the **genetics**society

ARTICLE

Seed germination schedule and environmental context shaped the population genetic structure of subtropical evergreen oaks on the Yun-Gui Plateau, Southwest China

Jin Xu [1,2] · Yi-Gang Song[1,3] · Min Deng[1,3] · Xiao-Long Jiang[1,3] · Si-Si Zheng[1,2] · Ying Li[1,2]

Abstract

The evergreen broadleaved forests (EBLFs) of Southwest China have a long-term stable environment and support a diverse flora, thus forming a global biodiversity hotspot. However, the key drivers that have shaped the genetic diversity patterns of species in these EBLFs are not well understood. *Quercus delavayi*, *Q. schottkyana*, and *Q. kerrii* are sympatric oak species with different seed biological traits that are typical for these EBLFs. This study combined multilocus phylogeography and ecological niche modeling to screen 33 *Q. delavayi* populations. Their population genetic structure was inferred in comparison with previous studies on *Q. schottkyana* and *Q. kerrii*. The seed germination traits of all three species were also investigated. cpDNAs showed a significant phylogeographic structure in *Q. delavayi*, which was not detected in *Q. schottkyana* or *Q. kerrii*. *Quercus delavayi*, *Q. kerrii*, and *Q. schottkyana* exhibited different pollen-to-seed migration ratios ($r = 219$, 117, and 22, respectively), which are linked to the germination schedules of acorns. The distributions of *Q. delavayi* and *Q. schottkyana* remained long-term stable since the last glacial maximum (LGM) with a similar nSSR genetic gradient change along latitude. Instead, *Q. kerrii* experienced a prominent range expansion since the LGM with genetic diversification between the East and the West of the Tanaka line due to environmental heterogeneity. These results identify seed traits and environmental heterogeneity as two key drivers that shape the population genetic structure of EBLF trees in Southwest China. These should be considered in regional forestry conservation and management.

Introduction

Southwest (SW) China is a key global biodiversity center (Myers et al. 2000) and climate refugium for many organisms; moreover, it features a high number of endemic and relic evolutionary lineages (Lu et al. 2018; Tang et al.

✉ Min Deng
   dengmin@sippe.ac.cn

1   Shanghai Chenshan Plant Science Research Center/Shanghai Chenshan Botanical Garden, Chinese Academy of Sciences, 201602 Shanghai, China

2   School of Ecological Technology and Engineering, Shanghai Institute of Technology, 201418 Shanghai, China

3   Southeast Asia Biodiversity Research Institute, Chinese Academy of Sciences, Yezin, Nay Pyi Taw 05282, Myanmar

2018). This region underwent three to four fast uplifts as a result of the collision of the Indian and Eurasian plates as well as the lateral extrusion of Indochina since the late Paleogene (Harrison et al. 1992; An et al. 2001; Liu et al. 2012, 2014). Furthermore, this region experienced significant climate change at the onset the of Asian monsoon regime since the late Oligocene (Wan et al. 2007). Therefore, the complex interplays between tectonic activities and climate drivers identify this region as an ideal model to study the macro- and microevolutionary processes that contribute to the observed high endemism and species richness (Qiu et al. 2011; Liu et al. 2012; Wang et al. 2013).

Several biogeographic boundaries have been proposed for SW China, which served as barriers to species dispersal and boost the allopatric divergence of this region (Ward 1921; Tanaka 1954; Fan et al. 2013). The tectonic-related paleodrainage basin evolution since the Neogene boosted the contemporary diversification pattern of the biota of SW China (Zhang et al. 2010; Yue et al. 2012). A further biogeographic line, the Tanaka line, is located at ~28°N, 98°E and 18°45′N, and 108°E with a series of (sub)tropical taxon

genealogy splits along its East and West sides (Li and Li 1997). The most widely accepted explanation for the existence of the Tanaka line is that it was caused by the presently differing monsoon regimes on either side since the late Pleistocene, which further increased population divergence (Fan et al. 2013). However, the current understanding of the biogeographical significance of the Tanaka line is far from sufficient, as this pattern does not universally exist in subtropical taxa in SW China, which still leaves the formation and function of the Tanaka line enigmatic.

Evergreen broadleaved forests (EBLFs) form a dominant vegetation type in SW China, span an elevation from 800 to 2300 m, and harbor a high level of biodiversity (Tang et al. 2018). Although EBLFs in SW China form a long-term stable refugium for many relic trees, recent phylogeographic studies in this region indicated that past tectonic movements, the onset of the East Asian monsoon, and changes in the environment greatly impacted the contemporary temporal-spatial diversity pattern of trees (Tian et al. 2015; Jiang et al. 2016, 2018). Nevertheless, the genetic diversity pattern of the EBLF biota is not identical, as different scenarios in subtropical lineages were detected to either agree with the paleodrainage system divergence (Zhang et al. 2010; Zhang and Sun 2011; Yue et al. 2012) or the divergence on each side of the Tanaka line (Zhang et al. 2006; Tian et al. 2015; Zhao and Gong 2015; Jiang et al. 2018; Ju et al. 2018). The most prevailing view suggests that EBLFs in SW China have distributed over a wide and stable range since the last glacial maximum (LGM), which is generally supported by species distribution models (SDMs) and demographic analyses (Zhao and Gong 2015; Jiang et al. 2016). However, few plant species in SW China experienced regional contractions and expansions (Yuan et al. 2008; Jiang et al. 2018). Therefore, it remains unclear whether the climatic factors universally imprint the genetic divergence of regional trees in the EBLFs.

With regard to another aspect, the biotic factors of the organisms are key factors that impact the temporal–spatial divergence pattern of species. For instance, two sympatric perennial shrubs of south-west Western Australia showed disparate levels of haplotype diversity and historical demographic signals. These were linked to different ecological preferences, and life history traits related to seed and pollination traits (Millar et al. 2016). However, the role of functional traits on the genetic diversity pattern of the organism has been largely overlooked in studies on the subtropic region of China. Therefore, comparative population genetic studies on multiple species of model lineages in SW China EBLFs with different functional traits and distribution patterns provide a unique chance to understand the underlying mechanisms that contribute to the genetic diversity pattern of forest trees at this biodiversity hotspot.

Oaks (genus *Quercus*) form a model clade and offer important ecosystem service and ecological functions (Cavender-Bares 2019). Moreover, oaks offer fundamental insights into the intersection of ecology and evolution, and specifically, the role of diversification in community assembly processes (Cavender-Bares 2019). They can be used to study the importance of flexibility in key functional traits when adapting to new environments. Seed traits of oaks regulate predation, hoarding strategies of rodents, seedling establishment, and the dispersal of seeds, which further impact the plant gene flow and results in different genetic structure (Nathan and Muller-Landau 2000; Jordano and Godoy 2002). Among seed traits that tightly link to their dispersal, seed size, and germination schedule have the most profound impacts on the regulation of the predation and hoarding behaviors of frugivores and their preference (Vander Wall 1995, 2001; Forget et al. 1998; Xiao et al. 2005, 2006). In general, larger acorns often contain more nutrients, but their consumption also requires more time; as a result they are often hoarded and dispersed farther away than smaller seeds (this summarizes the seed size hypothesis: Hadj-Chikh et al. 1996; Jansen et al. 2004; Xiao et al. 2005). In addition to the seed size hypothesis, the food perishability hypothesis reasons that frugivores often scatter and hoard acorns with delayed germination schedules, while consuming fast germinating seeds immediately (Hadj-Chikh et al. 1996; Steele et al. 2001a, 2001b, 2006; Chang et al. 2009; Xiao et al. 2013). Seed traits are diversified in oaks and their roles on the dispersal and adaptation of oaks have been controversial for a long time (Steele et al. 2006; Chang et al. 2009), let alone the impact of seed traits on the population genetic structure.

