Est	SE ^a	t	p	Est	SE ^a	t	p
F_{Pre}	1.13	0.32	3.48	<0.001***	0.09	0.03	3.23	<0.01**	-	-	-	-
F_{Stab}	-	-	-	-	-0.10	0.05	-2.11	<0.05*	0.54	0.23	2.36	<0.05*
Longitude	-0.03	0.01	-1.88	0.06	-	-	-	-	0.03	0.01	4.76	<0.001***
Latitude	0.04	0.03	1.41	0.16	-	-	-	-	0.04	0.02	2.43	<0.05*

Significance code ****0.001, ***0.01, **0.05; SE^a: standard error

promoted the spread of *Q. glauca* from western to eastern China. The Nanling Mountains may be both a dispersal corridor and glacial refugium, and the land bridges between mainland China and Taiwan islands during glacial periods may have contributed to the expansion of *Q. glauca* from the former to the latter. These results provide new insights into phylogeographic patterns and evolutionary histories of biotas in Southern China and adjacent islands, but also helpful for the management of forest biodiversity of East Asia subtropical EBLFs in the face of future climate change.

Materials and methods

Ensemble species distribution model

Distribution records of *Q. glauca* were obtained from field surveys, published documents, and records in the Chinese Virtual Herbarium (www.cvh.org.cn/). To reduce the effect of sampling bias on the prediction results, only one point was selected in each cell (size: $0.1^\circ \times 0.1^\circ$). Nineteen bioclimatic variables with a $2.5'$ resolution for the present (1970–2000s), LGM (Community Climate System Model version 4, CCSM4), and the future period (2081–2100s; Shared Socioeconomic Pathways, SSPs; The Beijing Climate Center Climate System Model, BCC-CSM2-MR) were downloaded from WorldClim (<http://www.worldclim.org/>) [93]. The present and future bioclimatic variables were version 2.1 [94], and the LGM period was version 1.4 [95]. Four common socioeconomic pathways (SSP126, SSP245, SSP370, and SSP585) were averaged for future climatic time frames to generate an integrated scenario [96, 97]. The climate variables at occurrence points in the present period were extracted, and we tested for correlations among them using the R package “*dismo*” [98]. High correlation variables ($|r| \geq 0.8$) were removed before further analysis.

We used an ensemble forecasting approach using the R package “*biomod2*” [99] to predict *Q. glauca* distribution with nine model algorithms (Table S2). Each model runs ten times with random seeds. For each model, the distribution data is randomly divided into two datasets: 75% is training data, and the remaining 25% is testing data. To build reliable species distribution models, 10,000 pseudo-absence coordinates were randomly generated [100, 101]. The presence and absence points were set with equal weights. The area under the receiver operating curve (AUC) and the true skill statistic (TSS) were used to evaluate the performance of the model. Models with higher TSS (>0.6) and AUC values (>0.8) were selected to construct an ensemble model. The weighted value of each model was proportional to the TSS value when constructing the ensemble model [102]. Predicted habitat was divided into categories of marginal (0.2–0.4), moderate (0.4–0.6), and highly suitable (>0.6). To quantify the distribution change of *Q. glauca*, we also used 0.5 as the

threshold value to convert the continuous suitable habitat from different periods to a presence/absence distribution [103, 104]. The distribution centroids of *Q. glauca* in the LGM, present, and future periods were calculated and compared using SDMToolbox [105].

Population sampling and genotyping

A total of 781 individuals were sampled from 77 populations, which covered nearly the entire range of subtropical China (Fig. 1). Sample identification was performed by Xiao-Long Jiang and Min Deng, and voucher specimens of each individual were preserved in the Herbarium of Shanghai Chenshan Botanical Garden (CSH). Fresh and healthy mature leaves were collected and dried with silica gel to preserve them until DNA extraction could be performed. Genomic DNA was extracted using a modified cetyltrimethylammonium bromide (CTAB) protocol [106]. Seven *n*SSR loci were genotyped in 620 individuals from 60 populations (Table 1). Primer sequences and amplification conditions for each primer set were described in Table S3. For the cp.DNA, a total of 534 samples from 57 populations were sequenced (Table 1), including 157 individuals (18 populations) newly collected and 377 samples (39 populations) used from a previous study [35]. Three primer pairs, *psbA-trnH* [107], *trnT-trnL* [108], and *atpI-atpH* [109], were amplified and sequenced following previous studies (Table S3) [35]. Successfully amplified PCR products of cp.DNA and *n*SSR markers were sequenced or genotyped by the Shanghai Majorbio Bio-pharm Technology Co., Ltd (Shanghai, China). All cp.DNA sequences were assembled and manually assessed for quality in Sequencher v5.4.6 (Gene Codes Corp., Ann Arbor, MI, United States), then aligned with Bioedit v7.2.5 [110]. We determined *n*SSR genotypes using GeneMapper v4.1 [111].

