Historical Dynamics of Semi-Humid Evergreen Forests in the Southeast Himalaya Biodiversity Hotspot: A Case Study of the *Quercus franchetii* Complex (Fagaceae)

Si-Si Zheng[1,2], Xiao-Long Jiang[3], Qing-Jun Huang[2] and Min Deng[4,5]*

[1] Shanghai Chenshan Botanical Garden, Shanghai, China, [2] School of Ecological Technique and Engineering, Shanghai Institute of Technology, Shanghai, China, [3] The Laboratory of Forestry Genetics, Central South University of Forestry and Technology, Changsha, China, [4] School of Ecology and Environmental Science, Yunnan University, Kunming, China, [5] Yunnan Key Laboratory of Plant Reproductive Adaptation and Evolutionary Ecology, Yunnan University, Kunming, China

The Oligocene and Miocene are key periods in the formation of the modern topography and flora of East Asian and Indo-China. However, it is unclear how geological and climatic factors contributed to the high endemism and species richness of this region. The *Quercus franchetii* complex is widespread in the southeast Himalaya fringe and northern Indo-China with a long evolutionary history. It provides a unique proxy for studying the diversity pattern of evergreen woody lineages in this region since the Oligocene. In this study, we combined chloroplast (*cp*DNA) sequences, nuclear microsatellite loci (nSSRs), and species distribution modeling (SDM) to investigate the impacts of geological events on genetic diversity of the *Q. franchetii* complex. The results showed that the initial *cp*DNA haplotype divergence was estimated to occur during the middle Oligocene (30.7 Ma), which might have been raised by the tectonic activity at this episode to the Miocene. The nSSR results revealed two major groups of populations, the central Yunnan-Guizhou plateau (YGP) group and the peripheral distribution group when *K* = 2, in responding to the rapid YGP uplift during the late Miocene, which restricted gene flow between the populations in core and marginal areas. SDM analysis indicated that the distribution ranges of the *Q. franchetii* complex expanded northwards after the last glacial maximum, but the core distribution range in YGP was stable. Our results showed that the divergence of *Q. franchetii* complex is rooted in the mid-Oligocene. The early geological events during the Oligocene, and the late Miocene may play key roles to restrict seed-mediated gene flow among regions, but the pollen-mediated gene flow was less impacted. The uplifts of the YGP and the climate since LGM subsequently boosted the divergence of the populations in core and marginal areas.

Keywords: population genetic structure, ecological niche modeling, *Quercus* section *Ilex*, geoclimatic events, phylogeography

INTRODUCTION

The late Paleogene (36~23.3 Ma) is a key period in the formation of the modern topography and flora of Asia (Akhmetiev and Zaporozhets, 2014; Li et al., 2019). During this period, an abrupt climate cooling at the Eocene-Oligocene (E-O) boundary (33.9 Ma) led to the turnover in regional biota and their distribution ranges. Meanwhile, the collisions between the Indian and Eurasian plates greatly changed the topography of Asia (Huchon et al., 1994; Chatterjee et al., 2013; Li S. H. et al., 2017). All these climatic and geological events had profound impacts on the distribution and divergence of the regional biota. However, little is known about how the timing and mechanisms of these ancient geological and climatic events that contributed to the high species diversity and high level of endemism in the southeast Himalayan fringe region.

One prevailing view addressed by numerous scholars is that the rapid uplifts of the Tibetan Plateau-Himalayas (TP) since the Miocene has created new habitats and niches, which have in turn promoted sympatric speciation (Liu et al., 2002; Wang et al., 2007; Meng et al., 2008). Meanwhile the uplift increased the complexity of the regional topography, which efficiently blocked gene flow among the populations, promoting allopatric speciation (Liu et al., 2006; Xu et al., 2010; Li et al., 2013; Qu et al., 2014). However, this view was challenged by Renner (2016) in a review that multiple line of evidences support TP had been 4–5 km high since the mid-Eocene, however, many phylogenetic works in Asia simply attributed the fast speciation between 0.5 and 15 Ma to the fast uplifts of TP, which is either miscited or outdated.
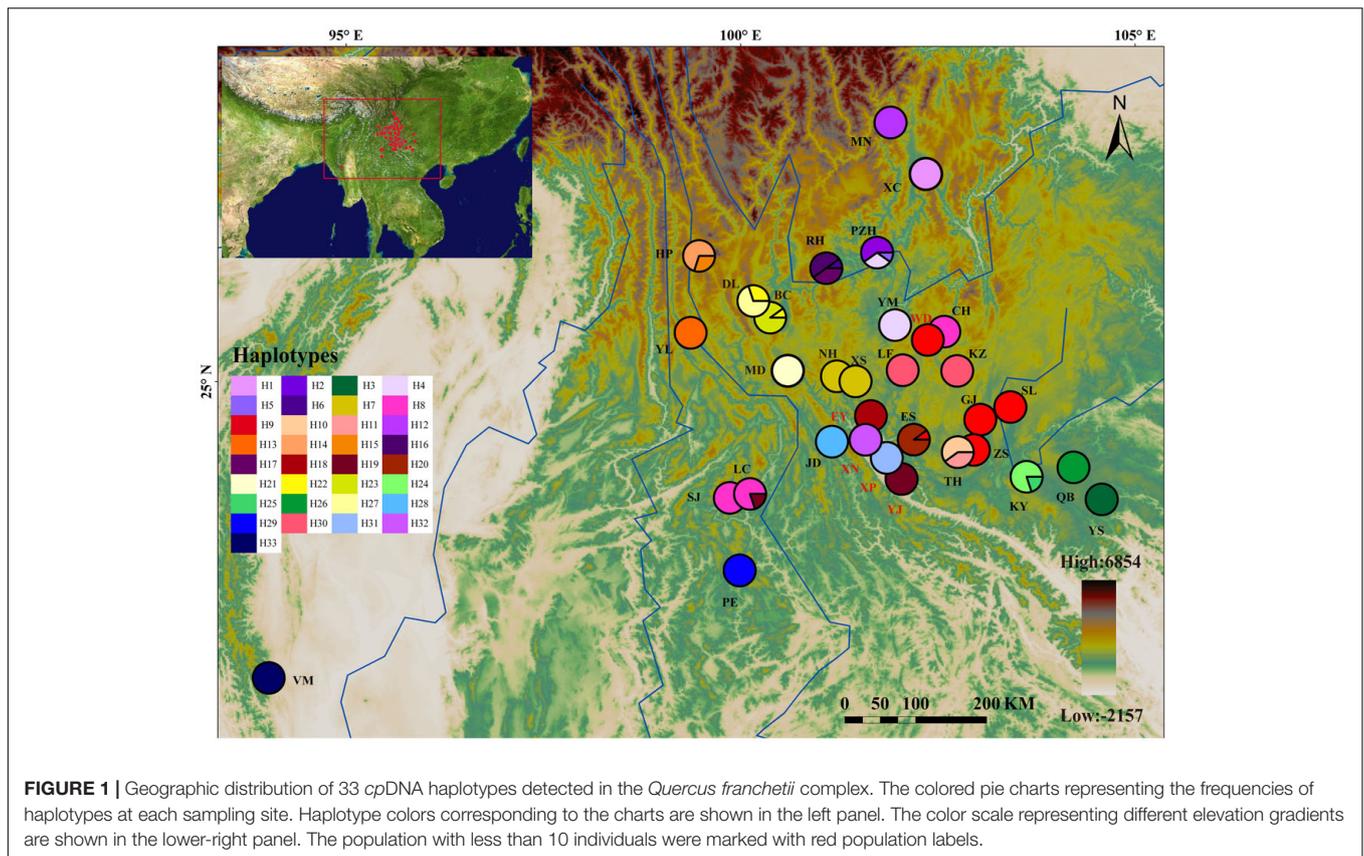
A recent study of Chinese angiosperms by Lu et al. (2018) indicated that the present Chinese flora is young, as a large number of genera did not originate until the Miocene (23 Ma), and their study determined that East China is a "museum" unlike West China, which is a "cradle" of herbaceous species. Likewise, the genome-wide analysis of *Salix brachista* (cushion willow) showed that the TP uplift induced sky island habitats, which increased population differentiation in combination with the Quaternary climate fluctuations, thus boosting *in situ* speciation (Chen et al., 2019). Other phylogenetic studies on relic woody species have also revealed recent divergence events, most dating back to the Miocene, with an intensification in the Pliocene-Pleistocene, regardless of whether these relic lineages had Paleogenic or even more ancient origins, e.g., *Cephalotaxus* (Cephalotaxaceae) (Wang et al., 2014), *Taxus* (Taxaceae) (Gao L. M. et al., 2007; Liu et al., 2013), and *Picea* (Pinaceae) (Shao et al., 2019).

Many study cases have collectively clarified the spatio-temporal diversity pattern of plant lineages in the East Himalayas since the Miocene. Indeed, the uplifts induced environmental heterogeneity and geographic barriers that together triggered the rapid diversification of these subtropical lineages. However, the ancient impacts of the geological events in the Oligocene to early Miocene remain a mystery, as the phylogeographic studies on the Paleogene diverged lineages are quite rare and fossil records in Asia at this epoch are scarce. Nevertheless, there are quite a few Paleogene relic genera also distributed in East Asian and the southwestern Himalayan fringe, e.g., *Ginkgo* (Shen et al., 2005;

Gong et al., 2008), *Taxus* (Gao L. M. et al., 2007; Ji et al., 2020), and *Eurycorymbus* (Wang et al., 2009). Phylogeographic studies on those relic species have illustrated that southwestern China, Dabashan, and the Wuyi Mountain regions served as important refugia during time periods with extreme climates (Shen et al., 2005; Gao L. M. et al., 2007; Wang et al., 2009). However, most of these lineages suffered massive extinctions at the geological timescale, with only few extant species with rather restricted distributions remaining. Therefore, investigations of their spatial genetic pattern can only provide limited information about the evolutionary dynamics occurring further back in geological time.

Quercus franchetii and *Q. lanata* belong to *Quercus* section *Ilex*. The two species are widespread in southwestern China and the southern Himalayas to Northern Indo-China, respectively, at an elevation range of approximately 800–2,600 m (Govaerts and Frodin, 1998; Huang et al., 1999). Their distribution ranges cover the main area of the ancient Red River drainage basin. They are both key regional trees in semi-humid evergreen broad-leaved forests and dry-hot river valleys and have important ecological service and functions (Wangda and Ohsawa, 2006; Liu et al., 2011, 2012). The early derived status of *Q. franchetii* in section *Ilex* at the E-O boundary was inferred by recent phylogenetic studies on oaks (Jiang et al., 2019; Hipp et al., 2020). Our recent phylogenetic study on *Quercus* section *Ilex* supplemented *Q. lanata* for analysis. The result showed that *Q. lanata* and *Q. franchetii* are sister taxa (unpublished data). However, after adding RAD-seq data from more individuals from both species and reconstructing the phylogenetic tree, neither of the two species form a monophyletic clade (unpublished data). Likewise, morphometric measurements showed no difference between *Q. franchetii* and *Q. lanata*, both at the species level as well as among different geographical regions (Zheng, 2021). These results together suggested that the two species are very closely related to each other or indeed represent the same species (hereafter called "*Quercus franchetii* complex"). Data beyond the molecular dating results concur that the *Q. franchetii* complex has a long evolutionary history. Fossils resembling the extant *Q. franchetii* complex have been widely reported along the Tethys/Paratethys Seaway dating to the late Eocene to Pliocene, and they were commonly used as a proxy indicating warm and semi-humid climates (Bouchal et al., 2017; Denk et al., 2017; Guner et al., 2017). Thus, the *Q. franchetii* complex offers a unique opportunity for untangling the timing and possible mechanisms by which geological events since the Oligocene have shaped the high biodiversity and endemism level of the southeastern Himalaya fringe.

In this study, we comprehensively sampled populations of the *Q. franchetii* complex throughout its distribution range (**Figure 1**). We used *cp*DNA and nSSR markers to scan the populations. By coupling population genetic structure analyses with species distribution modeling (SDM), we aimed to (1) illustrate the spatial genetic structure of the *Q. franchetii* complex, (2) identify the key environmental factors restricting the distribution and diversity pattern of the species complex, and (3) explore the key factors that drove the divergence of the *Q. franchetii* complex. This study provides deep insights into the distribution and evolutionary dynamics of this subtropical

FIGURE 1 | Geographic distribution of 33 *cp*DNA haplotypes detected in the *Quercus franchetii* complex. The colored pie charts representing the frequencies of haplotypes at each sampling site. Haplotype colors corresponding to the charts are shown in the left panel. The color scale representing different elevation gradients are shown in the lower-right panel. The population with less than 10 individuals were marked with red population labels.

woody lineage in the context of global environment change, informing efforts to safeguard this unique forestry ecosystem in the Southeast Himalaya biodiversity hotspot.

MATERIALS AND METHODS

Ethics Statement

Sampling of oaks were granted and supported by National Forestry Bureau of China, Local National Nature Reserves, and Ministry of Environmental Conservation and Forestry, Myanmar.

Population Sampling

Thirty-three *Q. franchetii* complex populations were sampled from Yunnan and Sichuan, China and Mount Victoria in Chin State, Myanmar. In total, 303 individuals from 33 populations were sampled for this study, covering the major known distribution range of the *Q. franchetii* complex (**Table 1**). Samples from the same population came from individuals that were separated by at least 50 m. At least 10 trees were sampled from each population, except for populations with very few individuals (WD, EY, YJ, XP, XN), in which case we sampled all the accessible adult trees in those populations. Fresh and healthy mature leaves were collected and put into containers with silica gel to dry them quickly until DNA extractions could be performed. The voucher

specimens of the DNA samples were deposited in the herbarium of Shanghai Chenshan Botanical Garden (CSH).