Quercus delavayi, *Q. schottkyana*, and *Q. kerrii* of section *Cyclobalanopsis* are dominant evergreen trees in the EBLFs of SW China. *Q. delavayi* and *Q. schottkyana* are sympatric species on the Yun-Gui Plateau (YGP, which encompasses the plateau and the surrounding regions) (Fig. S1). They commonly occur in mixed Lauraceous and pines forests at elevations of 1000–2800 m and 1500–2500 m, respectively (Huang et al. 1999). *Q. kerrii* is an Indo-China species commonly found in pine oak forests or river gorges at low-middle elevation (50–1300 m). Its northernmost distribution range overlaps with the southern distribution ranges of *Q. delavayi* and *Q. schottkyana* (Fig. S1). Among the three oaks, *Q. delavayi* and *Q. kerrii* are genetically closer (Deng et al. 2018). They have similar basal embryos close to the seed scars (Fig. 1) and typical rapidly germinating seeds. *Q. schottkyana* with apex embryo is genetically distant to *Q. delavayi* and *Q. kerrii* (Deng et al. 2018). Its seeds have a 3–4 months temperature dormancy before germination (Xia et al. 2015). Among these three species, *Q. kerrii* bears the largest acorns with diameters of ca. 20–28 mm, while *Q. delavayi* has acorns with diameters of
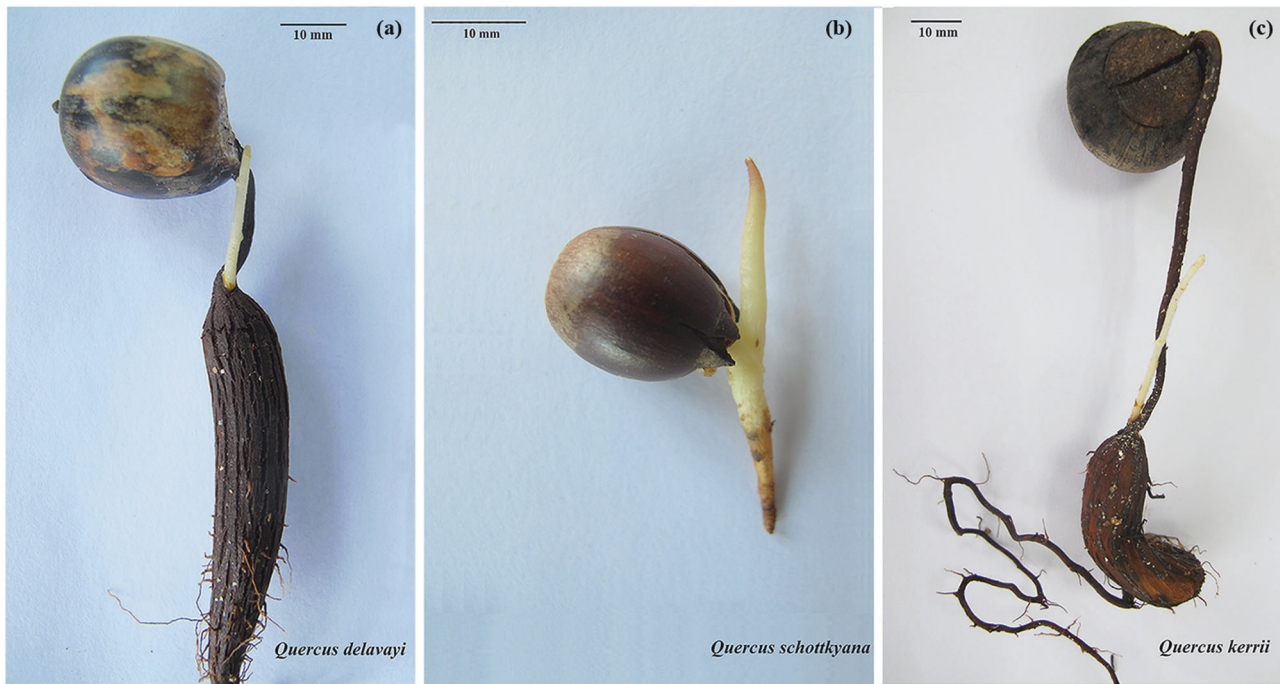
Fig. 1 Morphologies of acorn germination of (**a**) *Quercus delavayi*, (**b**) *Q. schottkyana*, and (**c**) *Q. kerrii*

10–15 mm and the acorns of *Q. schottkyana* are 7–10 mm (Huang et al. 1999).

Our previous studies on SDM and the phylogeography of *Q. kerrii* and *Q. schottkyana* showed that the distribution range of *Q. schottkyana* remained long-term stable after the LGM (Jiang et al. 2016), while *Q. kerrii* showed a northern range expansion during the interglacial period (Jiang et al. 2018). Moreover, the populations of *Q. kerrii* diverged along the Tanaka line due to longitudinal environmental heterogeneity (Jiang et al. 2018), while the population genetic structure of *Q. schottkyana* was mainly impacted by the climate gradient along the latitudinal cline (Jiang et al. 2016). Whether these observed differences on population genetic structures were caused by the climatic/geological factors or by biological traits (e.g., seed traits in particular) remains poorly understood.

This study coupled SDM and multilocus phylogeography to examine the population genetic structure and demographic dynamics of *Q. delavayi*, and compared these with *Q. schottkyana* and *Q. kerrii*. In addition, acorn sizes and germination traits of the three oaks were assessed. The objectives were: (1) to test whether these evergreen oaks of SW China responded to historical environmental changes in a similar or different manner; (2) to identify the key drivers that shaped the genetic diversity pattern of evergreen oaks in EBLFs of SW China; and (3), to explore the roles of acorn traits for the population genetic structure of regional woody species. This study presents the first case to disentangle the roles of biotic and abiotic factors for the

population genetic structures of forest trees. The results may provide comprehensive insight toward better conservation management and sustainable use of the EBLFs in SW China.

Materials and methods

Acorn size measurement and observation of germination traits

Mature acorns of *Q. delavayi*, *Q. schottkyana*, and *Q. kerrii* were collected from one population each in autumn 2015 (Table S1). The length (h) and width (d) (maximum diameter) of acorns were measured for 30 fruits per species. Scatter plots of the length and width of acorns were generated using R v3.3.3. The acorn volume (V) was estimated by the formula: $V = \pi d^2 h/4$, where π represents the ratio of the perimeter of the circle to the diameter of the circle.

Mature acorns of *Q. delavayi* and *Q. kerrii* were collected during 2014–2015 for germination tests during November and October, respectively (Table 1). The harvested seeds were X-ray scanned to identify aborted acorns and acorns that were infested with larvae. In total, 60 acorns of each species were planted in a tray with vermiculite, and kept in a growth chamber at temperatures of 15, 20, and 25 °C with a humidity of about 85–95%. Acorns were observed every 2 days until seedlings were established. Germination was defined as the emergence of a radicle of at

Table 1 Collection information and times until germination for *Quercus delavayi*, *Q. kerrii*, and *Q. schottkyana*

Species	Collection date (month/yr)	Collection location (altitude)	Time to 50% germination (T_{50}) (d)		
			15 °C	20 °C	25 °C
Q. delavayi	11/15	Puer, Yunnan (1500 m)	24[A]	13[A]	10[A]
Q. kerrii	10/14	Jinghong, Yunnan (1064 m)	18[A]	10[A]	6[A]
Q. schottkyana	11/06	Kunming, Yunnan (1900 m)	142[a]	119[a]	55[a]

The superscript letter ([A]) indicated a statistical significance; the same letters mean $P > 0.05$

[a]Data from Xia et al. (2015)

least 2 mm in length. The time to germination was the number of days acorns needed to reach 50% of maximum germination (T50 (d)) as described by Daws et al. (2008). The seed germination patterns were observed using a Canon G15 camera. The germination data of *Q. schottkyana* were obtained from Xia et al. (2015). One-way analyses of variance (ANOVAs) were used to test the differences in time to germination between *Q. delavayi* and *Q. kerrii*. However, this could not be used for the germination data of *Q. schottkyana*, since only the mean value of time to germination was available for this species. Therefore, *Q. schottkyana* was not included in ANOVAs. ANOVAs were also conducted to test the differences in acorn size among the three oak species. The statistical analyses were performed in R v3.3.3, and the significance level was set to $P < 0.05$.

Population sampling

A total of 493 samples, representing 33 populations and covering all major distribution sites of *Q. delavayi*, were collected at elevations of 1013–2760 m (Fig. 2 and Table 2). Each sampled tree was at least 30 m away from any other sampled trees. Sampled leaves were rapidly dried in silica gel until DNA extraction. Voucher specimens of each population were deposited in the herbarium of the Shanghai Chenshan (CSH) Botanical Garden. Provenance data on populations and their abbreviations are summarized in Table 2.

DNA extraction, sequencing, and microsatellite genotyping

Total genomic DNA was extracted using the modified CTAB method (Doyle and Doyle 1987). In total, the three cpDNA loci *trn*H-*psb*A (Shaw et al. 2005), *trn*T-*trn*L (Taberlet et al. 1991), and *pet*G-*trn*P (Huang et al. 2002) of 330 individuals from 33 populations were sequenced. The primers used to amplify cpDNAs are listed in Table S2. All 493 individuals were also genotyped at eight nSSR loci (Table S4). PCR were conducted to amplify the cpDNA and nSSR loci as described by Xu et al. (2016b) and An et al. (2016), respectively. PCR products of nSSR markers were mixed with fluorescence size standards at a ratio of 6-FAM:
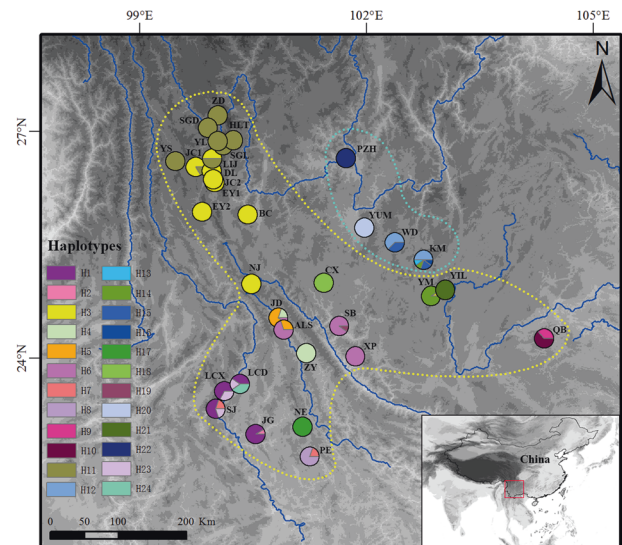

Fig. 2 Distribution of 33 populations of *Quercus delavayi* (abbreviations listed in Table 2) and 24 cpDNA haplotypes. The blue (North-central YGP) and yellow (Southwest YGP and South YGP) dashed lines delimit both population groups detected by BI analysis

HEX:ROX = 1:1:2. The PCR products of cpDNA and nSSR markers were bidirectionally sequenced or genotyped by Shanghai Majorbio Bio-pharm Technology Co., Ltd (Shanghai, China). All sequences obtained in this study have been deposited in GenBank (Table S3).