Population genetic analysis based on *n*SSR

Variation coefficients of *n*SSR data for each locus and population were calculated by GenAlEx v6.5 [112], including H_e (expected heterozygosity) and A (number of alleles) for loci, H_e , H_o (observed heterozygosity) for each population. To avoid bias from unbalanced sample sizes, A_r (allelic richness) for each population was calculated by rarefaction (here, $N=10$) using HP-Rare [113].

We used Bayesian clustering and PCoA to analyze the genetic structure of *Q. glauca*. For the former, we used STRUCTURE v2.3.4 [114] with 100,000 burn-in generations followed by 200,000 MCMC iterations. The number of clusters (K) varied from 1 to 10 with ten repetitions for each K value. The best fit K was determined by STRUCTURE HARVESTER [115] using both ΔK and $\ln \Pr(X|K)$ method [116, 117]. The ten repetitions with the optimum K value were aligned using CLUMPP v1.1.2 [118] based on a greedy algorithm. We performed the PCoA using

GenALEx v6.5 based on the genetic distances among populations. The first two principal coordinates (Gp1 and Gp2) at the population level were visualized in R. We used a Student's t-test in the “*dplyr*” package [119] to test these for significance. To examine whether there was a significant correlation between genetic divergence and geographic distance (i.e., isolation-by-distance; IBD), we used a Mantel test in GenALEx based on *n*SSR data, with 9,999 permutations. Genetic distance was estimated as $F_{ST}/(1-F_{ST})$ implemented on samples grouped by locality. Geographic distance was log transformed to account for dispersal in a two-dimensional habitat [120]. The correlation between geographical and genetic distance was plotted, and the correlation coefficient (r) and R^2 were estimated using GenALEx v6.5.

cp.DNA analysis

Haplotype diversity (H_d) and nucleotide diversity (π) of *Q. glauca* populations were calculated by ARLEQUIN v3.5 [121]. The median-joining network of haplotypes was constructed using NETWORK v10.2.8 [122]. Each indel and inversion was treated as a single mutation. The geographical distribution of the cp.DNA haplotypes was mapped with ArcGIS v10.8. Alleles in Space [123] was used to examine the distribution of genetic differentiation across species ranges according to cp.DNA and *n*SSR data. We used the Landscape Genetics GIS Toolbox [124] in ArcGIS v10.8 to create genetic landscape surfaces through interpolation based on genetic diversity (H_d , π , H_e , and A_r). The software DnaSP v6.0 [125] was used to calculate G_{ST} and N_{ST} . A significantly larger N_{ST} than G_{ST} implies the presence of a significant phylogeographic structure [126].

Habitat analyses

Climatic variation and habitat complexity have well-defined roles in influencing species demography and occupancy, and both have been used to project species distribution dynamics under climate change [127]. To estimate the correlation of genetic diversity (A_r , H_e) and genetic structure (cluster A , C_a) with climate and geographic (latitude, longitude) factors, a general linear model (GLM) analysis was performed in R. Four variables, E_{Pre} , E_{Stab} , latitude, and longitude, were used as explanatory covariates. The E_{Stab} was calculated using the formula $E_{Stab}=1-|E_{Pre}-E_{LGM}|$, considering E_{Pre} and E_{LGM} as habitat suitability at the present and LGM periods, respectively. These values were obtained using the “Extract by Points” tool in ArcGIS v10.8 from *Q. glauca* predicted distribution layers [11]. The initial model included all covariates, and the most suitable models were determined based on a backward elimination procedure.

Combined with our projections of suitable habitat and the spatial distribution of haplotypes, potential dispersal routes of the *Q. glauca* populations were inferred using the LCP function in the SDMToolbox [128]. To calculate least-cost corridors (LCCs), the LCP was weighted by resistance values with cutoffs for inclusion into high-, mid-, and low-classes set at 5%, 2%, and 1%, respectively. The weighted and categorized LCPs were then summed to create a LCCs dispersal network [129].

