forests

MDPI

Article

Genetic Diversity, Population Structure, and Conservation Units of *Castanopsis sclerophylla* (Fagaceae)

Shuang Chen, Risheng Chen, Xiaorong Zeng, Xing Chen, Xinsheng Qin, Zhuoxin Zhang and Ye Sun *

Guangdong Key Laboratory for Innovative Development and Utilization of Forest Plant Germplasm, College of Forestry and Landscape Architecture, South China Agricultural University, Guangzhou 510642, China
* Correspondence: sun-ye@scau.edu.cn; Tel./Fax: +86-20-85280256

Abstract: *Castanopsis sclerophylla* (Lindl.) Schott. is a canopy tree species of evergreen broad-leaved forests in subtropical China. In this study, the genetic diversity and population structure of *C. sclerophylla* were investigated by using chloroplast DNA sequences and nuclear microsatellite markers. Permutation tests with chloroplast DNA sequences indicated the presence of phylogeographic structure in *C. sclerophylla*. Based on nuclear microsatellite markers, Bayesian clustering analysis revealed eastern-to-western differentiation in *C. sclerophylla*, and the analysis of molecular variance suggested population divergence has arisen along the Xuefeng, Luoxiao, and Wuyi mountain ranges. The approximate Bayesian computation demonstrated that the genetic diversity pattern of *C. sclerophylla* could be explained by geographic isolation followed by secondary contact. Ecological niche modelling showed that distribution of *C. sclerophylla* shrank southward at the Last Glacial Maximum and expanded northward at the Mid Holocene. These results suggested that the uplift of the Xuefeng, Luoxiao, and Wuyi mountain ranges and the interglacial–glacial climate change shaped the genetic diversity of *C. sclerophylla*. The Luoxiao mountain range should be considered as a key conservation unit of *C. sclerophylla* due to its higher level of genetic diversity. Our study supplies important information for prioritizing the conservation and sustainable utilization of *C. sclerophylla*, and provides insight on the dynamics of evergreen broad-leaved forests in subtropical China.

Keywords: *Castanopsis sclerophylla*; genetic diversity; chloroplast DNA sequence; microsatellite

check for updates

Citation: Chen, S.; Chen, R.; Zeng, X.; Chen, X.; Qin, X.; Zhang, Z.; Sun, Y. Genetic Diversity, Population Structure, and Conservation Units of *Castanopsis sclerophylla* (Fagaceae). *Forests* **2022**, *13*, 1239. https://doi.org/10.3390/f13081239

Academic Editor: Arndt Hampe

Received: 24 June 2022
Accepted: 2 August 2022
Published: 5 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Subtropical evergreen broad-leaved forest (EBLF) is one of the most important vegetation types in China with a range from 22 to 34° N latitude and 99 to 123° E longitude [1]. Range shift of the EBLF during the Neogene and Quaternary climate changes remains a subject of dispute [2]. During the Last Glacial Maximum (LGM), the climate of this region was thought to be 4–6 °C cooler and 400–600 mm/year drier than today [3]. It was argued that the EBLF in subtropical China was forced to retreat southward as far as c. 24° N at LGM and re-colonized northward to mid-latitude after the Holocene [4,5]. Several studies on evergreen shrub or tree species revealed evidence for their extensive expansion-contraction across the EBLF [2,6,7]. However, some evergreen species displayed a pattern of multiple northern refugia with regional post-glacial expansion [2,8].

Population demographic history shape extant genetic diversity patterns thus have great impact on population genetic structure. Extensive expansion–contraction are expected to cause a significant negative correlation between genetic diversity and latitude [7], but long-term isolation in multiple refugia followed by limited-range expansion would provide opportunity for population divergence and lead to high population genetic structure [6,9]. Genetic diversity is an important part of biodiversity and the basis of species survival and evolution. The economic and ecological value of plants depends on the unique gene composition, and the stability and evolutionary potential of plant communities rely on the genetic diversity of their constituents. Biodiversity hot spots, such as subtropical EBLF

in China, have highest priority in ecosystem conservation due to their great biodiversity and vulnerability. Thus, great importance is placed on studying the genetic diversity and population structure of the EBLF components in subtropical China for their scientific conservation and utilization.