Analysis of cpDNAs

The sequence files of the bidirectionally sequenced cpDNAs were assembled and manually verified using SEQUENCHER 4.01 (Gene Codes Corp., Ann Arbor, MI, USA). Sequences were aligned with CLUSTAL_W using MEGA 6.06 (Tamura et al. 2013) and alignments were manually adjusted, where substitutions and indels were assumed to have evolved with equal weight. A 3-bp inversion was found in the *pet*G-*trn*P region and a 32-bp inversion was found in the *trn*H-*psb*A region. These were replaced with their reverse complements to maximize homology in the phylogenetic analyses. The inversions were coded as two single binary characters, comparable with an indel (Whitlock et al. 2010). Substitutions and

Table 2 Sampling sites, sample sizes, haplotype diversity (H_d), and nucleotide diversity (π) of the 33 populations of *Quercus delavayi* investigated in the Yun-Gui plateau

	Population abbre	Sample location	Latitude (°N)	Longitude (°E)	Altitude (m)	Sample size	Haplotypes (no. of individuals)	Haplotype diversity (H_d)	Nucleotide diversity ($\pi \times 10^3$)
North-Central populations	WD	Wuding, Yunnan	25.53	102.37	1890	8	Hap12(5)Hap15(3)	0.5357	0.30
	YUM	Yuanmo, Yunnan	25.73	101.97	2510	10	Hap20(10)	0.0000	0.00
	KM	Kunming, Yunnan	25.30	102.76	1013	10	Hap12(5)Hap13(1)Hap14(1)Hap15(2)Hap16(1)	0.7556	2.18
	PZH	Panzhihua, Sichuan	26.65	101.73	1865	10	Hap22(10)	0.0000	0.00
Northwest populations	JC1	Jianchuan, Yunnan	26.53	99.74	2475	12	Hap3(10)Hap11(2)	0.3030	0.17
	DL	Dali, Yunnan	26.48	99.95	2258	10	Hap3(10)	0.0000	0.00
	BC	Bingchuan, Yunnan	25.90	100.43	1862	12	Hap3(12)	0.0000	0.00
	EY1	Eryuan, Yunnan	26.33	99.99	2300	8	Hap3(8)	0.0000	0.00
	JC2	Jianchuan, Yunnan	26.37	99.97	2541	6	Hap3(6)	0.0000	0.00
	LIJ	Lijiang, Yunnan	26.64	99.96	2274	12	Hap3(6)Hap11(6)	0.5455	0.31
	SGL	Lijiang, Yunnan	26.82	100.10	2740	8	Hap11(8)	0.0000	0.00
	HLT	Lijiang, Yunnan	26.89	100.24	2442	6	Hap11(6)	0.0000	0.00
	ZD	Xianggelila, Yunnan	27.22	100.03	1983	10	Hap11(10)	0.0000	0.00
	SGD	Lijiang, Yunnan	27.05	99.90	1868	7	Hap11(7)	0.0000	0.00
	YL	Yulong, Yunnan	26.88	100.03	2159	6	Hap11(6)	0.0000	0.00
	EY2	Eryuan, Yunnan	25.94	99.83	1887	17	Hap3(17)	0.0000	0.00
	YS	Yongsheng, Yunnan	26.61	99.47	2760	9	Hap11(9)	0.0000	0.00
South populations	YM	Yimen, Yunnan	24.83	102.85	1701	11	Hap3(1)Hap14(10)	0.1818	0.10
	NJ	Nanjian, Yunnan	24.98	100.48	1845	12	Hap3(12)	0.0000	0.00
	JG	Jinggu, Yunnan	23.00	100.53	1430	11	Hap1(10)Hap2(1)	0.1818	0.31
	SB	Shuangbai, Yunnan	24.43	101.64	1013	10	Hap6(9)Hap19(1)	0.2000	0.34
	XP	Xinping, Yunnan	24.03	101.85	1225	6	Hap6(6)	0.0000	0.00
	SJ	Shuangjiang, Yunnan	23.33	100.00	1460	10	Hap1(6)Hap7(2)Hap23(2)	0.6222	0.60
	LCX	Lincang, Yunnan	23.57	100.12	1475	9	Hap1(6)Hap23(3)	0.5000	0.28
	LCD	Lincang, Yunnan	23.66	100.32	1203	5	Hap1(2)Hap23(1)Hap24(2)	0.8000	3.03
	JD	Jingdong, Yunnan	24.54	100.83	1265	10	Hap4(2)Hap5(7)Hap6(1)	0.5111	1.71
	ZY	Zhenyuan, Yunnan	24.07	101.20	1337	13	Hap4(13)	0.0000	0.00
	PE	Puer, Yunnan	22.70	101.25	1416	5	Hap7(1)Hap8(4)	0.4000	0.22
	QB	Qiubei, Yunnan	24.26	104.35	1678	11	Hap9(4)Hap10(7)	0.5091	0.29
	NE	Ninger, Yunnan	23.09	101.15	1547	10	Hap17(10)	0.0000	0.00
	CX	Chuxiong, Yunnan	25.00	101.43	2186	10	Hap18(10)	0.0000	0.00
	ALS	Ailaoshan, Yunnan	24.45	100.90	1431	10	Hap5(3)Hap6(7)	0.4667	0.26
	YIL	Yiliang, Yunnan	24.90	103.04	1886	10	Hap21(10)	0.0000	0.00
Total	—		—		—	314	—	0.8820	3.32

indels were assumed to have evolved with equal weight and chloroplast haplotypes were extracted by DnaSP 5 (Librado and Rozas 2009).

Genetic variation was measured by the average pairwise differences per base-pair between sequences (nucleotide diversity, π) (Nei and Li 1979) and haplotype diversity (H_d) using DnaSP version 5 (Librado and Rozas 2009). Total haplotype diversity (H_T), within-population diversity (H_S) (Nei 1975), and population differentiation indexes (G_{ST} and N_{ST}) were calculated by PERMUT 2.0 (Pons and Petit 1996). Population differentiation indexes (G_{ST} and N_{ST}) of Q. schottkyana were also calculated by PERMUT using data from a previous publication (Jiang et al. 2016). A haplotype network was constructed using the Median Joining model in NETWORK 5.0.0.0 (Bandelt et al. 1999). According to the geographic distribution of Q. delavayi and the network results, all 33 populations of Q. delavayi were divided into three groups (Table 2): (1) four populations from the North-Central YGP region, (2) 13 populations from the Northwest YGP region, and (3) 16 populations from the South YGP region. The variations within and among defined groups and populations were calculated using the hierarchical analysis of molecular variance (AMOVA) framework with 1000 permutations, using ARLEQUIN 3.5 (Excoffier and Lischer 2010). To examine the effect of geographic distance on the genetic structure and the relative contribution of both gene flow and drift to genetic structure (Hutchison and Templeton 1999), isolation-by-distance (IBD) was tested by Mantel tests in GENALEX 6.5 (Peakall and Smouse 2012). Possible genetic barriers were estimated using Monmonier's maximum difference algorithm (Monmonier 1973) based on the generated haplotype network, along with residual genetic distances among locations (Supplementary File S1).

The demographic history of Q. delavayi populations on the YGP was investigated using mismatch distribution analysis in ARLEQUIN 3.5 (Excoffier and Lischer 2010) to assess whether Q. delavayi populations experienced expansion in the past. Multimodal mismatch distributions of pairwise differences between haplotypes are expected for populations at a demographic equilibrium with a relatively stable size over time, while unimodal distributions are expected for populations that experienced demographic expansions (Harpending 1994). The sum of square deviations (SSD) between observed and expected distributions and the raggedness index (H_{Rag}) were used to validate the goodness of fit (Harpending 1994; Rogers and Harpending 1992) across 1000 parametric bootstrap replicates. Tajima's D (Tajima 1989) and Fu's F_S (Fu 1997) tests were calculated to discriminate mutations from drift equilibrium, and to evaluate population expansion via significant excess of low-frequency haplotypes. Significant negative values of D and F_S can be expected in expanding populations. All aforementioned demographic analyses were performed in ARLEQUIN 3.5 (Excoffier and Lischer 2010).

Phylogenetic relationships of sampled haplotypes were reconstructed using Bayesian inference (BI). Q. variabilis, Lithocarpus litseifolius, and Lithocarpus longinux were used as outgroups, and MrBayes 3.2 (Ronquist et al. 2012) was used to perform the BI analysis. A detailed description of the analytical procedure and utilized parameters is provided in the Supplementary File S1.

A Bayesian search for tree topologies and node ages of the cpDNA dataset was conducted in BEAST 1.8 (Drummond et al. 2012). A Hasegawa-Kishino-Yano (HKY) substitution model and an uncorrelated lognormal relaxed clock setting were used. The constant-size coalescent was applied as tree prior. The earliest acorn fossil of Q. sect. Cerris found in the Lower Miocene (15–17.5 Ma, Linqu County, Shandong Province, China; Song et al. 2000) was used as minimal age to constrain the stem of the haplotype tree with a median of 18 Ma and 95% highest probability density (HPD) from 14.88–34.12 Ma. For BEAST analysis, MCMC runs were performed for 1×10^8 generations and were sampled every 1000 generation, followed by a burn-in of the initial 20% cycles. MCMC samples were inspected in TRACER 1.6 (http://tree.bio.ed.ac.uk/software/tracer/) to confirm sampling adequacy and convergence of chains. The resulting chronograms were visualized and edited in FigTree 1.43 (http://tree.bio.ed.ac.uk/software/figtree/).

nSSR analysis

Descriptive statistics of nSSRs, including the number of alleles (N_A) and observed (H_o) and expected heterozygosity (H_e), were estimated using GENALEX 6.5 (Peakall and Smouse 2012). Allelic richness (A_r) was estimated using FSTAT 2.9.3.2. The genetic diversity indexes (A_r, H_o, and H_e) were also calculated at the group level using the Tanaka line as boundary to assess its western and eastern populations. The above-mentioned analyses were also conducted for Q. schottkyana based on data from a previous study (Jiang et al. 2016). IBD was tested by Mantel tests using GenAlEx 6.5 for the correlation of $F_{ST}/(1 - F_{ST})$ with geographic distance for all pairs of populations. $F_{ST}/(1 - F_{ST})$ was calculated using GenePop 4.6.9 (Rousset 2008). AMOVA analysis with 1000 permutations was conducted in ARLEQUIN 3.5 (Excoffier and Lischer 2010). The ratios (r) of pollen-to-seed flow of Q. delavayi, Q. schottkyana, and Q. kerrii were estimated using the following formula by Ennos (1994): $r = m_p/m_s = [(1/F_{ST(b)} - 1) - 2(1/F_{ST(m)} - 1)]/[1/F_{ST(m)} - 1]$. The F_{ST} values for nuclear and maternally inherited organelle markers of Q. schottkyana and Q. kerrii were calculated based on data obtained from previous studies (Jiang et al. 2016, 2018). To determine the spatial genetic structure (Supplementary File S1), the number

of genetic Bayesian clusters was inferred using STRUC-TURE 2.3.4 (Pritchard et al. 2000; Falush et al. 2003).