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-024-04830-1>.

Supplementary Material 1

Acknowledgements

We would like to thank Dr. Daniel Petticord at the University of Cornell for his assistance with the English language and grammatical editing of the manuscript. We are grateful to the Asian Elephant Yunnan Field Scientific Observation and Research Station, Yunnan Asian Elephant Field Scientific Observation and Research Station of the Ministry of Education, for their help on the fieldwork.

Author contributions

XLJ and GBX designed the research. MD and XLJ were responsible for field collections and specimen identification. XLJ, YS, CCW, RC and KXL conducted the experiments and data analyses. YS, XLJ, and HL wrote the manuscript. All authors contributed to the revised manuscript.

Funding

This work was supported by grants from the Natural Science Foundation of Hunan Province (2022JJ40861), the National Natural Science Foundation of China (31700174 and 31972858), and the Scientific Research Fund of Hunan Provincial Education Department (21A0158).

Data availability

The haplotype sequences used in this study were submitted to NCBI under accession numbers PP213780-PP213811. Meanwhile, the combined sequences of 58 haplotypes from three cpDNA sequences, *n*SSR data and code of SDM analysis submitted to FigShare (<https://doi.org/10.6084/m9.figshare.24798462.v2>).

Declarations

Ethics approval and consent to participate

All plant materials used in this study were obtained from wild plants and did not involve any endangered or protected species. The ecological studies and sample collection were permitted by the local government following the Regulations on the Protection of Wild Plants of China. The use of plants in this study complies with all local, national or international guidelines and legislation concerning research involving plants.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 13 December 2023 / Accepted: 16 February 2024