DNA Extraction, PCR Amplification, and Sequencing

Total genomic DNA was extracted using a modified cetyltrimethyl ammonium bromide (CTAB) protocol (Doyle and Doyle, 1987). Three pairs of *cp*DNA primers, namely *psb*A-*trn*H (Shaw et al., 2005), *trn*T-*trn*L (Taberlet et al., 1991), and *atp*I-*atp*H (Grivet et al., 2001), and eight highly polymorphic nuclear microsatellite loci, specifically nuclear simple sequence repeats (nSSRs), were selected for genotyping on all samples. Primer sequences and PCR amplification conditions are summarized in **Supplementary Appendix 1**. The *cp*DNA and nSSR amplification conditions followed the methods described by Xu et al. (2015) and An et al. (2016), respectively. PCR products of nSSR markers were 10 times diluted, then mixed with fluorescence size standards at a ratio of 6-FAM: HEX: ROX = 1:1:2, then genotyped by Shanghai Majorbio Bio-pharm Technology Co., Ltd. (Shanghai, China). The PCR products of *cp*DNAs were purified, then biodirectionally sequenced by the same company. All sequences obtained in this study have been deposited in GenBank (see "Data Availability").

The *cp*DNA sequences were assembled and checked using Sequencher 4.01 (Gene Codes Corp., Ann Arbor, MI, United States). The ClustalW implementation in MEGA X (Kumar et al., 2018) was used for sequence alignment with

TABLE 1 | Sampling information, nSSR and *cp*DNA genetic diversity, probabilities of populations belonging to each genetic cluster (C_A, C_B, C_C, C_D) inferred by InStruct analyses, locality habitat suitability, stability obtained from SDMs, and bottleneck effect test for the *Quercus franchetii* complex.

| Species | Pop code | Lon | Lat | n | cpDNA | | | SSRs | | | | | | SDM | | | Bottleneck |
					Haplotypes (no. of individuals)	h	π × 10³	C_A	C_B	C_C	C_D	A_r	H_e	N_{Pre}	N_{LGM}	N_{stab}	P_W_2t
QF	XC	102.35	27.66	10	H1 (10)	0	0	0.008	0.900	0.083	0.009	3.79	0.532	0.215	0.224	0.991	**0.020***
QF	PZH	101.73	26.65	10	H2 (6), H4 (3), H5 (1)	0.6	0.042	0.042	0.928	0.018	0.012	3.83	0.529	0.703	0.457	0.754	**0.012***
QF	LF	102.05	25.15	10	H30 (10)	0	0	0.038	0.028	0.932	0.006	3.85	0.557	0.778	0.686	0.909	0.313
QF	YM	101.96	25.72	10	H4 (10)	0	0	0.025	0.937	0.023	0.015	4.17	0.612	0.786	0.655	0.869	0.547
QF	RH	101.08	26.45	10	H6 (1), H16 (5), H17 (4)	0.64	0.028	0.011	0.964	0.009	0.016	3.55	0.519	0.757	0.471	0.714	0.250
QF	NH	101.22	25.07	10	H7 (10)	0	0	0.031	0.048	0.914	0.008	3.96	0.569	0.754	0.627	0.873	1.000
QF	ZXS	101.46	25.01	10	H7 (10)	0	0	0.87	0.033	0.08	0.017	3.87	0.564	0.731	0.598	0.867	0.688
QF	CH	102.58	25.64	10	H8 (10)	0	0	0.141	0.066	0.787	0.006	3.85	0.573	0.688	0.579	0.891	1.000
QF	SJ	99.86	23.51	10	H8 (10)	0	0	0.957	0.017	0.02	0.006	3.23	0.486	0.253	0.391	0.862	0.641
QF	LC	100.12	23.57	10	H8 (8), H19 (2)	0.36	0.011	0.687	0.022	0.286	0.005	3.18	0.518	0.391	0.323	0.932	0.938
QF	GJ	103.04	24.52	10	H9 (10)	0	0	0.978	0.006	0.01	0.006	3.47	0.523	0.760	0.637	0.878	0.383
QF	ZS	102.95	24.13	10	H9 (10)	0	0	0.954	0.008	0.032	0.006	3.38	0.486	0.767	0.658	0.891	0.313
QF	SL	103.42	24.68	10	H9 (10)	0	0	0.073	0.876	0.042	0.009	3.66	0.563	0.627	0.572	0.945	0.313
QF	WD	102.37	25.53	1	H9 (1)	1	0	0.78	0.006	0.209	0.005	–	–	0.684	0.550	0.867	–
QF	ES	102.19	24.26	10	H9 (1), H20 (9)	0.2	0.013	0.932	0.028	0.018	0.023	4.02	0.586	0.707	0.723	0.983	0.055
QF	TH	102.75	24.10	10	H10 (6), H11 (4)	0.53	0.028	0.904	0.022	0.061	0.014	3.46	0.523	0.736	0.635	0.899	0.313
QF	MN	101.90	28.31	10	H12 (10)	0	0	0.009	0.946	0.02	0.025	3.85	0.557	0.509	0.234	0.725	0.945
QF	YL	99.37	25.63	10	H13 (10)	0	0	0.028	0.928	0.04	0.009	3.5	0.536	0.520	0.305	0.785	0.461
QF	HP	99.47	26.61	10	H14 (7), H15 (3)	0.47	0.005	0.007	0.971	0.014	0.009	3.79	0.544	0.386	0.164	0.778	0.250
QF	EY	101.65	24.56	6	H18 (6)	0	0	0.781	0.184	0.028	0.006	3.5	0.568	0.770	0.687	0.917	0.945
QF	YJ	102.04	23.76	7	H19 (7)	0	0	0.841	0.029	0.119	0.011	3.23	0.437	0.638	0.631	0.993	**0.020***
QF	MD	100.60	25.14	10	H21 (10)	0	0	0.064	0.024	0.898	0.015	3.39	0.506	0.764	0.643	0.879	0.375
QF	BC	100.37	25.82	10	H22 (1), H23 (9)	0.2	0.026	0.018	0.019	0.958	0.005	3.69	0.530	0.703	0.486	0.783	0.195
QF	DL	100.16	26.03	10	H22 (3), H27 (7)	0.47	0.005	0.808	0.031	0.122	0.039	4.34	0.604	0.653	0.388	0.735	0.074
QF	KY	103.62	23.78	10	H24 (8), H25 (2)	0.36	0.004	0.029	0.772	0.174	0.025	4.01	0.588	0.766	0.667	0.902	0.461
QF	KZ	102.74	25.14	10	H30 (10)	0	0	0.108	0.102	0.765	0.025	3.67	0.540	0.681	0.559	0.879	0.313
QF	XN	101.58	24.26	3	H32 (3)	0	0	0.047	0.021	0.927	0.005	–	–	0.350	0.759	0.591	–
QL	YS	104.58	23.49	10	H3 (10)	0	0	0.925	0.046	0.011	0.018	3.73	0.579	0.465	0.574	0.891	1.000
QL	JD	101.16	24.23	12	H28 (12)	0	0	0.020	0.018	0.942	0.02	3.46	0.549	0.651	0.726	0.924	0.641
QL	QB	104.22	23.91	10	H26 (10)	0	0	0.022	0.94	0.031	0.007	3.95	0.603	0.600	0.565	0.965	0.742
QL	XP	101.85	24.03	3	H31 (3)	0	0	0.868	0.057	0.064	0.011	–	–	0.699	0.713	0.986	–
QL	PE	99.99	22.59	10	H29 (10)	0	0	0.007	0.026	0.014	0.953	3.28	0.522	0.151	0.132	0.981	0.297
QL	VM	94.02	21.21	11	H33 (11)	0	0	0.009	0.009	0.01	0.972	3.7	0.586	0.118	0.074	0.955	0.844

QF, Q. franchetii; QL, Q. lanata; Lon, longitude; Lat, latitude; n, number of individuals investigated in the population; h, haplotype diversity; π, nucleotide diversity; H_e, expected heterozygosity; A_r, standardized allelic richness; N_{Pre}, present habitat suitability; N_{LGM}, last glacial maximum (LGM) habitat suitability; N_{stab}, habitat stability since the LGM; P_W_2t, P-values of the TPM model based on the Wilcoxon sign-rank test; (the bold values)*, significant correlation (P < 0.05).

manual adjustment. All microsatellite loci were checked using GeneMarker® (Hulce et al., 2011) to detect and analyze allele sizes. Null alleles and stutter bands were checked with MicroChecker (Van Oosterhout et al., 2004).

Genetic Diversity and Structure

Haplotypes and polymorphism statistics for *cp*DNA loci were calculated with DnaSP 6.0 (Rozas et al., 2017). The haplotype geographic distribution was projected onto a map using ArcGIS 10.5.[1] Total haplotype diversity (H_T), within-population diversity (H_S), and coefficients of differentiation (G_{ST} and N_{ST}) for

[1]http://www.esri.com/software/arcgis/

*cp*DNA loci were estimated using PERMUT 2.0 (Pons and Petit, 1996). Haplotype diversity (h) and nucleotide diversity (π) for *cp*DNA loci were obtained with ARLEQUIN 3.5 (Excoffier and Lischer, 2010). The *cp*DNA haplotype network was constructed using the Median-Joining model in NETWORK 5.0.0.1 (Bandelt et al., 1999). GenAlEx 6.5 (Peakall and Smouse, 2012) was used to calculate the genetic diversity index of microsatellite data, including expected heterozygosity (H_e), observed heterozygosity (H_o), number of alleles (N_A), and effective number of alleles (N_E). Allelic richness (A_r) of nSSRs was determined using HP-RARE (Kalinowski, 2005).

The genetic differentiation values (F_{ST}) based on *cp*DNA and nSSR data were estimated using ARLEQUIN 3.5, respectively

(Excoffier and Lischer, 2010). The Genetic Landscape GIS toolbox (Vandergast et al., 2011) in ArcGIS 10.5 was used to generate a geographical landscape map based on both genetic diversity (A_r) for *cp*DNA loci and genetic divergence (F_{ST}) based on *cp*DNA and nSSR data according to inverse distance weighted interpolation. For *cp*DNA and nSSR data, analysis of molecular variance (AMOVA) was performed to assess the genetic variation among populations and within populations. The Mantel test was also performed based on the genetic structure and geographic distance matrix with 1,000 random permutations to evaluate their relationship and test isolation by distance (IBD). For nSSR data, Principal coordinate analysis (PcoA) based on genetic distance was performed using GenAlex 6.5 (Peakall and Smouse, 2012) to assess differences among individuals or groups. Barrier (Manni et al., 2004) was used to set up geographic barriers according to sample locations to detect the existence of genetic barriers among populations.

We used GenePop v 4.2 (Rousset, 2008) to test departures from Hardy-Weinberg equilibrium (HWE) for each of the eight nSSR loci. As the nSSR loci significantly deviated from HWE (see Results section "Genetic Diversity and Structure"), Bayesian assignment probability methods using the programs InStruct (Gao H. et al., 2007) and STRUCTURE 2.3.4 (Pritchard et al., 2000) were both used to infer the population genetic structure. The number of clusters (K) was varied from 1 to 10, with 10 replicates at each value. Each run consisted of a burn-in length of 25,000 iterations with a run length of 500,000 MCMC (Markov chain Monte Carlo) iterations. The optimal K (number of clusters) was determined using the $\triangle K$ method (Evanno et al., 2005). CLUMPP (Jakobsson and Rosenberg, 2007) was used to align 10 runs of InStruct with the optimum K using a greedy algorithm.

Among the three models in the Bottleneck program (Luikart et al., 1998), the "Two-phase mutation model" (TPM) is the most suitable for microsatellite loci, which was thus selected to test whether the populations of the *Q. franchetii* complex had experienced a bottleneck. We used "Wilcoxon sign-rank test" method for a significance test.

Divergence Time Estimation

The divergence time dating for *cp*DNA haplotypes of the *Q. franchetii* complex was estimated using a Bayesian approach as implemented in BEAST V2.4 (Suchard et al., 2018). *Quercus glauca* (*Quercus* section *Cyclobalanopsis*) was chosen as the outgroup to root the tree. An uncorrelated lognormal relaxed clock was applied with the K81uf + I substitution model, which was selected based on the Akaike information criterion (AIC) in Modeltest 3.7 (Posada and Crandall, 1998). The earliest conclusive leaf fossils of *Quercus* section *Ilex* were discovered in Tibet, dated to the late Eocene (ca. 34 Ma) (Su et al., 2019) was set as the minimum age to constrain the stem of the haplotype tree of the *Q. franchetii* complex with a lognormal distribution and a median of 37.8 Ma (95% HPD: 34.02–56.6 Ma). The MCMC chains were run for 100 million generations with a sampling frequency of once every 10,000 generations. Convergence was

assessed using Tracer v1.7[2] (Rambaut et al., 2018), and the effective sample sizes for all parameters were calculated. The resulting tree and log files from the two replicate runs were combined with LogCombiner v1.8. Then, we used TreeAnnotator v. 1.8[3] to generate the maximum clade credibility (MCC) tree after discarding the first 20% of the trees as burn-in. The results were visualized using FigTree v1.4.3.[4]

Population Demographic History and Ancestral Area Reconstruction Analyses

Pairwise mismatch distribution analysis for *cp*DNA loci, Tajima's (1989) *D* and Fu's (1997) *Fs* of neutrality tests were performed to detect possible demographic expansions of the *Quercus franchetii* complex using ARLEQUIN 3.5 (Excoffier and Lischer, 2010).