A general linear model was run in R 3.3.3 (http://www.r-project.org/) to estimate the correlation of genetic diversity (Ar and He) and genetic structure (cluster A, C_A) with the habitat and geographic location of populations. The habitat suitability of present (N_{Pre}) and LGM (N_{LGM}) periods, habitat stability (N_{Stab}), latitude, and longitude were used as explanatory covariates. Habitat stability (N_{Stab}) was estimated as $N_{Stab} = 1 - |N_{Pre} - N_{LGM}|$ (Ortego et al. 2012). The initial model contained all five explanatory variables related to habitat and geographic factors, and the final models were selected following a backward eliminating procedure.

Ecological distribution modeling

MAXENT 3.3.3 (Phillips et al. 2004, 2006) was used to estimate the distributions of *Q. delavayi* at present, and at the LGM (~21 Ka). In total, 61 occurrence points of *Q. delavayi* were obtained from field collections and from the Chinese Virtual Herbarium (http://www.cvh.ac.cn). After eliminating highly correlated variables ($r > 0.8$) (Kumar and Stohlgren 2009), eight climate variables, that were related to temperature and precipitation, were obtained from World-Clim (http://www.worldclim.org) at 2.5-arcmin resolution for both present and LGM (MIROC 3.2 scenario) time points. The mean value of 10 replicate results with random seeds was assumed as the potential distributions of species. The area under the curve (AUC) was used to evaluate model performance: 1 is the maximum prediction and 0.5 is considered a random prediction. The logistic output of MAX-ENT consists of a grid map where each cell has an index of habitat suitability between 0 and 1. Low values indicate that conditions are unsuitable for the species, while high values indicate suitable conditions. To quantitatively analyze the extent of species expansion, the ratio (Na) of the LGM was calculated and potential distribution areas are presented based on the maximum sensitivity plus specificity value as species presence/absence threshold. Na = (distribution areas of present)/(distribution areas of LGM) represents the extent of the distribution area change: a value close to 1 represents a scenario where the species distribution remains steady, while values below or above 1 represent scenarios where the species distribution area underwent contraction or expansion, respectively, from LGM to the present day. The annual mean temperatures (bio1) of sampling locations for the three oak species were extracted and used to construct a density plot in R v3.3.3. The sample locations of *Q. schottkyana* and *Q. kerrii* were obtained from Jiang et al. (2016) and Jiang et al. (2018), respectively. The differences between the niches that are occupied by the three oaks were evaluated using the niche overlap and niche identity test indexes: Schoener'D (D) and I statistic (I) as available in

ENMTools v1.3 (Warren et al. 2010). A detailed description of the analytical procedure is provided in Supplementary File S1.

Results

Acorn size and germination traits

The acorn width of *Q. kerrii* significantly exceeded that of *Q. delavayi* and *Q. schottkyana* ($P < 0.05$) (*Q. kerrii* = 21.64 ± 0.77 mm; *Q. delavayi* = 14.91 ± 1.05 mm; *Q. schottkyana* = 14.98 ± 0.58 mm; Table S1 and Fig. S2). *Q. kerrii* acorns were significantly shorter than the other two species ($P < 0.05$) (*Q. kerrii* = 13.75 ± 1.07 mm; *Q. delavayi* = 15.12 ± 0.81 mm; *Q. schottkyana* = 17.01 ± 0.81 mm; Table S1 and Fig. S2). The volume of *Q. kerrii* acorns was largest, followed by that of *Q. schottkyana*, and that of *Q. delavayi* (*Q. kerrii* = 5071.22 ± 595.47 mm³; *Q. delavayi* = 2667.48 ± 483.80 mm³; *Q. schottkyana* = 3008.01 ± 337.50 mm³; Table S1).

The T50 of the three species at the three temperature treatments showed great differences (Fig. S3 and Table 1). The germination pattern of *Q. schottkyana* was significantly different compared with *Q. delavayi* and *Q. kerrii* (Fig. 1). The germination schedule of *Q. schottkyana* was much later for all three temperature treatments compared with those of *Q. delavayi* and *Q. kerrii* (Fig. S3 and Table 1). *Q. schottkyana* has dormant seeds that germinated after 2 months at 25 °C, and 4–5 months at 20 °C and 15 °C. *Q. delavayi* and *Q. kerrii* germinated within 1 month under all three temperatures without significant difference in the time to germination ($P > 0.05$) (Table 1).

Phylogeny and network of haplotypes

The combined lengths of the three cpDNA fragments varied from 1792 to 1825 bp, with a consensus alignment length of 1827 bp. A total of 24 haplotypes with 23 nucleotide substitutions, five indels, and two inversions were identified from 330 samples in *Q. delavayi* (Table S5).

The phylogenetic network showed two distinct clades with nonoverlapping geographical ranges of *Q. delavayi* (Figs. 2 and 3). Clade I was distributed in North-Central YGP, and clade II was mainly distributed in Northwest and South YGP. Five median vectors were identified in the network frame. Haplotype H3 was the most common haplotype, which was found in 82 individuals from nine populations. Haplotype H11 was the second-most common haplotype, and occurred in 54 individuals from eight populations. Haplotypes H3 and H11 occurred across Northwest YGP, and haplotype H3 was also distributed in Central YGP. In total, 13 unique haplotypes were identified
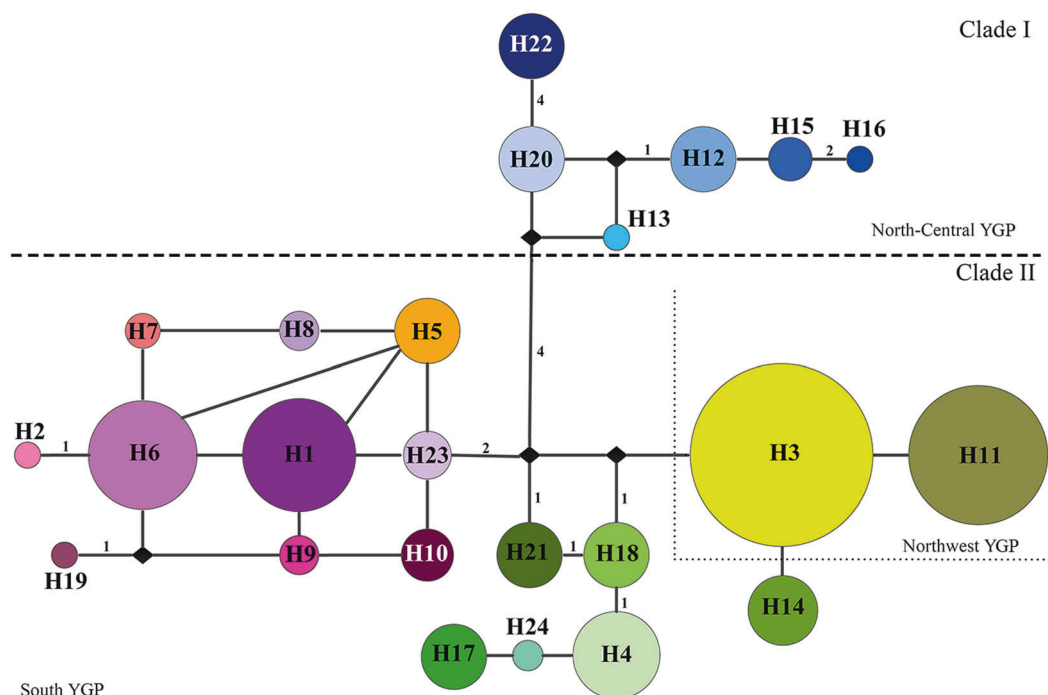
Fig. 3 Networks of 24 cpDNA haplotypes identified in *Quercus delavayi*. The circle size is proportional to the frequency of a haplotype over all populations. Each line between haplotypes represents a mutational step. Numbers at the middle of the line indicate the number of hypothetically missing haplotypes. Solid diamonds represent hypothetically unsampled haplotypes or extinct ancestral haplotypes

(Table 2 and Fig. 2) and the haplotype network had a similar haplotype structure as BEAST analysis (Fig. 4).

The BI topology (Fig. S4) was congruent with the result of the phylogenetic network. All 24 haplotypes were monophyletic with posterior probabilities (PP) of 1.00. Haplotypes from North-Central YGP (H12, H13, H15, H16, H20, and H22) were divergent and paraphyletic to a well-supported monophyletic clade of haplotypes from Northwest and South YGP (PP = 1.00).

Mismatch analysis and neutrality test

The observed mismatch distributions of pairwise nucleotide differences showed multimodal distributions for all populations (Fig. S5). The values of D and F_S were positive for *Q. delavayi* ($D = 0.675$, $P = 0.693$; $F_S = 0.196$, $P = 0.626$), suggesting that this species did not experience recent demographic expansion. However, nonsignificant SSD and H_{Rag} (SSD = 0.022, $P = 0.195$; $H_{Rag} = 0.030$, $P = 0.195$) indicated that this species might conform to a model of sudden population expansion. In summary, weak support for a model of sudden population expansion was obtained for *Q. delavayi*.