Published online: 04 March 2024

References

- Hewitt GM. Speciation, hybrid zones and phylogeography—or seeing genes in space and time. *Evol Appl*. 2001;10:537–49.
- Wang W, Xiang XG, Xiang KL, Ortiz RDC, Jabbour F, Chen ZD. A dated phylogeny of Lardizabalaceae reveals an unusual long-distance dispersal across the Pacific Ocean and the rapid rise of east Asian subtropical evergreen broadleaved forests in the late Miocene. *Cladistics*. 2020;36:447–57.
- Morris AB, Shaw J. Markers in time and space: a review of the last decade of plant phylogeographic approaches. *Mol Ecol*. 2018;27:2317–33.
- Weiskopf SR, Rubenstein MA, Crozier LG, Gaichas S, Griffis R, Halofsky JE, et al. Climate change effects on biodiversity, ecosystems, ecosystem services, and natural resource management in the United States. *Sci Total Environ*. 2020;733:137782.
- Céline B, Cleo B, Paul L, Wilfried T, Franck C. Impacts of climate change on the future of biodiversity. *Ecol Lett*. 2012;15:365–77.
- Chen IC, Hill JK, Ohlemüller R, Roy DB, Thomas CD. Rapid range shifts of species associated with high levels of climate warming. *Science*. 2011;333:1024–26.
- Warren R, VanDer-Wal J, Price J, Welbergen JA, Atkinson I, Ramirez-Villegas J, et al. Quantifying the benefit of early climate change mitigation in avoiding biodiversity loss. *Nat Clim Change*. 2013;3:678–82.
- Urban MC. Accelerating extinction risk from climate change. *Science*. 2015;348:571–73.
- Pacifici M, Foden WB, Visconti P, Watson JEM, Butchart SHM, Kovacs KM, et al. Assessing species vulnerability to climate change. *Nat Clim Change*. 2015;5:215–24.
- Kou Y, Cheng S, Tian S, Li B, Fan D, Chen Y, et al. The antiquity of *Cyclocarya paliurus* (Juglandaceae) provides new insights into the evolution of relict plants in subtropical China since the late early miocene. *J Biogeogr*. 2016;43:351–60.
- Jiang XL, Deng M, Li Y. Evolutionary history of subtropical evergreen broadleaved forest in Yunnan Plateau and adjacent areas: an insight from *Quercus schottkyana* (Fagaceae). *Tree Genet Genomes*. 2016;12:106.
- Harrison TM, Copeland P, Kidd WS, Yin A. Raising tibet. *Science*. 1992;255:1663–70.
- Replumaz A, Tapponnier P. Reconstruction of the deformed collision zone between India and Asia by backward motion of lithospheric blocks. *J Geophys Res: Solid Earth*. 2003;108:2285.
- Guo ZT, Ruddiman WF, Hao QZ, Wu HB, Qiao YS, Zhu RX, et al. Onset of Asian desertification by 22 myr ago inferred from loess deposits in China. *Nature*. 2002;416:159–63.
- An Z, John EK, Warren LP, Stephen CP. Evolution of Asian monsoons and phased uplift of the Himalaya–Tibetan plateau since late miocene times. *Nature*. 2001;411:62–6.
- Hewitt G. The genetic legacy of the Quaternary ice ages. *Nature*. 2000;405:907–13.
- Qiu YX, Fu CX, Comes HP. 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 Phylo Evol*. 2011;59:225–44.
- Qin HT, Möller M, Milne R, Luo YH, Zhu GF, Li DZ, et al. Multiple paternally inherited chloroplast capture events associated with *Taxus* speciation in the Hengduan Mountains. *Mol Phylog Evol*. 2023;189:107915.
- Birks HJB. Some reflections on the refugium concept and its terminology in historical biogeography, contemporary ecology and global-change biology. *Biodiversity*. 2015;16:196–212.
- Liu KB. Quaternary history of the temperate forests of China. *Quat Sci Rev*. 1988;7:1–20.
- Shi Y, Ren B, Wang J, Derbyshire E. Quaternary glaciation in China. *Quat Sci Rev*. 1986;5:503–7.
- Wu ZY. *Vegetation of China*. Beijing: Science; 1980.
- Kremer A, Hipp AL. Oaks: an evolutionary success story. *New Phytol*. 2020;226:987–1011.
- Crowl AA, Manos PS, McVay JD, Lemmon AR, Lemmon EM, Hipp AL. Uncovering the genomic signature of ancient introgression between white oak lineages (*Quercus*). *New Phytol*. 2020;226:1158–70.
- Leroy T, Louvet JM, Lalanne C, Le Provost G, Labadie K, Aury JM, et al. Adaptive introgression as a driver of local adaptation to climate in European white oaks. *New Phytol*. 2020;226:1171–82.
- Liang YY, Shi Y, Yuan S, Zhou BF, Chen XY, An QQ, et al. Linked selection shapes the landscape of genomic variation in three oak species. *New Phytol*. 2022;233:555–68.
- Yuan S, Shi Y, Zhou BF, Liang YY, Chen XY, An QQ, et al. Genomic vulnerability to climate change in *Quercus acutissima*, a dominant tree species in east Asian deciduous forests. *Mol Ecol*. 2023;32:1639–55.
- Plomion C, Aury JM, Amselem J, Leroy T, Murat F, Duplessis S, et al. Oak genome reveals facets of long lifespan. *Nat Plants*. 2018;4:440–52.