Ancestral range reconstruction was analyzed using statistical dispersal-vicariance (S-DIVA) analysis as implemented in RASP (Yu et al., 2015) and DEC models (Ree et al., 2005; Ree and Smith, 2008). Four areas were delimited for ancestral area reconstruction based on the geological characteristics and biogeographical division of southwestern China, including (A) the Nanpan River region (NPR); (B) the southwestern Red River (including southern Himalayas to Red River) (RR); (C) the Hengduan Mountains (HDM), and (D) the Yunnan-Guizhou Plateau (YGP) (**Figure 2A**).

Species Distribution Modeling

We used MaxEnt 3.4 (Phillips et al., 2006) to simulate the potential distribution range of the *Q. franchetii* complex (including *Q. franchetii* and *Q. lanata*) under the past, current, and projected future climate scenarios based on the maximum entropy model (Phillips et al., 2004). In total, 145 accurate occurrence points were collected from the Global Biodiversity Information Facility (GBIF),[5] Chinese Digital Herbarium (CVH),[6] and the field collection records of our research team. Each voucher specimen in the distribution record was inspected and checked carefully. Nineteen bioclimatic variables with a 2.5-arc-min resolution for the present (1950–2000), the last glacial maximum (LGM) (CCSM) period, and the future (2060–2080, RCP2.6; RCP4.5; RCP6.0; RCP8.5) were downloaded from WorldClim 2.[7] The nine environmental factors (bio1, Annual Mean Temperature; bio4, Temperature Seasonality; bio6, Min Temperature of Coldest Month; bio7, Temperature Annual Range; bio10, Mean Temperature of Warmest Quarter; bio11, Mean Temperature of Coldest Quarter; bio12, Annual Precipitation; bio13, Precipitation of Wettest Month; bio17, Precipitation of Driest Quarter) were selected after eliminating climatic variables such that none that were include were highly correlated (i.e., with correlation coefficients
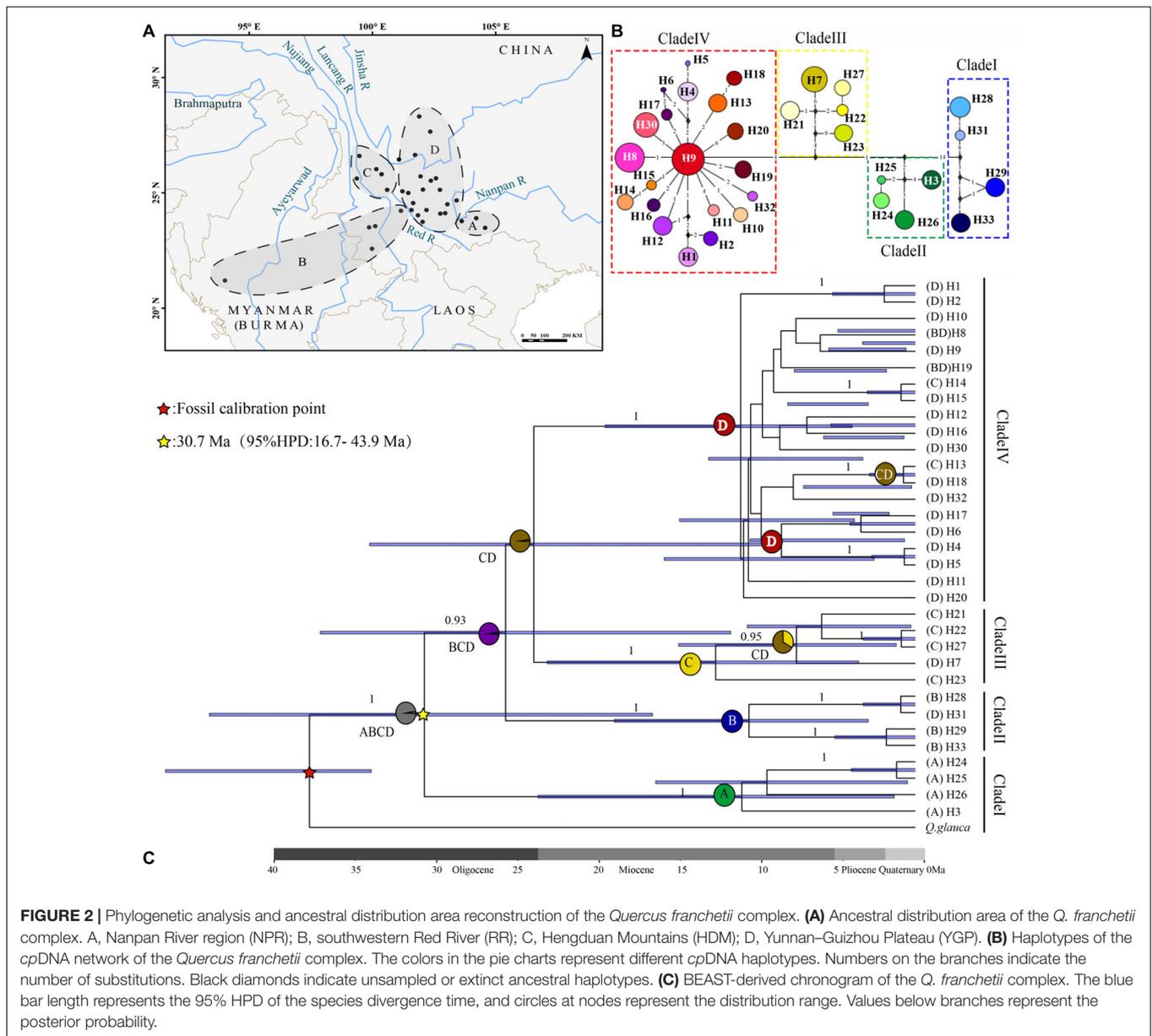
---

[2]http://tree.bio.ed.ac.uk/software/tracer/
[3]http://beast.bio.ed.ac.uk/TreeAnnotator
[4]http://tree.bio.ed.ac.uk/software/figtree/
[5]http://www.gbif.org
[6]https://www.cvh.ac.cn
[7]http://www.worldclim.org/

FIGURE 2 | Phylogenetic analysis and ancestral distribution area reconstruction of the *Quercus franchetii* complex. **(A)** Ancestral distribution area of the *Q. franchetii* complex. A, Nanpan River region (NPR); B, southwestern Red River (RR); C, Hengduan Mountains (HDM); D, Yunnan–Guizhou Plateau (YGP). **(B)** Haplotypes of the *cp*DNA network of the *Quercus franchetii* complex. The colors in the pie charts represent different *cp*DNA haplotypes. Numbers on the branches indicate the number of substitutions. Black diamonds indicate unsampled or extinct ancestral haplotypes. **(C)** BEAST-derived chronogram of the *Q. franchetii* complex. The blue bar length represents the 95% HPD of the species divergence time, and circles at nodes represent the distribution range. Values below branches represent the posterior probability.

greater than 0.8) using the Dismo package in R.[8] We also chose CHELSA climate data,[9] which had a higher prediction power in mountain regions (specifically the Himalayas), to simulate the present distribution (Bobrowski and Schickhoff, 2017; Karger et al., 2017).

In order to improve the prediction accuracy of the model, it was necessary to optimize the model in MaxEnt 3.4 (Phillips et al., 2006) by setting the β multiplier and environmental characteristic parameters. The MaxEnt model captures five features: linear (L), quadratic (Q), hinge (H), product (P), and threshold (T). In this research, we used seven feature combinations (auto, L, H, LQ, LPQ, LQH, LQHP) and set the regularization multiplier from

0.5 to 10 with increments of 0.5. Then, we used ENMTools (Warren et al., 2010) to calculate the lambdas file for the maxent result, and selected the model with the smallest AIC value as the optimal model parameter for further analyses. The occurrence points of *Q. franchetii* and *Q. lanata* were randomly divided into 75 and 25% of the data for training and testing, respectively. The area under the curve (AUC) was used to evaluate the model's accuracy. AUC values ranged from 0.5 to 1, where the higher AUC value indicated a better prediction. The maximum training sensitivity plus specificity thresholds for the presence or absence of species was used to draw a species distribution map in ArcGIS 10.5. In addition, we classified three categories equally between this threshold and 1, corresponding to low, medium, and high fitness areas, respectively. To compare the changes in species distribution across different periods, three indicators were

---

[8]https://CRAN.R-project.org/package=dismo
[9]https://chelsa-climate.org/

calculated: locality habitat stability (N_{Stab}), habitat distribution area ratio (N_a), and habitat expansion extent (N_e). These values were calculated using the following formulas: $N_{Stab} = 1 - | N_{Pre} - N_{LGM} |$, where N_{Pre} and N_{LGM} are the habitat suitability of the present and LGM distribution area; N_a = (present distribution areas)/(LGM or future distribution areas): a value close to 1 indicates a stable distribution of the species, while a value much higher or lower than 1 indicates that the distribution area of the species has expanded or contracted from the LGM to the present; $N_e = [1 - $ (Distribution area overlapping between the LGM and present or the future and present/present distributions area)$] \times 100\%$ represents the percentage of the distribution that has expanded from the LGM to the present.

The potential dispersal routes of the *Q. franchetii* complex in the past and present periods were inferred based on the least-cost path analysis method using SDM toolbox 2.0 (Brown, 2014) in ArcGIS 10.5. The specific steps are listed below: Firstly, we generated a resistance layer by inverting the SDMs (1-SDM). The resistance layer was used to create a cost distance raster for each sample locality. The corridor layers were built between two locations that only share haplotypes based on the cost distance raster. We used the categorical the least cost path (LCP) approach to better describe the habitat heterogeneity and its role in the dispersal. The value of each corridor layer was divided into low, medium, and high, and then these three intervals were re-divided into new values 5, 2, and 1. Finally, we reclassified all the corridor layers, summarized and standardized them from 0 to 1, and determined the dispersal corridors of the *Q. franchetii* complex.

Detection of Correlations Between Genetic Diversity and Climatic Factors

The linear model in R 3.5[10] was used to estimate the correlation of genetic diversity indexes (A_r and H_e) and genetic structure (cluster A from the Bayesian clustering, C_A) with habitat and geographic factors. Five variables—population longitude and latitude, habitat suitability for the present (N_{Pre}) and the LGM (N_{LGM}), and habitat stability (N_{Stab}) were used as explanatory covariates.

RESULTS

Genetic Diversity and Structure

The *psbA-trnH*, *atpI-atpH*, and *trnT-trnL* sequence alignments were 482–520, 899–1030, and 796–838 bp in length, respectively. The combined length of the three aligned chloroplast fragments was 2,319 bp, with 94 polymorphisms across a total of 33 haplotypes among the 303 individuals analyzed. The *cp*DNA haplotype diversity was $h = 0.956$ and nucleotide diversity was $\pi = 0.00536$. The nucleotide diversity and haplotype diversity within the populations were 0–0.04184 and 0–0.644, respectively (**Table 1** and **Figure 1**). The total diversity ($H_T = 0.982$) was much higher than the average within-population diversity ($H_S = 0.123$). The genetic diversity map showed that the NPR and HDM populations had high genetic diversity (**Figure 3C**). N_{ST} (0.959)

was significantly greater than G_{ST} (0.874) ($P < 0.05$), which indicated that a clear phylogeographic structure existed among the *Q. franchetii* complex populations. The network analysis resolved four distinct *cp*DNA haplotype clades in the *Q. franchetii* complex, which was consistent with the haplotype structure determined by BEAST. The haplotypes in Clade IV exhibited a star-like structure. The most common haplotype, H9, had the widest distribution, mainly within the central YGP (**Figure 2B**).

Among the 303 individuals genotyped, a total of 115 alleles were identified using eight pairs of microsatellite primers. All eight microsatellite loci were highly polymorphic, with allele numbers varying from 8 to 18 per locus. The genetic diversity indexes N_A, N_E, H_e, and H_o were 4.182 (SE = 0.103), 2.578 (SE = 0.066), 0.537 (SE = 0.013), and 0.536 (SE = 0.016), respectively, while allele richness (A_r) was 3.18–4.34.

The AMOVA on *cp*DNA sequence data revealed greater genetic variation among populations (87%; $F_{ST} = 0.87$) than within populations (13%). In contrast, the nSSRs showed substantial genetic differences within populations (63%; $F_{ST} = 0.369$, **Table 2**). Pairwise F_{ST} values calculated based on chloroplast and SSR data were shown in **Supplementary Appendix 2**. In order to more visually observe genetic differentiation, F_{ST} values were projected onto a map and a genetic divergence map was made. The genetic divergence map showed that the population in southwestern Yunnan had a relatively high genetic divergence, while the divergence of the populations from the northeast was comparatively lower (**Figures 3A,B**). The Mantel tests of *cp*DNA data ($r = 0.052$, $P = 0.001$) and microsatellite data ($r = 0.413$, $P = 0.001$) both revealed significant correlations between genetic and geographic distances, but the correlations among chloroplast markers were weaker (**Supplementary Appendix 3**). Barriers inferred from nuclear genes showed a barrier between the YGP and HDM regions (**Figure 4**). As well as, distinct geographical isolation had been detected between the populations in southwestern Yunnan and NPR (**Figure 4**). Based on the TPM model, three populations were determined to have experienced bottleneck (PZH, XC, YJ; **Table 1**).