Divergence time of cpDNA haplotypes

The BEAST tree of cpDNAs showed two main clades of *Q. delavayi* haplotypes: North-Central YGP and Northwest

and South YGP (Fig. 4). The divergence time of the most recent common ancestor (TMRCA) of *Q. delavayi* haplotypes was dated to 10.92 Ma (95% HPD 5.30–17.98 Ma). The estimated crown age of clade I was dated to 4.98 Ma (95% HPD 1.80–9.49 Ma), and clade II was dated to 7.02 Ma (95% HPD 3.19–12.46 Ma).

Genetic diversity and differentiation

cpDNA results

Total nucleotide diversity (π) and haplotype diversity (H_d) of *Q. delavayi* were 3.32×10^{-3} and 0.882, respectively (Table 2). The cpDNA was most diverse in the South populations, followed by the North-Central populations, and then the Northwest populations. Populations LCD, KM, and SJ (population abbreviations are summarized in Table 2) had the highest haplotype diversity ($H_d = 0.8000$, 0.7556, and 0.6222, respectively) and nucleotide diversity ($\pi = 3.03 \times 10^{-3}$, 2.18×10^{-3}, and 0.60×10^{-3}, respectively; see Table 2). Total haplotype diversity (H_T) and within-population diversity (H_S) were 0.907 and 0.197, respectively. A significant phylogeographic signal was detected in the cpDNA of *Q. delavayi* ($N_{ST} = 0.912 > G_{ST} = 0.782$; $P < 0.05$). Hierarchical AMOVA of the cpDNA showed that 63.28% of the variation was partitioned among three geographic regions (North-Central, Northwest, and South populations), 30.77% among populations within groups, and
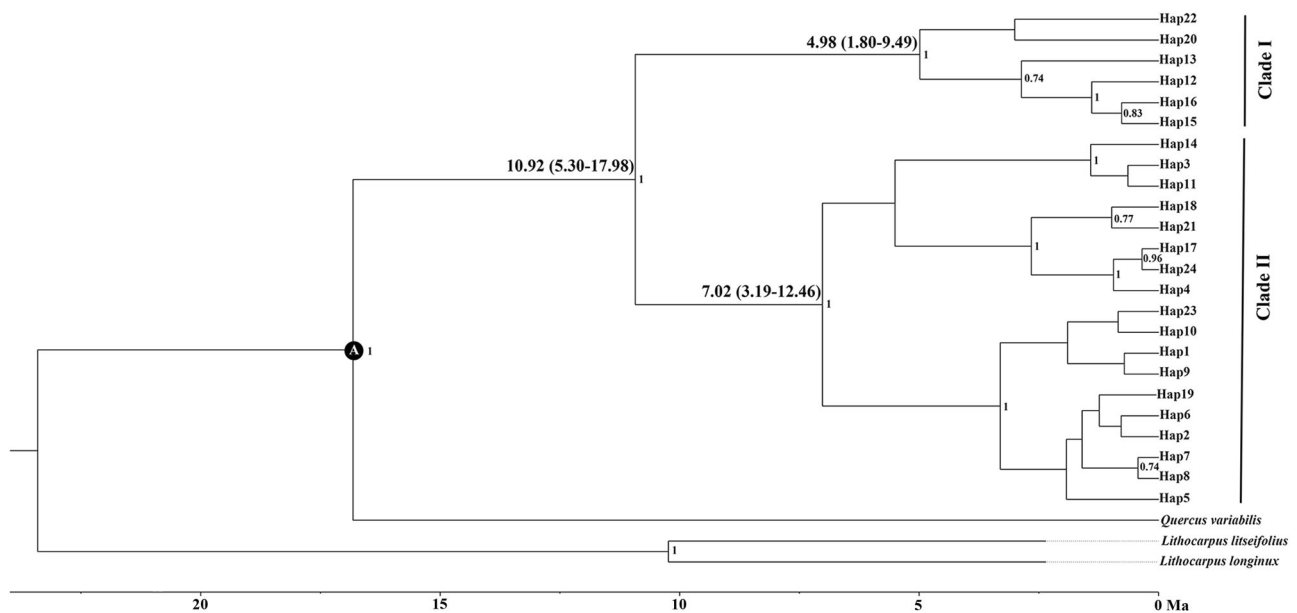
Fig. 4 Bayesian analysis chronogram of *Quercus delavayi* cpDNA haplotypes and three outgroups from genera *Quercus s.l.* and *Lithocarpus*. Fossil calibration point is marked with a letter (A) in the circle. Numbers on the right side of branches are posterior probabilities (PP). The divergence time (in Ma) and 95% highest probability density (HPD) for the most recent common ancestor (TMRCA) of sampled *Q. delavayi* populations and the two clades with 100% PP supports are shown above the node

5.95% within populations (Table 3). A Mantel test was conducted on the genetic and geographical distance matrices and a significant correlation was found ($R = 0.587$; $P = 0.001$), indicating a significant IBD level for cpDNA. A genetic barrier was detected for the cpDNA between North-Central YGP and Northwest and South YGP (Fig. 2).

nSSR results

All eight nSSR loci were highly variable, with 7–24 alleles per locus across 33 populations. In total, 114 alleles were identified. The diversity estimates of microsatellites differed among populations (Table S6). The lowest diversity was found for the PE population (in South YGP) and the highest diversity was found for the YUM population (in North-Central YGP). The genetic diversity indexes (Ar, Ho, and He) of the eastern group of *Q. delavayi* were slightly higher than in the western group of this species (Table S7). The genetic diversity indexes (Ar, Ho, and He) of both groups of *Q. schottkyana* were similar (Table S7). The largest variation in the nSSR loci (93.67%; $F_{ST} = 0.064$ in Table 3) was found within populations, and a small fraction of this variation was among populations, within lineages, and among lineages (Table 3). IBD was also detected for nSSR data ($R = 0.365$, $P = 0.001$).

Bayesian clustering analysis of the genetic structure demonstrated that the model had the highest ln Pr(X|K) and ΔK value for $K = 2$ (Fig. S6), indicating that the most probable number of groups was two. The Northwest populations (JC1, DL, BC, EY1, JC2, LIJ, SGL, HLT, ZD, SGD, YL, EY2, and YS) were primarily assigned to Cluster I (green), whereas other populations (North-Central and South populations) were assigned to Cluster II (red) (Fig. S7). STRUCTURE plots ($K = 2$–5) are shown in Fig. S8. When $K = 4$ and 5, the populations JC1 and DL differentiated from other populations in the northwestern group (Fig. S8). Ar or He was not associated with N_{Pre}, N_{LGM}, N_{Stab}, latitude, or longitude ($P > 0.15$ for all). However, the proportion of cluster A (C_A) was negatively associated with latitude (estimate ± SE: -0.226 ± 0.026, $t = -8.711$, $P < 0.001$).

Pollen-to-seed migration ratio

The pollen-to-seed migration ratios were 219 for *Q. delavayi*, 22 for *Q. schottkyana*, and 117 for *Q. kerrii* (Table 4). For comparison, Table S8 lists the pollen-to-seed migration ratios (r) of several oak species from the literature. A wide variation was found across species with r values ranging from 6.6 (*Q. rubra*) to 649 (*Q. aquifolioides*) (Table S8).

Ecological niche modeling

Model evaluation in MAXENT showed high performance scores (AUC = 0.9965, standard deviation (SD) = 0.0009). The predicted distribution, based on the model, was similar to the present-day distribution of *Q. delavayi* (Fig. 5a), containing extensive suitable ranges for the species across the YGP. The potential range of *Q. delavayi* shrank slightly and shifted to the south and east during the LGM (Fig. 5b).

Table 3 Analyses of molecular variance (AMOVA) based on cpDNA polymorphisms and nuclear microsatellite allele frequencies for populations of *Quercus delavayi*

Source of variation	cpDNA					nSSR				
	d.f.	Sum of squares	Variance components	Percentage of variation	Fixation indices	d.f.	Sum of squares	Variance components	Percentage of variation	Fixation indices
Among groups	2	435.209	2.172	55.96	$F_{CT}=0.560$**	2	45.791	0.055	1.79	$F_{CT}=0.018$**
Among populations within groups	30	424.445	1.470	37.88	$F_{SC}=0.860$**	30	209.688	0.140	4.57	$F_{SC}=0.047$**
Within populations	281	67.123	0.239	6.16	$F_{ST}=0.938$**	953	2732.273	2.867	93.64	$F_{ST}=0.064$**
Total	313	926.777	3.881	–		985	2987.753	3.062	–	

F_{CT}: correlation of haplotypes with groups relative to total; F_{SC}: correlation within populations relative to groups; F_{ST}: correlation within populations relative to total; *d.f.*: degrees of freedom

**$P < 0.001$

The largest contributor (32.30%, SD = 0.37) to identifying areas of occurrence for *Q. delavayi* was annual precipitation. Isothermality and temperature seasonality were the second (20.70%, SD = 0.32) and third largest contributors (19.85%, SD = 0.56), respectively. The rest of the variation (27.15%) was caused by all other climatic variables. The ratio of the present range to the LGM range for *Q. delavayi* was 1.14. The ranges of annual mean temperature of *Q. delavayi* were similar to that of *Q. schottkyana* (Fig. S9). The habitat of *Q. kerrii* was located at a narrower range of annual mean temperature compared with that of *Q. delavayi* and *Q. schottkyana*, and it was more often found in high-temperature zones (Fig. S9).