- Luo CS, Li TT, Jiang XL, Song Y, Fan TT, Shen XB et al. High-quality haplotype-resolved genome assembly for ring-cup oak (*Quercus glauca*) provides insight into oaks demographic dynamics. *Mol Ecol Resour*. 2023;e13914.
- Li Y, Wang L, Zhang X, Kang H, Liu C, Mao L, et al. Extensive sharing of chloroplast haplotypes among east Asian *Quercus* species: the imprints of shared ancestral polymorphism and introgression. *Ecol Evol*. 2022;12:e9142.
- Du FK, Qi M, Zhang YY, Petit RJ. Asymmetric character displacement in mixed oak stands. *New Phytol*. 2022;236:1212–24.
- Fu R, Zhu Y, Liu Y, Feng Y, Lu RS, Li Y, et al. Genome-wide analyses of introgression between two sympatric Asian oak species. *Nat Ecol Evol*. 2022;6:924–35.
- Chen X, Wang X, Song Y. Genetic diversity and differentiation of *Cyclobalanopsis glauca* populations in East China. *Acta Bot Sin*. 1997;39:149–55.
- Huang SS, Hwang SY, Lin TP. Spatial pattern of chloroplast DNA variation of *Cyclobalanopsis glauca* in Taiwan and East Asia. *Mol Ecol*. 2002;11:2349–58.
- Xu J, Deng M, Jiang XL, Westwood M, Song YG, Turkington R. Phylogeography of *Quercus glauca* (Fagaceae), a dominant tree of east Asian subtropical evergreen forests, based on three chloroplast DNA interspace sequences. *Tree Genet Genomes*. 2015;11:805.
- Tian XY, Ye JW, Wang TM, Bao L, Wang HF. Different processes shape the patterns of divergence in the nuclear and chloroplast genomes of a relict tree species in East Asia. *Ecol Evol*. 2020;10:4331–42.
- Ye JW, Bai WN, Bao L, Wang TM, Wang HF, Ge JP. Sharp genetic discontinuity in the aridity-sensitive *Lindera obtusiloba* (Lauraceae): solid evidence supporting the Tertiary floral subdivision in East Asia. *J Biogeogr*. 2017;44:2082–95.
- Araújo MB, New M. Ensemble forecasting of species distributions. *Trends Ecol Evol*. 2007;22:42–7.
- Ramirez-Reyes C, Nazeri M, Street G, Jones-Farrand DT, Vilella FJ, Evans KO. Embracing ensemble species distribution models to inform At-Risk species Status assessments. *J Fish Wildl Manag*. 2021;12:98–111.
- Toro M, Caballero A. Characterization and conservation of genetic diversity in subdivided populations. *Philos Trans R Soc B*. 2005;360:1367–78.
- Zheng SS, Jiang XL, Huang QJ, Deng M. Historical Dynamics of Semi-humid Evergreen Forests in the Southeast Himalaya Biodiversity Hotspot: a case study of the *Quercus Franchetii* Complex (Fagaceae). *Front Plant Sci*. 2021;12:774232.
- Xu J, Song YG, Deng M, Jiang XL, Zheng SS, Li Y. Seed germination schedule and environmental context shaped the population genetic structure of subtropical evergreen oaks on the Yun-Gui Plateau, Southwest China. *Heredity*. 2020;124:499–513.
- Di X. Population Genetic of *Quercus aliena* based on cpDNA and SSR marker. Xi-An: Northwest University; 2017.
- Chen XD, Yang J, Feng L, Zhou T, Zhang H, Li HM, et al. Phylogeography and population dynamics of an endemic oak (*Quercus fabri* Hance) in subtropical China revealed by molecular data and ecological niche modeling. *Tree Genet Genomes*. 2020;16:1–13.
- Zhang X, Li Y, Liu C, Xia T, Zhang Q, Fang Y. Phylogeography of the temperate tree species *Quercus acutissima* in China: inferences from chloroplast DNA variations. *Biochem Syst Ecol*. 2015;63:190–97.
- Meng L, Chen G, Li Z, Yang Y, Wang Z, Wang L. Refugial isolation and range expansions drive the genetic structure of *Oxyria Sinensis* (Polygonaceae) in the Himalaya-Hengduan Mountains. *Sci Rep*. 2015;5:1–14.
- Deli T, Kiel C, Schubart CD. Phylogeographic and evolutionary history analyses of the warty crab *Eriphia verrucosa* (Decapoda, Brachyura, Eriphiidae) unveil genetic imprints of a late pleistocene vicariant event across the Gibraltar Strait, erased by postglacial expansion and admixture among refugial lineages. *BMC Evol Biol*. 2019;19:1–20.
- Contreras-Moreira B, Serrano-Notivol R, Mohammed NE, Cantalapedra CP, Bequería S, Casas AM, et al. Genetic association with high-resolution climate data reveals selection footprints in the genomes of barley landraces across the Iberian Peninsula. *Mol Ecol*. 2019;28:1994–2012.
- Chen D, Zhang X, Kang H, Sun X, Yin S, Du H, et al. Phylogeography of *Quercus variabilis* based on chloroplast DNA sequence in East Asia: multiple glacial refugia and mainland-migrated island populations. *PLoS ONE*. 2012;7:e47268.
- Li Y, Zhang X, Fang Y. Landscape features and climatic forces shape the genetic structure and evolutionary history of an oak species (*Quercus chenii*) in East China. *Front Plant Sci*. 2019;10:1060.