A total of three loci of nSSRs conformed to HWE, and five loci deviated from HWE ($P < 0.05$, **Supplementary Appendix 1**). In the Bayesian clustering analysis, the optimal K-value for STRUCTURE was $K = 4$ (**Supplementary Appendix 4**). However, the optimal K-value selected by InStruct was 2, and the second highest peak occurred for $K = 4$ (**Supplementary Appendix 4**). Thus, we performed cluster analysis on the *Q. franchetii* complex using $K = 2$ and $K = 4$, respectively for both STRUCTURE and InStruct. Following the InStruct runs with $K = 2$, cluster a was mainly composed of the YGP group, with some populations in the HDM (DL, BC, MD), and southwestern Yunnan (LC, SJ) groups. Cluster b was composed of 12 populations peripheral to those of cluster a. When $K = 4$, the YGP group was further divided into subgroups, and the populations from RR (PE, VM) separated from the rest of cluster b to form a new subgroup (**Figure 4**). The STRUCTURE results were identical to the results obtained from STRUCTURE (**Supplementary Appendix 5**). The principal coordinate analysis (PcoA) of the clustering results found that 33 natural populations
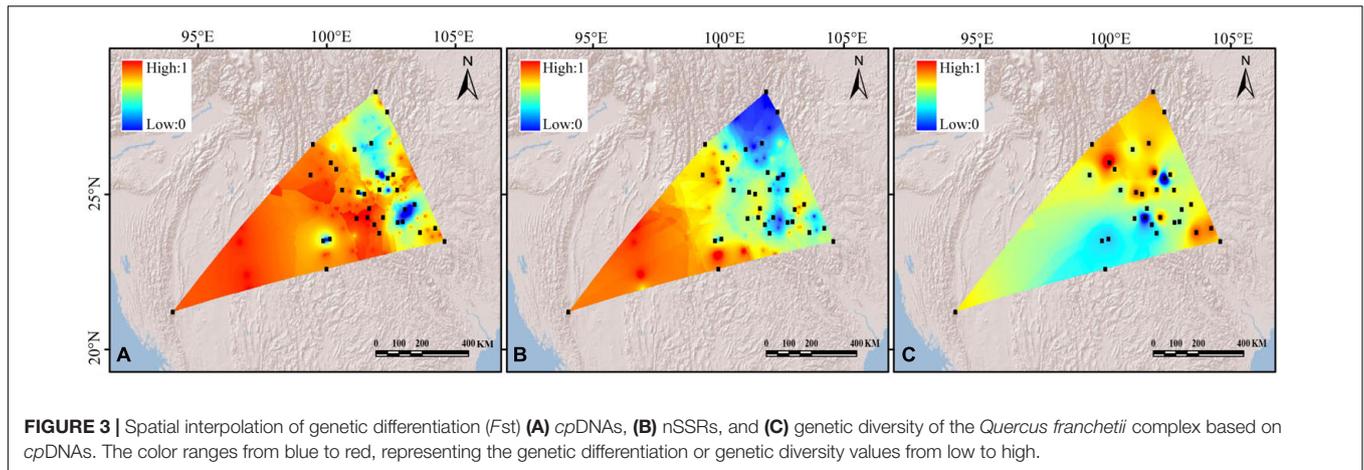
FIGURE 3 | Spatial interpolation of genetic differentiation (*Fst*) **(A)** *cp*DNAs, **(B)** nSSRs, and **(C)** genetic diversity of the *Quercus franchetii* complex based on *cp*DNAs. The color ranges from blue to red, representing the genetic differentiation or genetic diversity values from low to high.

TABLE 2 | Molecular variance analysis of the *Quercus franchetii* complex.

	Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation%
***cp*DNA**	Among populations	31	126.684	0.427	87%
	Within populations	270	17.200	0.064	13%
	Total	301	143.884	0.491	100%
	Among populations	31	1014.597	2.939	37%
SSR	Within populations	270	1356.910	5.026	63%
	Total	301	2371.507	7.965	100%

can be divided into four regions, which is mostly consistent with the results obtained using InStruct. The first and second principal components explained 16.24 and 6.91% of the genetic variation, respectively (**Supplementary Appendix 6**).

Divergence Time Estimation

The crown age of *cp*DNA haplotypes in the *Q. franchetii* complex was dated to the late Oligocene (30.7 Ma, 95% highest posterior density, HPD = 16.7–43.9 Ma), and Clade I (NPR) was the earliest derived. The haplotypes of the southwestern lineage (Clade II) began to diversify around 25.7 Ma (95% HPD = 11.9–37.2 Ma). The divergence of Clades III (HDM region) and IV (YGP region) was dated to ca. 24 Ma (95% HPD = 9.95–34.13 Ma), with subsequent rapid divergence of the haplotypes during the late Miocene (**Figure 2**).

Demographic History and Ancestral Area Reconstruction

Demographic History

Neutrality tests (Tajima's D = −0.50558, P = 0.325; Fu's Fs = 2.77576, P = 0.789) failed to identify population expansion. The mismatch distribution for *cp*DNA showed a multimodal distribution, consistent with a stable population size (**Supplementary Appendix 7**).
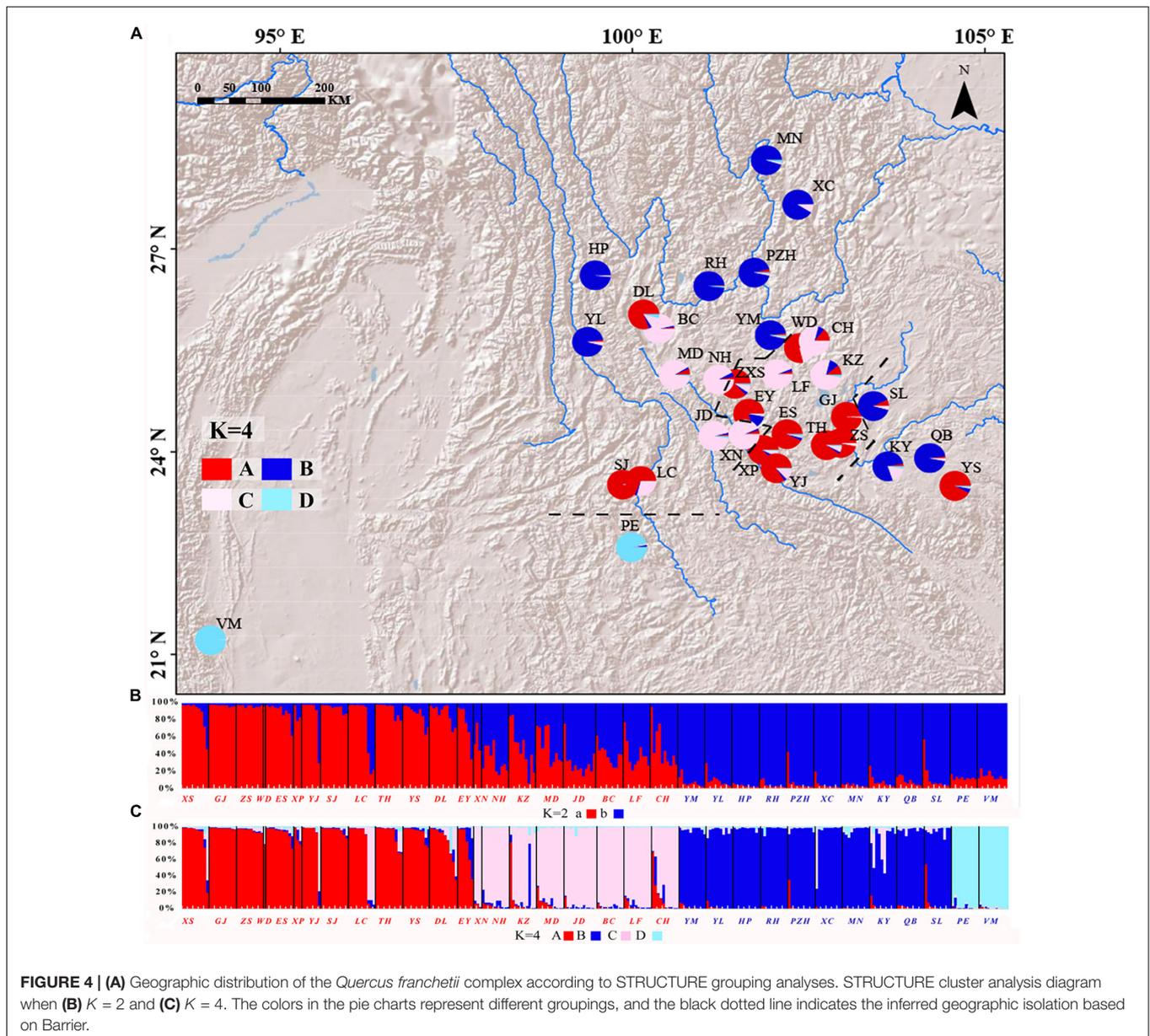
Ancestral Area Reconstruction

The ancestral distribution range reconstruction of the S-DIVA and DEC models both showed that the *Q. franchetii* complex had a wide distribution of ancestral populations. The S-DIVA model (95% HPD = 99%) provided a higher confidence estimate

than did the DEC model (95% HPD = 45.25%), and its results indicated that the ancestral distribution area of the *Q. franchetii* complex was widespread in southwestern Yunnan and the Southern Himalayas (including A, B, C, and D). Then followed three vicariance events, which led to the divergences of the NPR lineage (Clade I) during the Oligocene (95% HPD = 99%) and the RR lineage (Clade II) during the late Oligocene episode (95% HPD = 98.97%), with an increase in the divergence of the HDM lineage (Clade III) and YGP lineage (Clade IV) during the early Middle Miocene (**Figure 2**). The DEC results are shown in **Supplementary Appendix 8**.

Ecological Niche Modeling

Ecological niche modeling of the *Q. franchetii* complex in Maxent using WorldClim climate data revealed a high performance score (AUC = 0.9679–0.9719, standard deviation = 0.0157). Annual Mean Temperature (bio1) was the greatest contributor (36.46%, standard deviation = 1.3), followed by Temperature Seasonality (bio4) (34.7%, standard deviation = 1.3) and Annual Precipitation (bio12) (7.72%, standard deviation = 1.52) in identifying the areas of occurrence for *Q. franchetii* complex populations. The maximum sensitivity plus specificity value 0.13 was used as the species absence/presence threshold. The current distribution of the *Q. franchetii* complex was similar to the predicted distribution, except for quite a few occurrence sites in southwestern and southeastern Yunnan and Southern Himalayas that were located in predicted unsuitable areas (**Figure 5B**). The potential distribution range retreated to the south during the LGM (**Figure 5A**). The total distribution area of the LGM was greater than that of the present (N_a = 0.42,
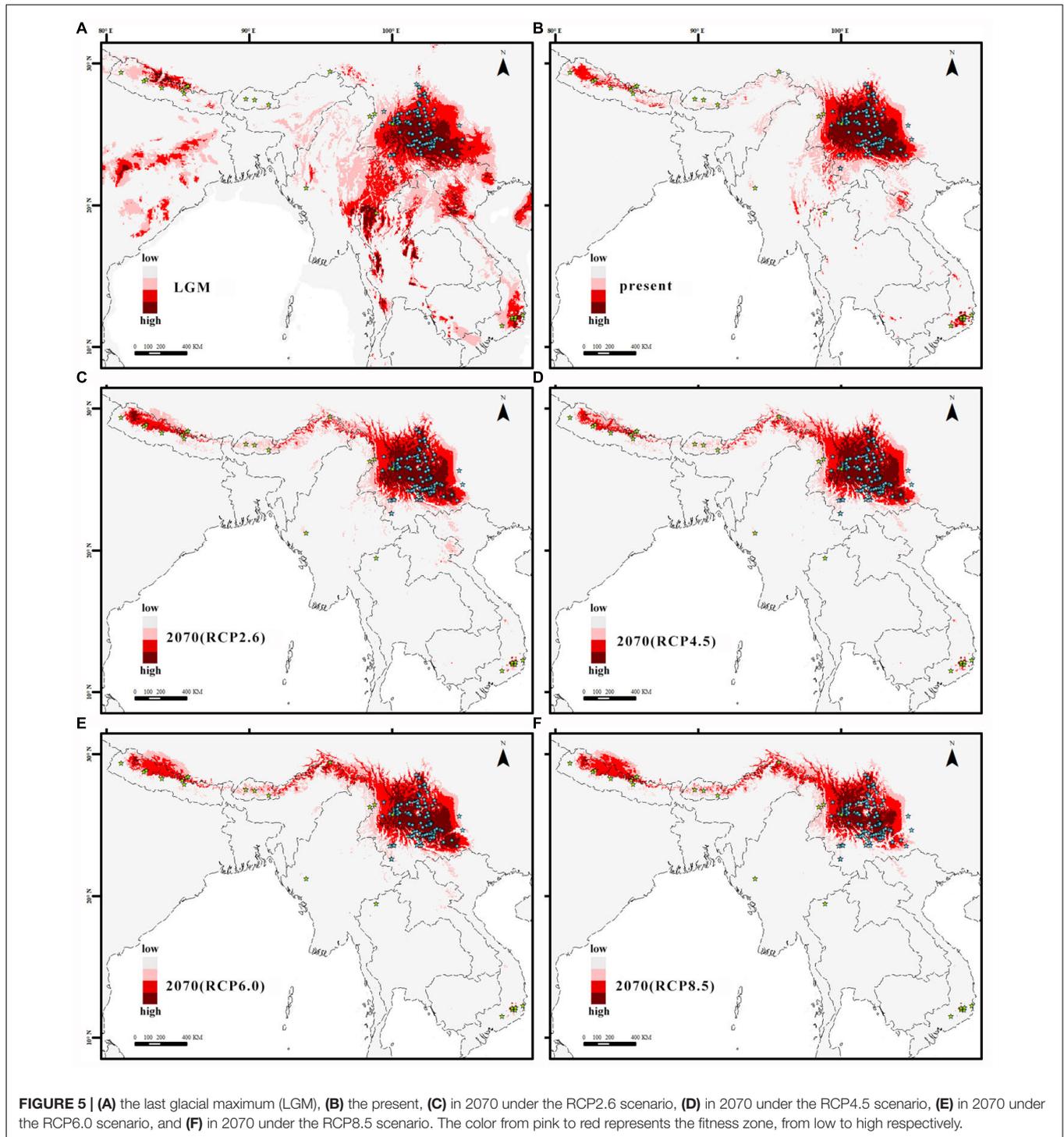
FIGURE 4 | (A) Geographic distribution of the *Quercus franchetii* complex according to STRUCTURE grouping analyses. STRUCTURE cluster analysis diagram when **(B)** $K = 2$ and **(C)** $K = 4$. The colors in the pie charts represent different groupings, and the black dotted line indicates the inferred geographic isolation based on Barrier.