Niche identity test indicated a relatively high similarity of the niches that are suitable for *Q. delavayi* and *Q. schottkyana* (*P* > 0.05), and a high ecological niche separation between *Q. delavayi* and *Q. schottkyana* compared with that of *Q. kerrii* (*P* < 0.001; Figs. S10 and S11).

Discussion

Disparate patterns in cpDNAs and nSSR loci

In oaks, cpDNA is maternally inherited, and nuclear DNA is inherited by both parents. The disparate patterns observed in the cpDNA and nuclear DNA of *Q. delavayi* are the result of different gene flow efficiencies caused by the dispersal abilities of seeds and pollen. The variation in cpDNA exceeded 50% among regions (North-Central YGP, South YGP, and Northwest YGP) (Table 3), indicating limited seed-mediated gene flow in these geographic regions. However, the pattern of nSSR was very different from cpDNAs, and most variation (93.64%) occurred within populations (Table 3). The barrier between populations on the North-Central YGP and the Northwest and South YGP, as inferred by cpDNA analysis, was not detected by the nSSR analysis. In summary, the result of this analysis suggests that pollen-mediated gene flow is not restricted by the topology of the YGP region.

As a typical outbreeding wind pollination lineage, oaks generally have much higher pollen-to-seed migration ratios compared with other plants (*r* = 17) (Petit et al. 2005) (Table S8). Similarly, *Q. delavayi*, *Q. schottkyana*, and *Q. kerrii* all have high pollen-to-seed migration ratios. Male oak catkins can release a vast number of pollen grains, which are dispersed by wind across long distances (Ennos 1994). However, acorns are predominantly dispersed by gravity, small rodents, and several bird species, and the long-distance dispersal of seeds is comparatively inefficient (Ramos-Palacios et al. 2014; Scofield et al. 2011). The YGP has high environment heterogeneity, with a number of deep river gorges and high mountains, which serve as

Table 4 Comparisons of the genetic diversity, genetic structure, and demographic dynamics of *Quercus delavayi* with *Q. schottkyana* and *Q. kerrii*

| Taxa | Genetic diversity | | Genetic structure | | Phylo structure | Network structure | $F_{ST(b)}$ | $F_{ST(m)}$ | r | Na | Demographic change | |
	H_T (se)	H_S (se)	G_{ST}(se)	N_{ST} (se)							Mismatch	Neutral test
Q. delavayi	0.907 (0.03)	0.197 (0.05)	0.782 (0.05)	0.912* (0.03)	Yes	Geographic structure	0.063	0.938	219	1.14	No expansion	No expansion
Q. kerrii	0.71 (0.06)[a]	0.05 (0.02)[a]	0.93 (0.03)[a]	0.92 (0.04)[a]	No[a]	Star-like[a]	0.066	0.894	117	1.27[a]	Expansion[a]	No expansion[a]
Q. schottkyana	0.828 (0.06)	0.341 (0.06)	0.588 (0.07)	0.615 (0.11)	No	Star-like[b]	0.075	0.665	22	1.03[b]	No expansion[b]	No expansion[b]

H_T: total genetic diversity; H_S: genetic diversity within populations; G_{ST}: interpopulation differentiation; N_{ST}: interpopulation differentiation taking similarities between haplotypes into account; Phylo structure: phylogeographic structure; $F_{ST(b)}$: population differentiation for nuclear; $F_{ST(m)}$: population differentiation for maternally inherited; r: the pollen-to-seed migration ratios in oaks

*N_{ST} differs from G_{ST} at $P < 0.05$

[a] Data from Jiang et al. (2018)

[b] Data from Jiang et al. (2016)

phylogeographic barriers and constrain the dispersal of seeds. Similar phylogeographic structures were also reported for other organisms with limited dispersal abilities, such as spiny frogs (Zhang et al. 2010) and fishes (Wu et al. 2013; Chen et al. 2017). Wind pollinating lineages can overcome these phylogeographic barriers via pollen flow, which was supported by the detected difference in divergence pattern between cpDNA and nSSR.

Seed germination traits and environmental heterogeneity shape the population genetic structure of evergreen oaks on the YGP

Different patterns of population genetic structures were found in three sympatric oak species on the YGP (Table 4). This suggests that the population genetic structure of trees on the YPG was shaped by both biological and environmental factors.

Assuming that pollen is abundantly produced by an organism rather than being a factor that restricts fertilization, the differences in r value observed in the three oak species likely resulted from differences in seed flow. Interestingly, *Q. schottkyana* had a tenfold higher seed-flow-ratio compared with *Q. delavayi*. The two species share largely overlapping climatic niches. The impact of environmental factors alone could not explain the differences in population genetic structure between both species. Their seed dispersal abilities are likely the key reason for the differences in population genetic structure. However, no correlation was found between acorn size and seed dispersal ability: The largest acorn (*Q. kerrii*) had mid-level dispersal, the middle-sized acorn (*Q. schottkyana*) had the highest dispersal ability, and the smallest acorn (*Q. delavayi*) had the lowest seed dispersal ability. In contrast, these results support the food perishability hypothesis. With their delayed germination, the acorns of *Q. schottkyana* had a higher dispersal ability than those of *Q. delavayi* and *Q. kerrii*. Acorns of *Q. delavayi* and *Q. kerrii* germinate soon after falling on the ground, and can even germinate on branches, while the acorns of *Q. schottkyana* remain dormant until the following spring (Xia et al. 2015, 2016, and personal communication). Therefore, *Q. schottkyana* acorns have more time for being transported to other places compared with the other two fast germinating species. Similarly, acorns from the red oak group (section *Lobatae*) have delayed germination and are thus more likely to be hoarded than the early germinating acorns of the white oak group (section *Quercus*) (Hadj-Chikh et al. 1996; Smallwood et al. 2001; Steele et al. 2001a, 2001b, 2006). In comparison, the fast germinating speeds of *Q. delavayi* and *Q. kerrii* may entice frugivores to consume the acorns or damage the existing embryo, which, in turn, increases the seed mortality and reduces seed-induced gene flow. In summary, the
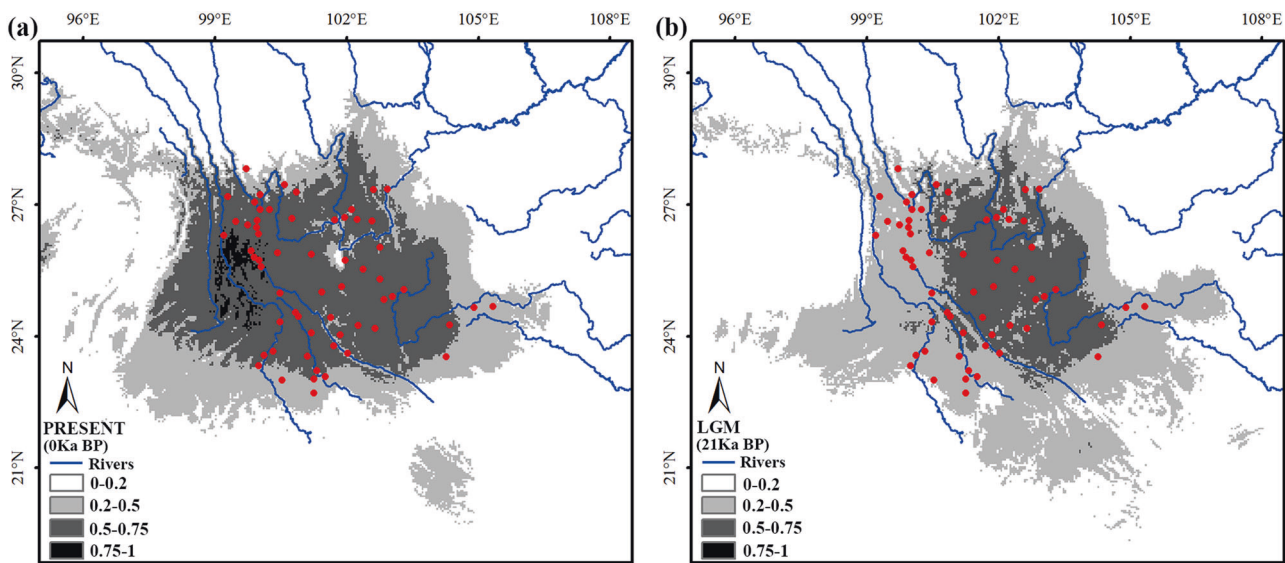
Fig. 5 Simulated distribution ranges of *Quercus delavayi* in the (**a**) present periods and (**b**) last glacial maximum (LGM). The different gray tones in rectangles in the lower right refer to different ranges of occurrence probability of occurrence (habitat suitability) from MAX-ENT output. Red dots represent the points used to simulate the species distributions

presented results suggest that the germination schedules of acorn influence seed dispersal by potentially altering frugivore behavior, thus affecting the population genetic structure of evergreen oaks on the YGP. However, this hypothesis requires the identification of the ecological functions of the different acorn germination patterns by increasing sampling from different populations and by setting up food trap experiments in the future.

Moreover, differences in the nSSR population genetic structure were found between *Q. delavayi*, *Q. kerrii*, and *Q. schottkyana*. *Q. delavayi* and *Q. schottkyana* showed a similar genetic gradient change along latitude, suggesting that environmental changes along latitude similarly impacted the population genetic structure of their nuclear genomes. However, the nSSR population genetic structure of *Q. kerrii* was significantly impacted by the Tanaka line. *Q. kerrii* is a low-elevation species and is adapted to a warmer climate (Fig. S9). Since the YGP is mostly at intermediate elevation, its available habitats are rather rare and highly fragmented on the YGP. The middle-elevation mountain chain along the Tanaka line can efficiently block gene flow among populations of *Q. kerrii*, resulting in genetic differentiation. In contrast, *Q. delavayi* and *Q. schottkyana* are both regionally widespread dominant trees at middle elevation on the YGP with large effective population sizes. The Tanaka line does not form a barrier for their nuclear gene flow. Therefore, the biogeographic significance of the Tanaka line is a phylogeographic barrier to the tropical organisms on the YGP.