51. Petit RJ, Brewer S, Bordács S, Burg K, Cheddadi R, Coart E, et al. Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. *Ecol Manage.* 2002;156:49–74.
52. Petit RJ, Excoffier L. Gene flow and species delimitation. *Trends Ecol Evol.* 2009;24:386–93.
53. Wolfe KH, Li WH, Sharp PM. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *PNAS.* 1987;84:9054–58.
54. Birky C Jr, Fuerst P, Maruyama T. Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics.* 1989;121:613–27.
55. Gugger PF, Ikegami M, Sork VL. Influence of late quaternary climate change on present patterns of genetic variation in valley oak, *Quercus lobata* Née. *Mol Ecol.* 2013;22:3598–612.
56. Anderson CD, Epperson BK, FORTIN MJ, Holderegger R, James PM, Rosenberg MS, et al. Considering spatial and temporal scale in landscape-genetic studies of gene flow. *Mol Ecol.* 2010;19:3565–75.
57. Epps CW, Keyghobadi N. Landscape genetics in a changing world: disentangling historical and contemporary influences and inferring change. *Mol Ecol.* 2015;24:6021–40.
58. Graham CH, Moritz C, Williams SE. Habitat history improves prediction of biodiversity in rainforest fauna. *PNAS.* 2006;103:632–6.
59. Araújo MB, Nogués-Bravo D, Diniz-Filho JAF, Haywood AM, Valdes PJ, Rahbek C. Quaternary climate changes explain diversity among reptiles and amphibians. *Ecography.* 2008;31:8–15.
60. Ortego J, Riordan EC, Gugger PF, Sork VL. Influence of environmental heterogeneity on genetic diversity and structure in an endemic southern Californian oak. *Mol Ecol.* 2012;21:3210–23.
61. Jiang XL, Xu GB, Deng M. Spatial genetic patterns and distribution dynamics of the rare oak *Quercus Chungii*: implications for biodiversity conservation in southeast China. *Forests.* 2019;10:821.
62. Jiang XL, An M, Zheng SS, Deng M, Su ZH. Geographical isolation and environmental heterogeneity contribute to the spatial genetic patterns of *Quercus kerrii* (Fagaceae). *Heredity.* 2018;120:219–33.
63. Hampe A, Petit RJ. Conserving biodiversity under climate change: the rear edge matters. *Ecol Lett.* 2005;8(5):461–7.
64. López-Pujol J, Zhang FM, Sun HQ, Ying TS, Ge S. Centres of plant endemism in China: places for survival or for speciation? *J Biogeogr.* 2011;38:1267–80.
65. Yan HF, Zhang CY, Wang FY, Hu CM, Ge XJ, Hao G. Population expanding with the phalanx model and lineages split by environmental heterogeneity: a case study of *Primula Obconica* in subtropical China. *PLoS ONE.* 2012;7:e41315.
66. Qiu YX, Guan BC, Fu CX, Comes HP. Did glacials and/or interglacials promote allopatric incipient speciation in east Asian temperate plants? Phylogeographic and coalescent analyses on refugial isolation and divergence in *Dyosma Versipellis*. *Mol Phylogenet Evol.* 2009;51:281–93.
67. Wang C, Wang T, Su Y. Phylogeography of *Cephalotaxus oliveri* (Cephalotaxaceae) in relation to habitat heterogeneity, physical barriers and the uplift of the Yungui Plateau. *Mol Phylogenet Evol.* 2014;80:205–16.
68. Zhong Y, Yang A, Liu S, Liu L, Li Y, Wu Z, et al. RAD-Seq data point to a distinct split in *Liriodendron* (Magnoliaceae) and obvious east–west genetic divergence in *L. Chinense*. *Forests.* 2018;10:13.
69. Valladares F, Matesanz S, Guilhaumon F, Araújo MB, Balaguer L, Benito-Garazón M, et al. The effects of phenotypic plasticity and local adaptation on forecasts of species range shifts under climate change. *Ecol Lett.* 2014;17:1351–64.
70. Aitken SN, Yeaman S, Holliday JA, Wang T, Curtis-McLane S. Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evol Appl.* 2008;1:95–111.
71. Gavin DG, Fitzpatrick MC, Gugger PF, Heath KD, Rodríguez-Sánchez F, Dobrowski SZ, et al. Climate Refugia: joint inference from fossil records, species distribution models and phylogeography. *New Phytol.* 2014;204:37–54.
72. Manel S, Schwartz MK, Luikart G, Taberlet P. Landscape genetics: combining landscape ecology and population genetics. *Trends Ecol Evol.* 2003;18:189–97.
73. Waltari E, Hijmans RJ, Peterson AT, Nyári AS, Perkins SL, Guralnick RP. Locating pleistocene refugia: comparing phylogeographic and ecological niche model predictions. *PLoS ONE.* 2007;2:e563.
74. Willis KJ, Bennett KD, Walker D, Hewitt GM. Genetic consequences of climatic oscillations in the Quaternary. *Philos Trans R Soc B.* 2004;359:183–95.
75. Ni J. Impacts of climate change on Chinese ecosystems: key vulnerable regions and potential thresholds. *Reg Environ Change.* 2011;11:49–64.