$N_e = 21\%$). In the future period, the higher the RCP value was, the more obvious of an expansion in the suitable area to the northwest was detected. The ratios of the present range to the future range under RCP2.6, RCP4.5, RCP6.0, and RCP8.5 were 1.07, 1.15, 1.04, and 1.18, respectively. Accordingly, the highly suitable habitat would be reduced by 7.2, 7.2, 16.5, and 40.1%, respectively (**Figures 5C–F**). The predicted present distribution of *Q. franchetii* complex using CHELSA climate data was similar to that from WorldClim climate data (**Supplementary Appendix 9**). The suitable distribution area in the current period from the two data sources was largely overlapped (0.94). In order to better compare the distribution dynamics with other oak, of which WorldClim database were generally used for distribution simulation, we selected the results of WorldClim for the subsequent analyses.

Putative dispersal corridors in the two periods were visualized based on *cp*DNA haplotype diversity (**Figure 6**). The dispersal corridors during the two periods are consistent, and both showed that the YGP area had a higher dispersal ratio, but the populations of the peripheral areas were rather isolated.

Correlation Between Genetic Diversity and Climatic Factors

The correlation analysis indicated that only the N_{LGM} was significantly correlated with GenPCoA1 ($P < 0.05$). A_r was associated with both longitude and latitude ($P < 0.05$). Cluster A (C_A) and GenPCoA1 were both associated with latitude ($P < 0.05$). However, H_e was not associated with N_{Pre}, $N_{stabLGM}$, N_{LGM}, latitude, nor longitude ($P > 0.05$) (**Table 3**).

FIGURE 5 | (A) the last glacial maximum (LGM), **(B)** the present, **(C)** in 2070 under the RCP2.6 scenario, **(D)** in 2070 under the RCP4.5 scenario, **(E)** in 2070 under the RCP6.0 scenario, and **(F)** in 2070 under the RCP8.5 scenario. The color from pink to red represents the fitness zone, from low to high respectively.

DISCUSSION

Genetic Diversity Pattern of the *Quercus franchetii* Complex

Our study revealed a high genetic diversity estimate for the *Q. franchetii* complex (H_T, 0.982; A_r, 3.18–4.34), which is similar to the genetic diversity of other evergreen oaks in southwestern China, e.g., *Q. schottkyana* (H_T, 0.828; A_r, 4.83–7.78; Jiang et al., 2016), *Q. kerrii* (H_T, 0.71; A_r, 2.27–3.20; Jiang et al., 2018), *Q. delavayi* (H_T, 0.907; A_r, 3.750–5.237; Xu et al., 2020), and the *Q. cocciferoides* complex (H_T, 0.904; Liu, 2019). Comparatively, the genetic diversity for deciduous oaks seems lower, e.g., in the eight European white oaks (H_T, 0.635–0.847) (Petit et al., 2002), *Quercus variabilis* (H_T, 0.888, in 50 populations in East

FIGURE 6 | Potential dispersal corridors of the *Quercus franchetii* complex during **(A)** the last glacial maximum (LGM) and **(B)** the present. Both white and black dots are the occurrence points of the *Q. franchetii* complex, with the black dots indicating sampling sites used in this study and white points indicating occurrence sites obtained from herbarium records. Colors in **(A,B)** from blue to red represent the potential of species dispersal corridors from low to high.

TABLE 3 | Correlation between genetic diversity, habitat and geographical factors of *Quercus franchetii* complex.

	Ar				**He**				C_A				**GenPCoA1**			
	Estimate	**Se[a]**	**t**	**p**	**Estimate**	**Se[a]**	**t**	**p**	**Estimate**	**Sea**	**t**	**p**	**Estimate**	**Sea**	**t**	**p**
(Intercept)	−7.44	4.00	−1.86	0.07	−0.70	0.59	−1.18	0.25	4.03	1.50	2.68	0.01*	14.42	6.35	2.27	0.03*
N_{Pre}	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
$N_{stadLGM}$	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
N_{LGM}	–	–	–	–	–	–	–	–	–	–	–	–	1.31	0.51	−2.98	0.01*
longitude	0.08	0.04	2.16	0.04*	0.01	0.01	1.89	0.07	–	–	–	–	−0.11	0.06	−1.75	0.09
latitude	0.13	0.04	3.20	0.004 *	0.01	0.01	1.50	0.15	−0.15	0.06	−2.44	0.02*	−0.17	0.06	−2.98	0.01*

*Significant correlation (P < 0.05), Se[a], standard error.

Asia) (Chen et al., 2012), *Q. mongolica* var. *crispula* (H_T, 0.827, in Japan) (Okaura et al., 2007), *Q. acutissima* (H_T, 0.791, in Southeast China) (Zhang et al., 2015). The deciduous species are mainly Miocene-Pliocene derived young lineages, but the divergence time of the evergreen oaks are much older that dated to the Eocene-Miocene (Hipp et al., 2020). It is easy

to understand that oak taxa with long evolutionary histories and wide distribution can accumulate higher levels of genetic diversity than those "young" species. Additionally, the high environmental heterogeneity of southwestern China can buffer the climate extremes of the Quaternary episode (Zhang et al., 2010; Kai and Jiang, 2014; Jiang et al., 2018). As a result, the habitats of southwestern China have had long-term stability without significant regional extinction events or distribution range shifts to allow the genetic diversity of the species can be maintained. All these factors contributed to the high genetic diversity level found in evergreen oaks in southwestern China.

Comparing the genetic structure of the sympatric close related species can better illustrate the factors determining genetic diversity pattern. Notably, there was significant genetic differentiation among populations in the *Q. franchetii* complex, with F_{ST} estimates for cpDNA and nSSRs of 0.87 and 0.369, respectively, which is very similar to that found in *Q. delavayi*, as the cpDNA of the both species show significant phylogeographic structure, IBD pattern, and high differentiation among the populations (Xu et al., 2020). However, in two other sympatric/parapatric oaks, *Q. kerrii* (Jiang et al., 2018) and *Q. schottkyana* (Jiang et al., 2016), only low differentiation among the populations without phylogenetic structure were found (**Table 4**). The biological traits restrict gene flow and its efficiency, no doubt, can greatly impact the population genetic structure (Petit et al., 2003; Cavender-Bares et al., 2015; Xu et al., 2020). Pollen and seed mediated gene flow among the populations is different, as they have different dispersal efficiencies when barriers exist (e.g., rivers and high mountains). Generally, the seed-mediated gene flow among populations is more restricted in the species with instant germination seeds comparing to those species that seed with a period of dormancy. Consequently, the typical recalcitrant seed species shows significant phylogeographic structure in cpDNA makers. Vice versa, as pollen of oaks can disperse

long distance, population differentiation revealed by biparental markers was much lower than that in maternal makers (Xu et al., 2020). However, the genetic structure of *Q. franchetii* seems not only determine by seed/pollen mediated gene flow efficiency. Although *Q. franchetii* has temporary seed dormancy (2–4 months; Xia et al., 2012), its seed size and tannin content similar to those of *Q. schottkyana*, but its population genetic structure is similar to that of *Q. delavayi*—a typical recalcitrant small seeds species. Thus, factors beyond seed germination schedule, seed size, and dispersal abilities also played important roles in shaping the population structure of these oaks in southwestern China, e.g., their ancestor distribution range and evolutionary history, etc.

Contemporary and historical factors shaped the genetic structure of organisms (Van Oppen et al., 2011; Hernawan et al., 2017; Li J. J. et al., 2017). Geological and climatic factors have been shown to influence the evolutionary histories of taxa and shape their genetic structures (Feng et al., 2014; Xing et al., 2014; Lu et al., 2018; Chen et al., 2020). Thus, the genetic structures of ancient lineages with long evolutionary histories and wide distribution range can essentially record more ancient geological events than in those of young lineages. The evergreen oak lineages in YGP and southwestern China, e.g., *Q. schottkayana* at 6.37 Ma (Jiang et al., 2016), *Q. kerrii* at 6–7 Ma (Jiang et al., 2018), and *Q. cocciferoides* at ca. 5 Ma (Jiang et al., 2019; Liu, 2019) are later derived "young" (originated at the late Neogene) and they generally show no (or very limited) IBD pattern among the populations. In contrasts, the early derived species *Q. delavayi* with crown node age at 10.92 Ma, and the *Q. franchetii* complex with crown node age at 30.7 Ma (95% HPD 16.7–43.9 Ma) had similar genetic structures and distinct IBD pattern. The different genetic structures detected in the "young" and "old" species might reflect the outcomes of past geological events on the biota at the different epochs, as the ancient geological events may imprint a genetic structure

TABLE 4 | Comparisons of the genetic diversity, genetic structure, and demographic dynamics of *Quercus franchetii* complex with *Q. schottkyana*, *Q. kerrii*, *Q. schottkyana,* and *Q. cocciferoides* complex.

Taxa	Genetic diversity				Genetic structure				Demographic change			IBD	
	H_T (se)	H_S (se)	G_{ST} (se)	N_{ST} (se)	Phylo structure	Network structure	F_{ST} (b)	F_{ST} (m)	Na	Mismatch	Neutral test	R (m)	R (b)
Q. franchetii complex	0.982 (0.01)	0.123 (0.04)	0.874 (0.04)	0.959* (0.02)	No	Star-like	0.369	0.87	0.42	No expansion	No expansion	0.052*	0.413*
Q. delavayi	0.907 (0.03)[a]	0.197 (0.05)[a]	0.782 (0.05)[a]	0.912* (0.03)[a]	Yes[a]	Geographic structure[a]	0.063[a]	0.938[a]	1.14[a]	No expansion[a]	No expansion[a]	0.587*[a]	0.365*[a]
Q. kerrii	0.71 (0.06)[b]	0.05 (0.02)[b]	0.93 (0.03)[b]	0.92 (0.04)[b]	No[b]	Star-like[b]	0.066[a]	0.894[a]	1.27[b]	Expansion[a]	No expansion[b]	–	–
Q. schottkyana	0.828 (0.06)[a]	0.341 (0.06)[a]	0.588 (0.07)[a]	0.615 (0.11)[a]	No[a]	Star-like[c]	0.075[a]	0.665[a]	1.03[c]	No expansion[c]	No expansion[c]	–	–
Q. cocciferoides Complex	0.904[d]	0.140[d]	0.845[d]	0.860[d]	Yes[d]	Geographic structure[d]	0.134[d]	0.946[d]	1.29[d]	No expansion[d]	No expansion[d]	0.312[d]	0.886[d]

H_T, total haplotype diversity; H_S, within population diversity; G_{ST}, coefficient of genetic variation over all populations; N_{ST}, coefficient of genetic variation influenced by both haplotype frequencies and genetic distances between haplotypes; Phylo structure, phylogeographic structure; F_{ST}(b), population differentiation for nuclear; F_{ST}(m), population differentiation for maternally inherited; Na, habitat distribution area ratio; R(m), correlation of two matrices for maternally inherited; R(b), correlation of two matrices for nuclear; *N_{ST} differs from G_{ST} at P < 0.05.
[a]Xu et al. (2020); [b]Jiang et al. (2018); [c]Jiang et al. (2016); [d]Articles to be published.
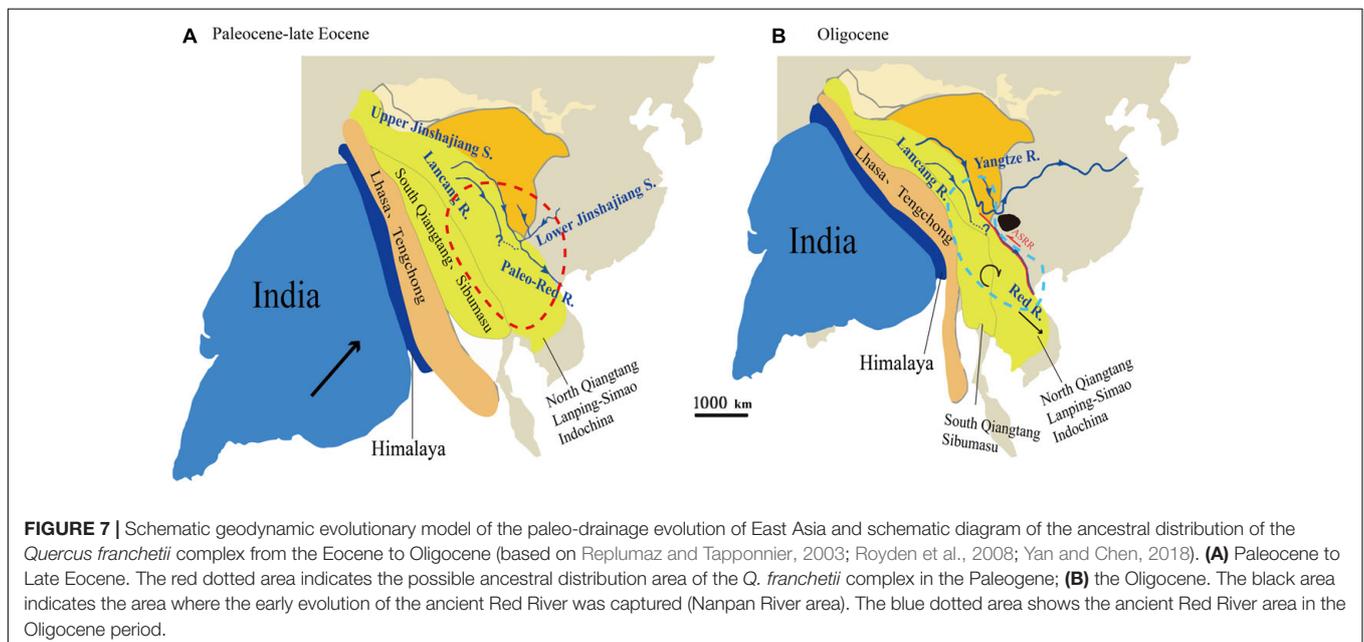
pattern in an "old" and widespread lineages, but not in those of the "young" lineages. However, our hypothesis requires further investigation, as the molecular markers we used in previous phylogeographical studies on oaks of southwestern China are not all the same. Regardless, these markers are informative to reveal the spatial genetic structures of these oaks, but it contains potential bias when comparing across the species to reveal the underlying mechanisms that shaped the genetic structures of these oaks. Further studies using universal high-throughput markers to scan oak populations in southwestern China and incorporating seed functional traits analyses are necessary to reveal the underlying drivers lead to the contemporary genetic structure found in these oaks.