Q. delavayi and *Q. kerrii* have relatively closely related phylogenetic relationships and share a similar germination schedule; however, they have different phylogeographic

structures. This phenomenon is likely caused by three factors: first, *Q. kerrii* is a late Miocene derived species, while *Q. delavayi* derived at the early Miocene (Deng et al. 2018). The ancestral lineage of *Q. delavayi* might once have had a much wider distribution during the Miocene than at present, since fossils were found even on areas with high elevation (ca. 4000 m) of the East Tibet plateau (Mangkang), where *Q. delavayi* no longer exists (Xu et al. 2016a). During the long evolutionary time and wide distribution range, a large number of mutations could have accumulated in different regional populations. This may underlie the high genetic diversity and significant population differentiation observed for *Q. delavayi*. Second, the evolution rate of the chloroplast genome of Fagaceae is lower than that of other angiosperm lineages (Frascaria et al. 1993; Graur and Li 2000; Simeone et al. 2013). The high sequence conservation of the cpDNA genome may lead to a difficulty to detect population-level differentiation in *Q. kerrii*. Third, a previous study on phylogeography of *Q. kerrii* mainly sampled populations from its northern distribution range. These populations may have been established via relatively recent demographic expansion. These factors likely cause a lack of phylogeographic structure in the cpDNAs of *Q. kerrii*.

Different demographic dynamics of sympatric evergreen oaks in section *Cyclobalanopsis* on the YGP

SDM analysis, neutrality tests, and mismatch distribution analysis of *Q. delavayi* indicated that this species did not experience recent demographic expansion (Table 4 and Fig. S5). Similarly, *Q. schottkyan*a had a relatively stable

distribution since the LGM (Table 4). Although neutral test for the cpDNA sequences of *Q. kerrii* suggested that this species did not undergo significant demographic expansion, the star-shaped haplotype network and unimodal distribution of the mismatch distribution analysis implied demographic expansion (Table 4).

Quantitative analysis for the SDM results indicated that *Q. kerrii* underwent a larger expansion of its potential distribution range since the LGM compared with *Q. delavayi* and *Q. schottkyana* (Table 4). *Q. kerrii* has a significantly different ecological niche compared with *Q. delavayi* and *Q. schottkyana* (Figs. S10 and S11). During glacial-interglacial periods, the two mid-elevation oak species were able to move to suitable habitats by shifting their distribution elevation into montane regions. In contrast, during the glacial period, suitable habitats for *Q. kerrii* moved to lower elevations, or retreated to southern refugia, then expanded northward during interglaciation. Due to their different ecological tolerances, these three species may have experienced different demographic dynamics during the Quaternary glacial cycles.

This study demonstrated that both the climatic heterogeneity and biological functional traits play important roles for the population genetic structure and demographic dynamics of trees of EBLFs of the YGP. The latitude–longitude climate gradient universally imprinted the nuclear population genetic structure of evergreen oaks on the YGP. The pollen-mediated gene flow of middle-elevation oaks was not strongly impacted by regional topological barriers, except for the lowland tropical species, due to limited habitat availability. Furthermore, the seed germination pattern plays a key role that restricts their seed dispersal distance, and might regulate predation, hoarding, and seed preference by frugivores, which impacts the cpDNA divergence pattern. The interaction between seed germination and both distance- and density-dependent mortality on seedling recruitment patterns needs to be investigated at a finer scale to better understand the ecological functions of different germination patterns in sympatric oaks. This study provides important information on how tree genetic diversity was maintained within the EBLFs on the YGP.

Data availability

All the sequences were deposited as GenBank (Accession numbers MN607123-MN607150); the nuclear microsatellite loci data (nSSR) is available from Dryad Digital Repository: https://doi.org/10.5061/dryad.2v6wwpzhd.

Compliance with ethical standards

References

An M, Deng M, Zheng SS, Song YG (2016) *De novo* transcriptome assembly and development of SSR markers of oaks *Quercus austrocochinchinensis* and *Q. kerrii* (Fagaceae) Tree Genet Genomes 12(6):103

An ZS, Kutzbach JE, Prell WL, Porter SC (2001) Evolution of Asian monsoons and phased uplift of the Himalaya–Tibetan plateau since Late Miocene times Nature 411(6833):62–66

Bandelt HJ, Forster P, Rohl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16(1):37–48

Cavender-Bares J (2019) Diversification, adaptation, and community assembly of the American oaks (*Quercus*), a model clade for integrating ecology and evolution. New Phytol 221:669–692

Chang G, Xia ZS, Zhang ZB (2009) Hoarding decisions by Edward's long-tailed rats (*Leopoldamys edwardsi*) and South China field mice (*Apodemus draco*): the responses to seed size and germination schedule in acorns. Behav Process 82(1):7–11

Chen WT, Zhong ZX, Dai W, Fan Q, He SP (2017) Phylogeographic structure, cryptic speciation and demographic history of the sharpbelly (*Hemiculter leucisculus*), a freshwater habitat generalist from southern China. BMC Evol Biol 17(1):216

Daws MI, Crabtree LM, Dalling JW, Mullins CE, Burslem DFRP (2008) Germination responses to water potential in neotropical pioneers suggest large-seeded species take more risks. Ann Bot 102(6):945–951

Deng M, Jiang XL, Hipp AL, Manos PS, Hahn M (2018) Phylogeny and biogeography of East Asian evergreen oaks (*Quercus* section *Cyclobalanopsis*; Fagaceae): insights into the Cenozoic history of evergreen broad-leaved forests in subtropical Asia. Mol Phylogen Evol 119(2018):170–181

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol 29(8):1969–1973

Ennos RA (1994) Estimating the relative rates of pollen and seed migration among plant populations. Heredity 72(3):250–259

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour 10(2):564–567

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele grequencies. Genetics 164(4):1567–1587

Fan DM, Yue JP, Nie ZL, Li ZM, Comes HP, Sun H (2013) Phylogeography of *Sophora davidii* (Leguminosae) across the "Tanaka-Kaiyong Line", an important phytogeographic boundary in Southwest China. Mol Ecol 22(16):4270–4288

Forget PM, Milleron T, Feer F (1998) Patterns in post-dispersal seed removal by neotropical rodents and seed fate in relation to seed size. In: Newbery DM, Prins HHT, Brown ND (eds) Dynamics of tropical communities. Blackwell Science, Qxford, UK, p 25–49

Frascaria N, Maggia L, Michaud M, Bousquet J (1993) The rbcL gene sequence from chestnut indicates a slow rate of evolution in the Fagaceae. Genome 36(4):668–671

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147(2):915–925

Graur D, Li WH (2000) Fundamentals of molecular evolution, 2nd edn. Sinauer Associates Inc, Sunderland

Hadj-Chikh LZ, Steele MA, Smallwood PD (1996) Caching decisions by grey squirrels: a test of the handling time and perishability hypotheses. Anim Behav 52(5):941–948

Harpending HC (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Hum Biol 66(4):591–600

Harrison TM, Copeland P, Kidd W, Yin A (1992) Raising Tibet. Science 255(5052):1663–1670

Huang CC, Chang YT, Bartholomew B (1999) Fagaceae. In: Wu CY, Raven PH (eds) Flora of China. Science Press and Missouri Botanical Garden Press, Beijing and St. Louis, p 314–400

Huang SSF, Hwang SY, Lin TP (2002) Spatial pattern of chloroplast DNA variation of *Cyclobalanopsis glauca* in Taiwan and East Asia. Mol Ecol 11(11):2349–2358

Hutchison DW, Templeton AR (1999) Correlation of pairwise genetic and geographic distance measures: inferring the relative influences of gene flow and drift on the distribution of genetic variability. Evolution 53(6):1898–1914

Jansen PA, Bongers F, Hemerik L (2004) Seed mass and mast seeding enhance dispersal by a neotropical scatter-hoarding rodent. Ecol Monogr 74(4):569–589

Jiang XL, An M, Zheng SS, Deng M, Su ZH (2018) Geographical isolation and environmental heterogeneity contribute to the spatial genetic patterns of *Quercus kerrii* (Fagaceae) Heredity 120 (3):219–233

Jiang XL, Deng M, Li Y (2016) Evolutionary history of subtropical evergreen broad-leaved forest in Yunnan Plateau and adjacent areas: an insight from *Quercus schottkyana* (Fagaceae). Tree Genet Genomes 12(6):104

Jordano P, Godoy JA (2002) Frugivore-generated seed shadows: a landscape view of demographic and genetic effects. In: Levey DJ, Silva WR, Galetti M (eds) Seed dispersal and frugivory: ecology, evolution and conservation. CAB International, Wallingford, Oxfordshire, UK, p 305–321

Ju MM, Fu Y, Zhao GF, He CZ, Li ZH, Tian B (2018) Effects of the Tanaka Line on the genetic structure of *Bombax ceiba* (Malvaceae) in dry-hot valley areas of southwest China. Ecol Evol 8 (7):3599–3608

Kumar S, Stohlgren TJ (2009) Maxent modeling for predicting suitable habitat for threatened and endangered tree *Canacomyrica monticola* in New Caledonia. J Ecol Nat Environ 1(1):94–98

Li XW, Li J (1997) The Tanaka-Kaiyong line—an important floristic line for the study of the flora of East Asia. Ann Mo Botanical Gard 84(4):888–892

Librado P, Rozas J (2009) DnaSPv5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25 (11):1451–1452