76. Li X, Li Y, Fang Y. Prediction of potential suitable distribution areas of *Quercus fabri* in China based on an optimized maxent model. *Scientia Silvae Sinicae.* 2018;08:153–14.
77. Zhang XW, Li Y, Fang YM. Geographical distribution and prediction of potential ranges of *Quercus acutissima* in China. *Northwest J Bot.* 2014;8:1685–92.
78. Tian S, Kou Y, Zhang Z, Yuan L, Li D, López-Pujol, Jet. 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.* 2018;18:1–12.
79. Chen Y, Li H, Sun W, Ireland T, Tian X, Hu Y, et al. Generation of late mesozoic Qianlishan A2-type granite in Nanling Range, South China: implications for Shizhuyuan W–Sn mineralization and tectonic evolution. *Lithos.* 2016;266–267:435–52.
80. Lu LM, Mao LF, Yang T, Ye JF, Liu B, Li HL, et al. Evolutionary history of the angiosperm flora of China. *Nature.* 2018;554:234–8.
81. Luo Y, Zhou ZK. Phylogeography of *Quercus* subg. *Cyclobalanopsis* Acta Bot Yunnanica. 2001;23:1–16.
82. Gou S. Pliocene flora of western Sichuan. *Acta Palaeontol Sin.* 1978;17:343–9.
83. Gou S. The late Miocene Bangmai flora from Lincang county of Yunnan, southwestern China. *Acta Palaeontol Sin.* 2011;50:353–408.
84. Sibuet JC, Hsu SK. Geodynamics of the Taiwan arc-arc collision. *Tectonophysics.* 1997;274:221–51.
85. Sibuet JC, Hsu SK. How was Taiwan Created? *Tectonophysics.* 2004;379:159–81.
86. Chen Z, Ying J, Lu A. Disjunct distribution of seed plants between southwestern China and Taiwan Island of China. *Chi Bull Bot.* 2013;47:551–70.
87. Tsun Y. An analysis of the flora of seed plants of Taiwan, China: its nature, characteristics, and relations with the flora of the mainland. In: 2002.
88. Lu A. Origin diversification and relationships of primitive angiosperms on both sides of the Taiwan Strait, China. *Acta Bot Yunnanica.* 2001;23:269–77.
89. Ye JW, Yang ZZ, Tian B. Tempo-spatial evolution of seed plant endemism in Taiwan island. *J Biogeogr.* 2023;50:1981–91.
90. Jiang XL, Gardner EM, Meng HH, Deng M, Xu GB. Land bridges in the Pleistocene contributed to flora assembly on the continental islands of South China: insights from the evolutionary history of *Quercus championii*. *Mol Phylogenet Evol.* 2019;132:36–45.
91. Wang YH, Comes HP, Cao YN, Guo R, Mao YR, Qiu YX. Quaternary climate change drives allo-peripatric speciation and refugial divergence in the *Dyosma versipellis-pleiantha* complex from different forest types in China. *Sci Rep.* 2017;7:40261.
92. Bai WN, Wang WT, Zhang DY. Contrasts between the phylogeographic patterns of chloroplast and nuclear DNA highlight a role for pollen-mediated gene flow in preventing population divergence in an east Asian temperate tree. *Mol Phylogenet Evol.* 2014;81:37–48.
93. Booth TH, Nix HA, Busby JR, Hutchinson MF. Bioclim: the first species distribution modelling package, its early applications and relevance to most current MaxEnt studies. *Divers Distrib.* 2014;20:1–9.
94. Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. Very high resolution interpolated climate surfaces for global land areas. *Int J Biometeorol.* 2005;25:1965–78.
95. Fick SE, Hijmans RJ. WorldClim 2. New 1-km spatial resolution climate surfaces for global land areas. *Int J Biometeorol.* 2017;37:4302–15.
96. Jiang XL, Su ZH, Xu GB, Deng M. Genomic signals reveal past evolutionary dynamics of *Quercus schottkyana* and its response to future climate change. *J Syst Evol.* 2021;59:985–97.
97. Martins K, Gugger PF, Llanderal-Mendoza J, et al. Landscape genomics provides evidence of climate-associated genetic variation in Mexican populations of *Quercus rugosa*. *Evol Appl.* 2018;11:1842–58.
98. Hijmans R, Phillips S, Leathwick J, Elith J. Dismo: species distribution modelling. R package version 0.7–23. In: R Foundation for Statistical Computing Vienna, Austria; 2012.
99. Thuiller W, Lafourcade B, Engler R, Araújo MB. BIOMOD—a platform for ensemble forecasting of species distributions. *Ecography.* 2009;32:369–73.
100. Barbet-Massin M, Jiguet F, Albert CH, Thuiller W. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods Ecol Evol.* 2012;3:327–38.
101. Iturbide M, Bedia J, Herrera S, Hierro OD, Pinto M, Gutiérrez JM. A framework for species distribution modelling with improved pseudo-absence generation. *Ecol Model.* 2015;312:166–74.
102. Marmion M, Parviainen M, Luoto M, Heikkinen RK, Thuiller W. Evaluation of consensus methods in predictive species distribution modelling. *Divers Distrib.* 2009;15:59–69.