The Impacts of Ancient Geological Events on the Divergence of Subtropical Lineages in the Southeastern Himalaya Fringe

The Oligocene was a key period in the formation of the modern topography of China. During this period, the high eastern and low western topographies of China were totally reversed (Ming, 2007; Di et al., 2018). The tectonic plates associated with the spreading of the East Asian marginal sea and the uplift of the eastern margin of the Tibetan Plateau exerted important influence on the topography and drainage of East Asia (Clark et al., 2004). This was accompanied by dramatic tectonically induced topographic changes and landscape development, which resulted in repeated river captures and reversals (Zheng, 2015). Paleomagnetic and stratigraphic evidence suggests that there was a wide Paleo-Red River drainage basin between the southeastern Tibetan Plateau and the South China Sea, which included Hoh-Xil, Songpan-Ganzi, northern Qiangtang, Yidun, and western Yangtze Terranes (**Figure 7**). Since the late Eocene, the hard

collision between the Indian and Eurasian plates began. These massifs/blocks were extruded eastward under the resulting pressure (Chen et al., 2017). Meanwhile, the Ailao Shan-Red River fault zone began to slip rapidly. The Lanpin-Simao block began to rotate clockwise substantially, which blocked the upper and lower Yangtze River from continually flowing to the south along the paleo-drainage (**Figure 7**). As a result, the northern drainage basin of the paleo Red River disappeared, and the modern Red River began to become established (Replumaz and Tapponnier, 2003; Royden et al., 2008; Chen et al., 2017; He et al., 2021). These the Oligocene to early Miocene events greatly changed the regional biota, e.g., giving rise to the diversification of *Cautleya*, *Roscoea* (Zhao et al., 2016), *Badidae* (Rüber et al., 2004), and spiny frogs (*Paini*) (Che et al., 2010).

Our ancestral range reconstruction based on *cp*DNA data showed that the *Q. franchetii* complex once had a wide distribution in southwestern China and the southern Himalaya regions during the early Oligocene, followed by three vicariance events. Among these events, the NPR (Clade I) was first diversified during the mid-Oligocene (ca. 30 Ma), and then, during the late Oligocene and the early Miocene, the RR lineage and HDM lineage were derived, respectively. Within the main lineages, the fast divergence of the *cp*DNA haplotypes occurred during the late Miocene (**Figure 7A**). The NPR was located at the core area during early river capture; the hard collision between the Indian and Eurasian plates might have squeezed the plates in this region leading to river re-alignment, which induced the divergence of the NPR lineage (Clark et al., 2004; Yang et al., 2012; **Figure 7B**). During the late Oligocene, the Ailao Shan-Red River fault had an early left strike-slip that may have raised the barrier that blocked gene flow among the populations and promoted allopatric divergence (Hall et al., 2008; Fyhn et al., 2009; Lin et al., 2009; Deng et al., 2014). Thus, the RR lineage diverged (Clade II). During the early Miocene, large fault basins

FIGURE 7 | Schematic geodynamic evolutionary model of the paleo-drainage evolution of East Asia and schematic diagram of the ancestral distribution of the *Quercus franchetii* complex from the Eocene to Oligocene (based on Replumaz and Tapponnier, 2003; Royden et al., 2008; Yan and Chen, 2018). **(A)** Paleocene to Late Eocene. The red dotted area indicates the possible ancestral distribution area of the *Q. franchetii* complex in the Paleogene; **(B)** the Oligocene. The black area indicates the area where the early evolution of the ancient Red River was captured (Nanpan River area). The blue dotted area shows the ancient Red River area in the Oligocene period.

were established corresponding to the fast uplift of the YGP and HDM regions (Zhang, 2012), which might dramatically change the regional topography and climate, eventually blocking gene flow in the two regions and leading to the divergence of the YGP lineage (Clade III) and HDM lineage (Clade IV). Followed by the late Neogene period fast HDM uplifts, the complex topography of southwestern China was eventually formed, which further restricted regional seed-mediated gene flow and promoted the divergence of *cp*DNA haplotypes. A similar scenario was also detected in other oaks, e.g., *Q. delavayi* (Xu et al., 2020), and *Q. aquifolioides* (Du et al., 2017), as well as plant lineages with wide distribution in semi-humid evergreen broadleaved forests on YGP region, e.g., *Primula secundiflora* (Wang et al., 2008), *Terminalia franchetii* (Zhang et al., 2011), and *Cycas multipinnata* (Gong et al., 2015). Such phenomenon suggested the regional biota might be impacted by similar environmental drivers.

In summary, the high biodiversity levels found in southwestern China are rooted deeply in the Oligocene. The early tectonic events during the Oligocene drove the main lineage splits, while the fast uplifts of the Himalayas during the Miocene-Pliocene increased environmental heterogeneity and established substantial dispersal barriers (Meng et al., 2017; Xing and Ree, 2017; Yuan et al., 2019). Then, the Quaternary climate fluctuation led to distribution range contractions and expansions of the species, as well as the occurrence of Asian winter monsoons, and the dry season in winter and spring in southwestern China further restricted gene flow between the core YGP region and its periphery (Su et al., 2013; Ye et al., 2019). All these geological and climatic factors interacted during different timespans to shape the contemporary divergence pattern of the biota of the southeastern Himalaya biodiversity hotspot. Our phylogeographic study indicated the Oligocene tectonic induced divergence in *Q. franchetii* complex, which is a supplement to the review of Renner (2016) concerning the Paleogene events contributing to the species richness of the East Himalayan biodiversity hotspot.

Moreover, the population genetic structures inferred from nSSR and *cp*DNA markers of *Q. franchetii* complex were dissimilar. Notably, nSSRs mainly reflected the divergence between the populations in the core YGP region and the peripheral populations. The similar population genetic structure on nuclear genome was also reported in another sympatric species (*Q. cocciferoides*; Liu, 2019). All these evidences suggested the two species might underwent the similar selection pressure to trigger their divergence. The potential migration corridor analysis on the *Q. franchetii* complex showed that the populations in core YGP region maintained strong gene flows, but the marginal populations were mostly isolated since LGM (**Figure 6**). The rugged topography induced by the rapid YPG uplift during the late Neogene and (or) highly fragmented habitat of semi-humid evergreen forests in the peripheral areas around YGP might boost the allopatric divergence of these oaks. Nevertheless, another possibility of this pattern is that the quick evolution and possible backward evolution of SSR markers blurs the geological pattern. Further investigation using high throughput marker to illustrate the genetic structure at fine scale can provide a better understanding on the interplays between genetic diversity and environmental factors.

In contrast, the *cp*DNA data of *Q. franchetii* complex showed a much clearer phylogeographic structure, which has also been shown in other oaks, e.g., *Q. delavayi* (Xu et al., 2020), *Q. cocciferoides* (Liu, 2019), and *Q. aquifoliodes* (Du et al., 2017). Pollen- and seed-mediated gene flow have very different dispersal efficiencies in oaks (Du et al., 2017; Liu, 2019; Xu et al., 2020). In this study, nSSRs showed no significant differentiation of the populations in the East Red River, West Red River, and HDM regions, but *cp*DNA data showed a clear phylogeographic structure. This result suggested that these early tectonic activities during the Oligocene to early Miocene might have restricted seed-mediated gene flow in different regions, but only had limited impacts on pollen-induced gene flow.

In contrast, the niche modeling result suggested that the *Q. franchetii* complex populations are mainly located in the predicted suitable area. While the species complex experienced a southern contraction during the LGM, the distribution area at present and during the LGM largely overlapped in YGP. The quaternary glaciation had only minor impacts on its distribution. Likewise, the niche modeling results of *Osteomeles schwerinae* (Wang et al., 2015), *Q. schottkyana* (Jiang et al., 2016), and *Q. kerrii* (Jiang et al., 2018) from southwestern China show the similar pattern. Collectively, these studies suggest that central YGP region is an important refugia for species in semi-humid evergreen broadleaved forests in southwestern China.

CONCLUSION AND PERSPECTIVE

The spatial genetic structure is subject to environmental factors and the evolutionary process of the organism that affect genetic and genomic variation. The southeastern Himalaya fringe with extensive environmental changes since Cenozic high biodiversity. In our case, the population genetic diversity pattern of the *Q. franchetii* complex showed that the divergence of this subtropical lineage is rooted at the Oligocene. The tectonic events ever since this epoch might have restricted the regional seed-mediated gene flow, in turn triggered the early divergences of this subtropical woody lineage (Replumaz and Tapponnier, 2003; Clark et al., 2004; Zhang et al., 2011). Following, the rapid uplift-induced environmental heterogeneity in the Miocene in the southeastern Himalayas fringe, with subsequent Quaternary climatic fluctuations inducing distribution range expansions and contractions might further restrict the gene flow among the populations in core distribution and the peripheral areas (Huchon et al., 1994; Liu et al., 2006; Xu et al., 2010). These geological and climatic factors acted in a combined manner to boost the diversification of the subtropical biota in the southeastern Himalaya fringe. Our study provides an example that clearly reveals the evolutionary dynamics of the subtropical evergreen forests since the Oligocene in southwestern China for the first time, and demonstrated that except for the biological traits, the evolutionary history of the lineages are important factors impact the spatial genetic structures found in the evergreen oaks in YGP region. These results can provide important information on the formation of high biodiversity

level in southeast Himalayas, as well as conservation and safeguard this unique ecosystem on the background of global climate change.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm. nih.gov/, MW201294-MW201314, MW201315- MW201334, and MW201335- MW201352.

AUTHOR CONTRIBUTIONS

MD and X-LJ conceived and designed the experiments and were responsible for field collections and specimen identification. S-SZ performed the experiments. S-SZ and X-LJ analyzed the data. MD, S-SZ, and X-LJ wrote and revised the manuscript. Q-JH gave much advice on the details of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021. 774232/full#supplementary-material

REFERENCES

Akhmetiev, M. A., and Zaporozhets, N. I. (2014). Paleogene events in Central Eurasia: their role in the flora and vegetation cover evolution, migration of phytochore boundaries, and climate changes. *Stratigr. Geol. Correlat.* 22, 312–335.

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

Bandelt, H. J., Forster, P., and Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evolut.* 16, 37–48.

Bobrowski, M., and Schickhoff, U. (2017). Why input matters: Selection of climate data sets for modelling the potential distribution of a treeline species in the Himalayan region. *Ecol. Model.* 359, 92–102.

Bouchal, J. M., Mayda, S., Zetter, R., Grímsson, F., Akgün, F., and Denk, T. (2017). Miocene palynofloras of the Tınaz lignite mine, Muğla, southwest Anatolia: Taxonomy, palaeoecology and local vegetation change. *Rev. Palaeobot. Palynol.* 243, 1–36. doi: 10.1016/j.revpalbo.2017.02.010

Brown, J. L. (2014). SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods Ecol. Evolut.* 5, 694–700. doi: 10.7717/peerj.4095

Cavender-Bares, J., Gonzalez-Rodriguez, A., Eaton, D. A. R., Hipp, A. L., Beulke, A., and Manos, P. S. (2015). Phylogeny and biogeography of the American live oaks (*Quercus* subsection *Virentes*): a genomic and population genetics approach. *Mol. Ecol.* 24, 3668–3687. doi: 10.1111/mec.13269

Chatterjee, S., Goswami, A., and Scotese, C. R. (2013). The longest voyage: Tectonic, magmatic, and paleoclimatic evolution of the Indian plate during its northward flight from Gondwana to Asia. *Gondwana Res.* 23, 238–267. doi: 10.1016/j.gr.2012.07.001

Che, J., Zhou, W. W., Hu, J. S., Yan, F., Papenfuss, T. J., Wake, D. B., et al. (2010). Spiny frogs (*Paini*) illuminate the history of the Himalayan region and Southeast Asia. *Proc. Natl. Acad. Sci.* 107, 13765–13770. doi: 10.1073/pnas. 1008415107

Chen, D. M., Zhang, X. X., Kang, H. Z., Sun, X., Yin, S., Du, H. M., et al. (2012). Phylogeography of *Quercus variabilis* based on chloroplast DNA sequence in East Asia: multiple glacial refugia and mainland-migrated island populations. *PLoS One* 7:e47268. doi: 10.1371/journal.pone.0047268

Chen, J. H., Huang, Y., Brachi, B., Yun, Q. Z., Zhang, W., Lu, W., et al. (2019). Genome-wide analysis of Cushion willow provides insights into alpine plant divergence in a biodiversity hotspot. *Nat. Commun.* 10, 5230–5230. doi: 10. 1038/s41467-019-13128-y

Chen, X. D., Yang, J., Feng, L., Zhou, T., Zhang, H., Li, H. M., et al. (2020). 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* 16:2.