Liu JL, Tang Y, Tran M-D, Cao SY, Zhao L, Zhang ZC et al. (2012) The nature of the Ailao Shan-Red River (ASRR) shear zone: constraints from structural, microstructural and fabric analyses of metamorphic rocks from the Diancang Shan, Ailao Shan and Day Nui Con Voi massifs. J Asian Earth Sci 47(2012):231–251

Liu JQ, Duan YW, Hao G, Ge XJ, Sun H (2014) Evolutionary history and underlying adaptation of alpine plants on the Qinghai-Tibet Plateau. J Syst Evol 52(3):241–249

Lu LM, Mao LF, Yang T, Ye JF, Liu B, Li HL et al. (2018) Evolutionary history of the angiosperm flora of China. Nature 554:234–238

Millar MA, Byrne M, Coates DJ, Roberts JD (2016) Contrasting diversity and demographic signals in sympatric narrow-range endemic shrubs of the south-west Western Australian semi-arid zone. Bot J Linn Soc 118(2):315–329

Monmonier MS (1973) Maximum-difference barriers: an alternative numerical regionalization method. Geogr Anal 5(3):245–261

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. Nature 403(6772):853–858

Nathan R, Muller-Landau HC (2000) Spatial patterns of seed dispersal, their determinants and consequences for recruitment. Trends Ecol Evol 15(7):278–285

Nei M (1975) Molecular population genetic and evolution. North-Holland, Amsterdam

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76(10):5269–5273

Ortego J, Riordan EC, Gugger PF, Sork VL (2012) Influence of environmental heterogeneity on genetic diversity and structure in an endemic southern Californian oak. Mol Ecol 21 (13):3210–3223

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28(19):2537–2539

Petit RJ, Duminil J, Fineschi S, Hampe A, Salvini D, Vendramin GG (2005) Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. Mol Ecol 14 (3):689–701

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecol Model 190 (3–4):231–259

Phillips SJ, Dudik M, Schapire RE (2004) A maximum entropy approach to species distribution modeling. In: Proceeding of twenty-first international conference on Machine Learning. ACM Press, New York, USA, p 655–662

Pons O, Petit RJ (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. Genetics 144(3):1237–1245

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155 (2):945–959

Qiu YX, Fu CX, Comes HP (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 Phylogen Evol 59(1):225–244

Ramos-Palacios CR, Badano EI, Flores J, Flores-Cano JA, Flores-Flores JL (2014) Distribution patterns of acorns after primary dispersion in a fragmented oak forest and their consequences on predators and dispersers. Eur J For Res 133(3):391–404

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol 9 (3):552–569

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61(3):539–542

Rousset F (2008) Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Mol Ecol Resour 8 (1):103–106

Scofield DG, Alfaro VR, Sork VL, Grivet D, Martinez E, Papp J et al. (2011) Foraging patterns of acorn woodpeckers (*Melanerpes formicivorus*) on valley oak (*Quercus lobata* Née) in two California oak savanna-woodlands. Oecologia 166(1):187–196

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu WS, Miller J et al. (2005) The tortoise and the hare II: relative utility of 21 non-coding chloroplast DNA sequences for phylogenetic analysis. Am J Bot 92(1):142–166

Simeone MC, Piredda R, Papini A, Vessella F, Schirone B (2013) Application of plastid and nuclear markers to DNA barcoding of Euro-Mediterranean oaks (*Quercus*, Fagaceae): problems, prospects and phylogenetic implications. Bot J Linn Soc 172 (4):478–499

Smallwood PD, Steele MA, Faeth SH (2001) The ultimate basis of the caching preferences of rodents, and the oak-dispersal syndrome: tannins, insects, and seed germination. Am Zool 41 (4):840–851

Song SY, Krajewska K, Wang YF (2000) The first occurrence of the *Quercus* section *Cerris* Spach fruits in the Miocene of China. Acta Palaeobot 40(2):153–163

Steele MA, Manierre S, Genna T, Contreras TA, Smallwood PD, Pereira ME (2006) The innate basis of food-hoarding decisions in grey squirrels: evidence for behavioural adaptations to the oaks. Anim Behav 71(1):155–160

Steele MA, Smallwood PD, Spunar A, Nelsen E (2001a) The proximate basis of the oak dispersal syndrome: detection of seed dormancy by rodents. Am Zool 41(4):852–864

Steele MA, Turner G, Smallwood PD, Wolff JO, Radillo J (2001b) Cache management by small mammals: experimental evidence for the significance of acorn-embryo excision. J Mammal 82 (1):35–42

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17(5):1105–1109

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123(3):585–595

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30(12):2725–2729

Tanaka T (1954) Species problem in citrus. Japanese society for the promotion of science, ToKyo, Ueno

Tang CQ, Matsui T, Ohashi H, Dong YF, Momohara A, Herrando-Moraira S et al. (2018) Identifying long-term stable refugia for relict plant species in East Asia. Nat Commun 9:4488

Tian B, Zhou ZL, Du FK, He CZ, Xin PY, Ma HC (2015) The Tanaka Line shaped the phylogeographic pattern of the cotton tree (*Bombax ceiba*) in southwest China. Biochem Syst Ecol 60 (2015):150–157

Vander Wall SB (1995) The effects of seed value on the caching behavior of yellow pine chipmunks. Oikos 74(3):533–537

Vander Wall SB (2001) The evolutionary ecology of nut dispersal. Bot Rev 67(1):74–117

Wan S, Li A, Clift PD, Stuut J-BW (2007) Development of the East Asian monsoon: mineralogical and sedimentologic records in the northern South China Sea since 20 Ma. Paleogeogr Paleoclimatol Paleoecol 254(3–4):561–582

Wang BS, Mao JF, Zhao W, Wang XR (2013) Impact of geography and climate on the genetic differentiation of the subtropical pine *Pinus yunnanensis*. PLoS ONE 8(6):e67345

Ward FK (1921) The Mekong-Salween Divide as a geographical barrier. Geogr J 58(1):49–56

Warren DL, Glor RE, Turelli M (2010) ENMTools: a toolbox for comparative studies of environmental niche models. Ecography 33(3):607–611

Whitlock BA, Hale AM, Groff PA (2010) Intraspecific inversions pose a challenge for the *trn*H-*psb*A plant DNA barcode. PLoS ONE 5 (7):e11533

Wu XY, Luo J, Huang S, Chen ZM, Xiao H, Zhang YP (2013) Molecular phylogeography and evolutionary history of *Poropuntius huangchuchieni* (Cyprinidae) in southwest China. PLoS ONE 8(11):e79975

Xia K, Daws MI, Zhou ZK, Pritchard HW (2015) Habitat-linked temperature requirements for fruit germination in *Quercus* species: a comparative study of *Quercus* subgenus *Cyclobalanopsis* (Asian evergreen oaks) and *Quercus* subgenus *Quercus*. S Afr J Bot 100(2015):108–113

Xia K, Harrower WL, Turkington R, Tan HY, Zhou ZK (2016) Pre-dispersal strategies by *Quercus schottkyana* to mitigate the effects of weevil infestation of acorns. Sci Rep 6:37520

Xiao ZS, Gao X, Zhang ZB (2013) Sensitivity to seed germination schedule by scatter-hoarding Pére David's rock squirrels during mast and non-mast years. Ethology 119(6):472–479

Xiao ZS, Wang YS, Harris M, Zhang ZB (2006) Spatial and temporal variation of seed predation and removal of sympatric large-seeded species in relation to innate seed traits in a subtropical forest, Southwest China. Ecol Manag 222(1–3):46–54

Xiao ZS, Zhang ZB, Wang YS (2005) Effects of seed size on dispersal distance in five rodent-dispersed fagaceous species. Acta Oecol 28(3):221–229

Xu H, Su T, Zhang ST, Deng M, Zhou ZK (2016a) The first fossil record of ring-cupped oak (*Quercus* L. subgenus *Cyclobalanopsis* (Oersted) Schneider) in Tibet and its paleoenvironmental implications. Paleogeogr Paleoclimatol Paleoecol 442(2016):61–71

Xu J, Jiang XL, Deng M, Westwood M, Song YG, Zheng SS (2016b) Conservation genetics of rare trees restricted to subtropical montane cloud forests in southern China: a case study from *Quercus arbutifolia* (Fagaceae). Tree Genet Genomes 12(5):90

Yuan QJ, Zhang ZY, Peng H, Ge S (2008) Chloroplast phylogeography of *Dipentodon* (Dipentodontaceae) in southwest China and northern Vietnam. Mol Ecol 17(4):1054–1065

Yue LL, Chen G, Sun WB, Sun H (2012) Phylogeography of *Buddleja crispa* (Buddlejaceae) and its correlation with drainage system evolution in southwestern China. Am J Bot 99(10):1726–1735

Zhang DR, Chen MY, Murphy RW, Che J, Pang JF, Hu JS et al. (2010) Genealogy and palaeodrainage basins in Yunnan Province: phylogeography of the Yunnan spiny frog, *Nanorana yunnanensis* (Dicroglossidae). Mol Ecol 19(16):3406–3420

Zhang L, Li QJ, Li HT, Chen J, Li DZ (2006) Genetic diversity and geographic differentiation in *Tacca chantrieri* (Taccaceae): an autonomous selfing plant with showy floral display. Ann Bot 98 (2):449–457

Zhang TC, Sun H (2011) Phylogeographic structure of *Terminalia franchetii* (Combretaceae) in southwest China and its implications for drainage geological history. J Plant Res 124(1):63–73

Zhao YJ, Gong X (2015) Genetic divergence and phylogeographic history of two closely related species (*Leucomeris decora* and *Nouelia insignis*) across the "Tanaka Line" in Southwest China. BMC Evol Biol 15(1):134