103. Grenouillet G, Buisson L, Casajus N, Lek S. Ensemble modelling of species distribution: the effects of geographical and environmental ranges. *Ecography*. 2011;34:9–17.
104. Liu C, Newell G, White M. On the selection of thresholds for predicting species occurrence with presence-only data. *Ecol Evol*. 2016;6:337–48.
105. Brown JL, Bennett JR, French CM. SDMtoolbox 2.0: the next generation Python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *PeerJ*. 2017, 5.
106. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemistry*. 1980;11–3.
107. Joey S, Lickey EB, Beck JT, Farmer SB, Liu WS, Miller J, et al. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Am J Bot*. 2005;92:142–66.
108. Taberlet P, Gielly L, Pautou G, Bouvet J. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Pl Molec Biol*. 1991;17:1105–09.
109. Grivet D, Heinze B, Vendramin G, Petit R. Genome walking with consensus primers: application to the large single copy region of chloroplast DNA. *Molec Ecol Notes*. 2001;1:345–9.
110. Hall T. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. In: *Nucleic acids symposium series*: 1999. [London]: Information Retrieval Ltd., c1979-c2000.: 95–8.
111. Rinehart TA. AFLP analysis using GeneMapper® software and an Excel® macro that aligns and converts output to binary. *Biotechniques*. 2004;37:186–8.
112. Peakall R, Smouse PE. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*. 2012;28:2537–39.
113. Kalinowski ST. Hp-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol Ecol Notes*. 2005;5:187–9.
114. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics*. 2000;155:945–59.
115. Earl DA, VonHoldt BM. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour*. 2012;4:359–61.
116. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*. 2005;14:2611–20.
117. Pritchard J, Wen X, Falush D. Documentation for STRUCTURE software: version 2.3. 2003.
118. Jakobsson M, Rosenberg NA. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*. 2007;23:1801–06.
119. Wickham H, François R, Henry L, Müller K. dplyr. A Grammar of Data Manipulation. R package version 0.7. 6. Computer software. <https://CRAN.R-project.org/package=dplyr> 2018.
120. Rousset F. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*. 1997;145:1219–28.
121. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour*. 2010;10:564–7.
122. Bandelt HJ, Forster P, Röhl A. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol*. 1999;16:37–48.
123. Miller MP. Alleles in space (AIS): computer software for the joint analysis of interindividual spatial and genetic information. *J Hered*. 2005;96:722–4.
124. Vandergast AG, Perry WM, Lugo RV, Hathaway SA. Genetic landscapes GIS toolbox: tools to map patterns of genetic divergence and diversity. *Mol Ecol Resour*. 2011;11:158–61.
125. Julio R, Albert FM, Carlos SDJ, Sara GR, Pablo L. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol Biol Evol*. 2017;34:3299–330.
126. Pons O, Petit RJ. Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics*. 1996;144:1237–45.
127. Oliver TH, Morecroft MD. Interactions between climate change and land use change on biodiversity: attribution problems, risks, and opportunities. *Wires Clim Change*. 2014;5:317–35.
128. Brown JL. SDM toolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods Ecol Evol*. 2014;5:694–700.
129. Jensen AM, O'Neil NP, Iwaniuk AN, Burg TM. Landscape effects on the contemporary genetic structure of Ruffed Grouse (*Bonasa umbellus*) populations. *Ecol Evol*. 2019;9:5572–92.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.