Chen, Y., Yan, M. D., Fang, X. M., Song, C. H., Zhang, W. L., Zan, J. B., et al. (2017). Detrital zircon U–Pb geochronological and sedimentological study of the Simao Basin, Yunnan: Implications for the Early Cenozoic evolution of the Red River. *Earth Planet. Sci. Lett.* 476, 22–33. doi: 10.1016/j.epsl.2017.07.025

Clark, M. K., Schoenbohm, L. M., Royden, L. H., Whipple, K. X., Burchfiel, B. C., Zhang, X., et al. (2004). Surface uplift, tectonics, and erosion of eastern Tibet from large-scale drainage patterns. *Tectonics* 23:TC1006.

Deng, J., Wang, Q. F., Li, G. J., and Santosh, M. (2014). Cenozoic tectono-magmatic and metallogenic processes in the Sanjiang region, southwestern China. *Earth Sci. Rev.* 138, 268–299. doi: 10.1016/j.earscirev.2014.05.015

Denk, T., Velitzelos, D., Güner, T. H., Bouchal, J. M., Grímsson, F., and Grimm, G. W. (2017). Taxonomy and palaeoecology of two widespread western Eurasian Neogene sclerophyllous oak species: *Quercus drymeja* Unger and *Q. mediterranea* Unger. *Rev. Palaeobot. Palynol.* 241, 98–128. doi: 10.1016/j. revpalbo.2017.01.005

Di, H. Z., Deng, B., Zhao, G. P., Ye, Y. H., and Qiu, J. W. (2018). The Evolution of the river system on the Yunnan-Guizhou plateau and formation process of the plateau based on modern river sediment. *Acta Geol. Sichuan* 38, 536–541.

Doyle, J. J., and Doyle, J. L. (1987). A rapid DNA isolation precedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* 19, 11–15.

Du, F. K., Hou, M., Wang, W. T., Mao, K. S., and Hampe, A. (2017). Phylogeography of *Quercus aquifolioides* provides novel insights into the Neogene history of a major global hotspot of plant diversity in south-west China. *J. Biogeogr.* 44, 294–307.

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294x.2005.02553.x

Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Feng, X. Y., Wang, Y. H., and Gong, X. (2014). Genetic diversity, genetic structure and demographic history of *Cycas simplicipinna* (Cycadaceae) assessed by DNA sequences and SSR markers. *BMC Plant Biol.* 14:187. doi: 10.1186/1471-2229-14-187

Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147, 915–925. doi: 10.1093/genetics/147.2.915

Fyhn, M. B. W., Boldreel, L. O., and Nielsen, L. H. (2009). Geological development of the Central and South Vietnamese margin: Implications for the establishment of the South China Sea, Indochinese escape tectonics and Cenozoic volcanism. *Tectonophysics* 478, 184–214. doi: 10.1016/j.tecto.2009.08.002

Gao, L. M., MÖller, M., Zhang, X. M., Hollingsworth, M. L., Liu, J., Mill, R. R., et al. (2007). High variation and strong phylogeographic pattern among *cp*DNA haplotypes in Taxus wallichiana (Taxaceae) in China and North Vietnam. *Mol. Ecol.* 16, 4684–4698. doi: 10.1111/j.1365-294X.2007.03537.x

Gao, H., Williamson, S., and Bustamante, C. D. (2007). A markov chain monte carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype. *Genetics* 176, 1635–1651. doi: 10.1534/genetics.107. 072371

Gong, W., Chen, C., Dobeš, C., Fu, C. X., and Koch, M. A. (2008). Phylogeography of a living fossil: Pleistocene glaciations forced *Ginkgo biloba* L. (Ginkgoaceae) into two refuge areas in China with limited subsequent postglacial expansion. *Mol. Phylogenet. Evolut.* 48, 1094–1105. doi: 10.1016/j.ympev.2008.05.003

Gong, Y. Q., Zhan, Q. Q., Khang Sinh, N., Hiep Tien, N., Wang, Y. H., and Gong, X. (2015). The historical demography and genetic variation of the endangered *Cycas multipinnata* (Cycadaceae) in the Red River region, examined by chloroplast DNA sequences and microsatellite markers. *PLoS One* 10:e0117719. doi: 10.1371/journal.pone.0117719

Govaerts, R., and Frodin, D. G. (1998). *World Checklist and Bibliography of Fagales (Betulaceae, Corylaceae, Fagaceae and Ticodendraceae)*. London: Kew Publishing.

Grivet, D., Heinze, B., Vendramin, G. G., and Petit, R. J. (2001). Genome walking with consensus primers: application to the large single copy region of chloroplast DNA. *Mol. Ecol. Notes* 1, 345–349. doi: 10.1046/j.1471-8278.2001. 00107.x

Guner, T. H., Bouchal, J. M., Kose, N., Goktas, F., Mayda, S., and Denk, T. (2017). Landscape heterogeneity in the Yatagan Basin (southwestern Turkey) during the middle Miocene inferred from plant macrofossils. *Palaeontogr. Abteilung B Palaeophytol. Palaeobot. Palaeophytol.* 296, 113–171. doi: 10.1127/palb/296/ 2017/113

Hall, R., Hattum, M. W. A. V., and Spakman, W. (2008). Impact of India–Asia collision on SE Asia: The record in Borneo. *Tectonophysics* 451, 0–389.

He, M. Y., Zheng, H. B., Clift, P. D., Bian, Z., Yang, Q., Zhang, B. H., et al. (2021). Paleogene Sedimentary Records of the Paleo-Jinshajiang (Upper Yangtze) in the Jianchuan Basin, Yunnan, SW China. *Geochem. Geophys. Geosyst.* 22:e2020GC009500.

Hernawan, U. E., Van Dijk, K. J., Kendrick, G. A., Feng, M., Biffin, E., Lavery, P. S., et al. (2017). Historical processes and contemporary ocean currents drive genetic structure in the seagrass Thalassia hemprichii in the Indo-Australian Archipelago. *Mol. Ecol.* 26, 1008–1021. doi: 10.1111/mec.13966

Hipp, A. L., Manos, P. S., Hahn, M., Avishai, M., Bodenes, C., Cavender-Bares, J., et al. (2020). Genomic landscape of the global oak phylogeny. *New Phytol.* 226, 1198–1212.

Huang, C. C., Chang, Y. T., and Bartholomew, B. (1999). "Fagaceae," in *Flora of China*, Vol. 4, eds C. Y. Wu and P. H. Raven (Beijing: Science Press and Missouri Botanical Garden Press), 380–400.

Huchon, P., Pichon, X. L., and Rangin, C. (1994). Indochina Peninsula and the collision of India and Eurasia. *Geology* 22, 27–30.

Hulce, D., Li, X., Snyder-Leiby, T., and Johathan Liu, C. S. (2011). GeneMarker® Genotyping Software: Tools to Increase the Statistical Power of DNA Fragment Analysis. *J. Biomol. Techniq. JBT* 22, S35–S36.

Jakobsson, M., and Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23, 1801–1806. doi: 10.1093/ bioinformatics/btm233

Ji, Y. H., Liu, C. K., Landis, J. B., Deng, M., and Chen, J. H. (2020). Plastome phylogenomics of *Cephalotaxus* (Cephalotaxaceae) and allied genera. *Ann. Bot.* 127, 697–708. doi: 10.1093/aob/mcaa201

Jiang, X. L., An, M., Zheng, S. S., Deng, M., and Su, Z. H. (2018). Geographical isolation and environmental heterogeneity contribute to the spatial genetic patterns of *Quercus kerrii* (Fagaceae). *Heredity* 120, 219–233. doi: 10.1038/ s41437-017-0012-7

Jiang, X. L., Deng, M., and 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* 12:104.

Jiang, X. L., Hipp, A. L., Deng, M., Su, T., Zhou, Z. K., and Yan, M. X. (2019). East Asian origins of European holly oaks (*Quercus* section *Ilex* Loudon) via the Tibet-Himalaya. *J. Biogeogr.* 46, 2203–2214. doi: 10.1111/jbi.13654

Kai, H., and Jiang, X. L. (2014). Sky islands of southwest China I: an overview of phylogeographic patterns. *Chin. Sci. Bull.* 59, 585–597.

Kalinowski, S. T. (2005). HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. *Mol. Ecol. Notes* 5, 187–189.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., et al. (2017). Climatologies at high resolution for the Earth land surface areas. *Sci. Data* 4:170122.

Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evolut.* 35, 1547–1549. doi: 10.1093/molbev/msy096

Li, J. G., Wu, Y. X., Batten, D. J., and Lin, M. Q. (2019). Vegetation and climate of the central and northern Qinghai–Xizang plateau from the Middle Jurassic to the end of the Paleogene inferred from palynology. *J. Asian Earth Sci.* 175, 35–48. doi: 10.1016/j.jseaes.2018.08.012

Li, S. H., Advokaat, E. L., Van Hinsbergen, D. J. J., Koymans, M., Deng, C. L., and Zhu, R. X. (2017). Paleomagnetic constraints on the Mesozoic-Cenozoic paleolatitudinal and rotational history of Indochina and South China: Review and updated kinematic reconstruction. *Earth Sci. Rev.* 171, 58–77.

Li, J. J., Hu, Z. M., Sun, Z. M., Yao, J. T., Liu, F. L., Fresia, P., et al. (2017). Historical isolation and contemporary gene flow drive population diversity of the brown alga Sargassum thunbergii along the coast of China. *BMC Evolution. Biol.* 17:246. doi: 10.1186/s12862-017-1089-6

Li, L., Abbott, R. J., Liu, B. B., Sun, Y. S., Li, L. L., Zou, J. B., et al. (2013). Pliocene intraspecific divergence and Plio-Pleistocene range expansions within *Picea likiangensis* (Lijiang spruce), a dominant forest tree of the Qinghai-Tibet Plateau. *Mol. Ecol.* 22, 5237–5255. doi: 10.1111/mec.12466

Lin, T. H., Lo, C. H., Chung, S. L., Hsu, F. J., Yeh, M. W., Lee, T. Y., et al. (2009). 40Ar/39Ar dating of the Jiali and Gaoligong shear zones: Implications for crustal deformation around the Eastern Himalayan Syntaxis. *J. Asian Earth Sci.* 34, 674–685.

Liu, F. Y., Wang, X. Q., Li, K., Sun, Y. Y., Zhang, Z. X., and Zhang, C. H. (2012). Species composition and diversity characteristics of *Quercus franchetii* communities in dry-hot valley of Jinsha River. *Guihaia* 32, 56–62.

Liu, F. Y., Zhang, Z. X., Wang, X. Q., Li, K., Sun, Y. Y., and Zhang, C. H. (2011). Effects of habitat heterogeneity on early growth of *Quercus franchetii* natural regeneration seedlings in the Jinsha river dry-hot valley. *Chin. J. Appl. Environ. Biol.* 17, 338–344. doi: 10.3724/sp.j.1145.2011.00338

Liu, J. Q., Gao, T. G., Chen, Z. D., and Lu, A. M. (2002). Molecular phylogeny and biogeography of the Qinghai-Tibet Plateau endemic *Nannoglottis* (Asteraceae). *Mol. Phylogenet. Evolut.* 23, 307–325. doi: 10.1016/s1055-7903(02)00 039-8

Liu, J. Q., Wang, Y. J., Wang, A. L., Hideaki, O., and Abbott, R. J. (2006). Radiation and diversification within the *Ligularia-Cremanthodium-Parasenecio* complex (Asteraceae) triggered by uplift of the Qinghai-Tibetan Plateau. *Mol. Phylogenet. Evolut.* 38, 31–49. doi: 10.1016/j.ympev.2005.09.010

Liu, J., Möller, M., Provan, J., Gao, L. M., Poudel, R. C., and Li, D. Z. (2013). Geological and ecological factors drive cryptic speciation of yews in a biodiversity hotspot. *New Phytol.* 199, 1093–1108. doi: 10.1111/nph.12336

Liu, R. B. (2019). *Phylogeography of Quercus cocciferoides complex*. Shanghai: Shanghai Normal University.

Lu, L. M., Mao, L. F., Yang, T., Ye, J. F., Liu, B., Li, H. L., et al. (2018). Evolutionary history of the angiosperm flora of China. *Nature* 554:234.

Luikart, G., Sherwin, W. B., Steele, B. M., and Allendorf, F. W. (1998). Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Mol. Ecol.* 7, 963–974. doi: 10.1046/j.1365-294x.1998.00 414.x

Manni, F., Guérard, E., and Heyer, E. (2004). Geographic patterns of (genetic, morphologic, linguistic, etc.) variation : how barriers can be detected by "Monmonier's algorithm". *Hum. Biol.* 76, 173–190. doi: 10.1353/hub.2004.0034

Meng, H. H., Su, T., Gao, X. Y., Li, J., Jiang, X. L., Sun, H., et al. (2017). Warm–Cold colonization: Response of oaks to uplift of the Himalaya–Hengduan Mountains. *Mol. Ecol.* 26, 3276–3294. doi: 10.1111/mec.14092

Meng, Y., Wen, J., Nie, Z. L., Sun, H., and Yang, Y. P. (2008). Phylogeny and biogeographic diversification of *Maianthemum* (Ruscaceae: Polygonatae). *Mol. Phylogenet. Evolut.* 49, 424–434. doi: 10.1016/j.ympev.2008.07.017

Ming, Q. Z. (2007). A study on the neotectonic division & environment evolution of Qing-Zang plateau & three parallel rivers area. *Yunnan Geol.* 26, 387–396. doi: 10.1093/rpd/ncw067

Okaura, T., Quang, N. D., Ubukata, M., and Harada, K. (2007). Phylogeographic structure and late Quaternary population history of the Japanese oak *Quercus mongolica* var. *crispula* and related species revealed by chloroplast DNA variation. *Genes Genet. Syst.* 82, 465–477. doi: 10.1266/ggs.82.465

Peakall, R., and Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28, 2537–2539. doi: 10.1093/bioinformatics/bts460

Petit, R. J., Bodénès, C., Ducousso, A., Roussel, G., and Kremer, A. (2003). Hybridization as a mechanism of invasion in oaks. *New Phytol.* 161, 151–164. doi: 10.1046/j.1469-8137.2003.00944.x

Petit, R. J., Csaikl, U. M., Bordacs, S., Burg, K., Coart, E., Cottrell, J., et al. (2002). Chloroplast DNA variation in European white oaks: Phylogeography and patterns of diversity based on data from over 2600 populations. *For. Ecol. Manage.* 156:5.

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

Phillips, S. J., Dudík, M., and Schapire, R. (2004). "A maximum entropy approach to species distribution modeling," in *Proceedings of the twenty-first international conference on Machine learning*, (Alberta: ICML).

Pons, O., and Petit, R. J. (1996). Measwring and testing genetic differentiation with ordered versus unordered alleles. *Genetics* 144, 1237–1245. doi: 10.1093/genetics/144.3.1237

Posada, D., and Crandall, K. A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14, 817–818. doi: 10.1093/bioinformatics/14.9.817

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155:945.

Qu, Y. H., Ericson, P. G. P., Quan, Q., Song, G., Zhang, R. Y., Gao, B., et al. (2014). Long-term isolation and stability explain high genetic diversity in the Eastern Himalaya. *Mol. Ecol.* 23, 705–720. doi: 10.1111/mec.12619

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A. (2018). Posterior summarization in bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* 67, 901–904. doi: 10.1093/sysbio/syy032

Ree, R. H., and Smith, S. A. (2008). Maximum likelihood inference of geographic range evolution by dispersal, local extinction, and cladogenesis. *Systemat. Biol.* 57, 4–14. doi: 10.1080/10635150701883881

Ree, R. H., Moore, B. R., Webb, C. O., and Donoghue, M. J. (2005). A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. *Evolution* 59, 2299–2311.

Renner, S. S. (2016). Available data point to a 4-km-high Tibetan Plateau by 40 Ma, but 100 molecular-clock papers have linked supposed recent uplift to young node ages. *J. Biogeogr.* 43, 1479–1487.

Replumaz, A., and Tapponnier, P. (2003). Reconstruction of the deformed collision zone Between India and Asia by backward motion of lithospheric blocks. *J. Geophys. Res. Solid Earth* 108:2285.

Rousset, F. (2008). genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Mol. Ecol. Resour.* 8, 103–106. doi: 10.1111/j.1471-8286.2007.01931.x

Royden, L. H., Burchfiel, B. C., and Hilst, R. D. (2008). The geological evolution of the Tibetan Plateau. *Science* 321, 1054–1058. doi: 10.1126/science.1155371

Rozas, J., Ferrer-Mata, A., Sánchez-Delbarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., et al. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evolut.* 34, 3299–3302. doi: 10.1093/molbev/msx248

Rüber, L., Britz, R., Kullander, S. O., and Zardoya, R. (2004). Evolutionary and biogeographic patterns of the Badidae (Teleostei: Perciformes) inferred from mitochondrial and nuclear DNA sequence data. *Mol. Phylogenet. Evolut.* 32, 1010–1022. doi: 10.1016/j.ympev.2004.04.020

Shao, C. C., Shen, T. T., Jin, W. T., Mao, H. J., Ran, J. H., and Wang, X. Q. (2019). Phylotranscriptomics resolves interspecific relationships and indicates multiple historical out-of-North America dispersals through the Bering Land

Bridge for the genus *Picea* (Pinaceae). *Mol. Phylogenet. Evolut.* 141:106610. doi: 10.1016/j.ympev.2019.106610

Shaw, J., Lickey, E. B., Beck, J. T., Farmer, S. B., Liu, W., Miller, J., et al. (2005). The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Am. J. Bot.* 92, 142–166. doi: 10.3732/ajb.92.1.142

Shen, L., Chen, X. Y., Zhang, X., Li, Y. Y., Fu, C. X., and Qiu, Y. X. (2005). Genetic variation of *Ginkgo biloba* L. (Ginkgoaceae) based on *cp*DNA PCR-RFLPs: inference of glacial refugia. *Heredity* 94, 396–401. doi: 10.1038/sj.hdy.6800616

Su, T., Farnsworth, A., Spicer, R. A., Huang, J., Wu, F. X., Liu, J., et al. (2019). No high Tibetan Plateau until the Neogene. *Sci. Adv.* 5:eaav2189.

Su, T., Liu, Y. S., Jacques, F. M. B., Huang, Y. J., Xing, Y. W., and Zhou, Z. K. (2013). The intensification of the East Asian winter monsoon contributed to the disappearance of *Cedrus* (Pinaceae) in southwestern China. *Quaternary Res.* 80, 316–325. doi: 10.1016/j.yqres.2013.07.001

Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., and Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolut.* 4:vey016. doi: 10.1093/ve/vey016

Taberlet, P., Gielly, L., Pautou, G., and Bouvet, J. (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Mol. Biol.* 17, 1105–1109. doi: 10.1007/bf00037152

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595. doi: 10.1093/genetics/123.3.585

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., and Shipley, P. (2004). Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol. Ecol. Notes* 4, 535–538. doi: 10.1111/j.1471-8286.2004.00684.x

Van Oppen, M. J. H., Peplow, L. M., Kininmonth, S., and Berkelmans, R. A. Y. (2011). Historical and contemporary factors shape the population genetic structure of the broadcast spawning coral, Acropora millepora, on the Great Barrier Reef. *Mol. Ecol.* 20, 4899–4914. doi: 10.1111/j.1365-294X.2011.05328.x

Vandergast, A. G., Perry, W. M., Lugo, R. V., and Hathaway, S. A. (2011). Genetic landscapes GIS Toolbox: tools to map patterns of genetic divergence and diversity. *Mol. Ecol. Resour.* 11, 158–161. doi: 10.1111/j.1755-0998.2010.02904.x

Wang, C. B., Wang, T., and Su, Y. J. (2014). Phylogeography of *Cephalotaxus oliveri* (Cephalotaxaceae) in relation to habitat heterogeneity, physical barriers and the uplift of the Yungui Plateau. *Mol. Phylogenet. Evolut.* 80, 205–216. doi: 10.1016/j.ympev.2014.08.015

Wang, F. Y., Gong, X., Hu, C. M., and Hao, G. (2008). Phylogeography of an alpine species *Primula secundiflora* inferred from the chloroplast DNA sequence variation. *J. Systemat. Evolut.* 46, 13–22.

Wang, J., Gao, P. X., Kang, M., Lowe, A. J., and Huang, H. W. (2009). Refugia within refugia: the case study of a canopy tree (*Eurycorymbus cavaleriei*) in subtropical China. *J. Biogeogr.* 36, 2156–2164.

Wang, Y. J., Liu, J. Q., and Miehe, G. (2007). Phylogenetic origins of the Himalayan endemic *Dolomiaea*, *Diplazoptilon* and *Xanthopappus* (Asteraceae: Cardueae) based on three DNA regions. *Ann. Bot.* 99, 311–322. doi: 10.1093/aob/mcl259

Wang, Z. W., Chen, S. T., Nie, Z. L., Zhang, J. W., Zhou, Z., Deng, T., et al. (2015). Climatic factors drive population divergence and demography: Insights based on the phylogeography of a riparian plant species endemic to the Hengduan Mountains and adjacent regions. *PLoS One* 10:e0145014. doi: 10.1371/journal.pone.0145014

Wangda, P., and Ohsawa, M. (2006). Structure and regeneration dynamics of dominant tree species along altitudinal gradient in a dry valley slopes of the Bhutan Himalaya. *For. Ecol. Manage.* 230, 136–150. doi: 10.1016/j.foreco.2006.04.027

Warren, D. L., Glor, R. E., and Turelli, M. (2010). ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography* 33, 607–611.

Xia, K., Daws, M. I., Stuppy, W., Zhou, Z.-K., and Pritchard, H. W. (2012). Rates of water loss and uptake in recalcitrant fruits of *Quercus* species are determined by pericarp anatomy. *PLoS One* 7:e47368. doi: 10.1371/journal.pone.0047368

Xing, Y. W., and Ree, R. H. (2017). Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot. *Proc. Natl. Acad. Sci.* 114, E3444–E3451. doi: 10.1073/pnas.1616063114

Xing, Y. W., Onstein, R. E., Carter, R. J., Stadler, T., and Peter Linder, H. (2014). Fossils and a large molecular phylogeny show that the evolution of species richness, generic diversity, and turnover rates are disconnected. *Evolution* 68, 2821–2832. doi: 10.1111/evo.12489

Xu, J., Deng, M., Jiang, X. L., Westwood, M., Song, Y. G., and Turkington, R. (2015). Phylogeography of *Quercus glauca* (Fagaceae), a dominant tree of East Asian subtropical evergreen forests, based on three chloroplast DNA interspace sequences. *Tree Genet. Genomes* 11:805.

Xu, J., Song, Y. G., Deng, M., Jiang, X. L., Zheng, S. S., and Li, Y. (2020). Seed germination schedule and environmental context shaped the population genetic structure of subtropical evergreen oaks on the Yun-Gui Plateau, Southwest China. *Heredity* 124, 499–513. doi: 10.1038/s41437-019-0283-2

Xu, T. T., Abbott, R. J., Milne, R. I., Mao, K., Du, F. K., Wu, G. L., et al. (2010). Phylogeography and allopatric divergence of cypress species (*Cupressus* L.) in the Qinghai-Tibetan Plateau and adjacent regions. *BMC Evolution. Biol.* 10:194. doi: 10.1186/1471-2148-10-194

Yan, M. D., and Chen, Y. (2018). Detrital zircon U-Pb age analyses of the Early Cenozoic sediments from the Simao Basin and evolution of the paleo-Red River drainage system. *Quat. Sci.* 38, 130–144. doi: 10.11928/j.issn.1001-7410.2018.01.11

Yang, J., Yang, J. X., and Chen, X. Y. (2012). A re-examination of the molecular phylogeny and biogeography of the genus *Schizothorax* (Teleostei: Cyprinidae) through enhanced sampling, with emphasis on the species in the Yunnan–Guizhou Plateau, China. *J. Zool. Syst. Evolution. Res.* 50, 184–191. doi: 10.1111/j.1439-0469.2012.00661.x

Ye, X. Y., Ma, P. F., Yang, G. Q., Guo, C., Zhang, Y. X., Chen, Y. M., et al. (2019). Rapid diversification of alpine bamboos associated with the uplift of the Hengduan Mountains. *J. Biogeogr.* 46, 2678–2689. doi: 10.1111/jbi.13723

Yu, Y., Harris, A. J., Blair, C., and He, X. J. (2015). RASP (Reconstruct Ancestral State in Phylogenies): A tool for historical biogeography. *Mol. Phylogenet. Evolut.* 87, 46–49. doi: 10.1016/j.ympev.2015.03.008

Yuan, Z., Jiang, J. B., Dong, Y., Zhao, Q., Gao, X., and Qiu, J. P. (2019). The dispersal and diversification of earthworms (Annelida: Oligochaeta) related to paleogeographical events in the Hengduan Mountains. *Eur. J. Soil Biol.* 94:103118.

Zhang, L. S. (2012). *Palaeogeography of China: the Formation of China's Natural Environment*. Beijing: Science Press.

Zhang, M. W., Rao, D. Q., Yang, J. X., Yu, G. H., and Wilkinson, J. A. (2010). Molecular phylogeography and population structure of a mid-elevation montane frog *Leptobrachium ailaonicum* in a fragmented habitat of southwest China. *Mol. Phylogenet. Evolut.* 54, 47–58. doi: 10.1016/j.ympev.2009.10.019

Zhang, T. C., Comes, H. P., and Sun, H. (2011). Chloroplast phylogeography of *Terminalia franchetii* (Combretaceae) from the eastern Sino-Himalayan region and its correlation with historical river capture events. *Mol. Phylogenet. Evolut.* 60, 1–12. doi: 10.1016/j.ympev.2011.04.009

Zhang, X. W., Li, Y., Liu, C. Y., Xia, T., Zhang, Q., and Fang, Y. M. (2015). Phylogeography of the temperate tree species *Quercus acutissima* in China: Inferences from chloroplast DNA variations. *Biochem. Syst. Ecol.* 63, 190–197. doi: 10.1016/j.bse.2015.10.010

Zhao, J. L., Xia, Y. M., Cannon, C. H., Kress, W. J., and Li, Q. J. (2016). Evolutionary diversification of alpine ginger reflects the early uplift of the Himalayan–Tibetan Plateau and rapid extrusion of Indochina. *Gondwana Res.* 32, 232–241. doi: 10.1016/j.gr.2015.02.004

Zheng, H. B. (2015). Birth of the Yangtze River: age and tectonic-geomorphic implications. *Natl. Sci. Rev.* 2, 438–453. doi: 10.1093/nsr/nwv063

Zheng, S. S. (2021). *Phylogeography of Quercus franchetii complex (Fagaceae)*. Shanghai: Shanghai Institute of Technology.