Castanopsis sclerophylla (Lindl.) Schott. is a canopy tree species of the EBLF in subtropical China. It is mainly distributed in areas south of the Yangtze River and north of the Nanling mountains in China. The wood of *C. sclerophylla* is dense, tough, and elastic, thus it is widely used in industrial and agricultural production. The nuts of *C. sclerophylla* are rich in starch with a unique flavor and aroma, and have a long history of consumption as a specialty food in some provinces of China [10]. Long-term utilization and habitat fragmentation are expected to reduce genetic diversity of *C. sclerophylla* because of increased genetic drift and inbreeding [11]. Only small-scale studies on the genetic diversity of *C. sclerophylla*, such as in the region of Qiandao Lake, have been reported [12–14], and the results indicated that pollen-mediated gene flow could maintain within-population genetic diversity and low differentiation among populations. Therefore, there is an urgent need to comprehensively investigate the genetic diversity of *C. sclerophylla* from the whole distribution range.

Microsatellites refer to simple sequence repeats (SSRs), which are very abundant in eukaryotic genomes [15]. Due to the advantages of Mendelian codominant inheritance and high polymorphism and repeatability, SSRs are widely used to study the genetic diversity, population structure, and phylogeographic history of broad-leaved forest plants and their genomic response to geo-climate changes [6,14,16]. In contrast, organelle DNA sequences have been widely applied in phylogeographic studies because of uniparental inheritance and low mutation rate [17,18]. In this study, we used nuclear SSR markers and chloroplast DNA sequences to study the genetic diversity and population structure of *C. sclerophylla* in subtropical China with the specific aim of inferring the population evolutionary history of this canopy tree species in EBLF.

2. Materials and Methods

A total of 320 individuals of *C. sclerophylla* were sampled from 15 natural populations (Table 1, Figure 1). The sampled individuals were at least 20 m apart. Fresh leaves were collected and immediately dried with silica gel. Total genomic DNA was extracted from 100 mg of leaf tissue using the DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany), according to the manufacturer's instructions. DNA quality and concentration were evaluated by using Nano Drop TM 2000 Spectrophotometers (Thermo Scientific, Waltham, MA, USA) and agarose gel electrophoresis.

Chloroplast intergenic spacers of psbA-trnH and trnM-trnV were amplified and sequenced using previously reported primers [19,20]. The total volume of the polymerase chain reaction (PCR) was 30 μL, which contained 1 × ES Tag Master Mix (Cwbiotech, Beijing, China), 0.5 μM each of primer, and 20 ng of DNA. PCR was conducted on a programmable thermal cycler (EDC-810, Suzhou, China) using 33 cycles of 9 °C for 45 s, 58 °C for 45 s and 7 °C for 45 s. An initial denaturation at 9 °C for 5 min and a final extension at 7 °C for 7 min were added.

A total of 32 pairs of nuclear SSR primers with high polymorphism and repeatability were screened from those reported in previous studies [21,22]. Forward primers were labeled with fluorochromes including TAMRA, HEX, 6-FAM, and ROX. The Type-it Microsatellite PCR Kit (QIAGEN, Hilden, Germany) was used to prepare the multiple PCR with a total volume of 10 μL, which contained 1 × PCR Master Mix, 1 × Q-Solution, 0.2 μM each of primer, and 20 ng of DNA. The initial denaturation was carried out at 9 °C for 5 min, which was followed by 28 cycles of 95 °C for 30 s, 57 °C for 90 s, and 72 °C for 30 s, and a final extension at 60 °C for 30 min. PCR products were visualized on an ABI-3730XL fluorescence sequencer (Applied Biosystems, Foster City, CA, USA) by using LIZ500 as an internal size standard.

Table 1. Location and sample size of 15 *Castanopsis sclerophylla* populations.

Region Group	Population Location	Population Code	Latitude (N)	Longitude (E)	Sample Size (Sequenced Individuals)
Western	Xiaojiaqiao	XJQ	27°19′	110°39′	24 (18)
	Longhui	LH	27°19′	111°01′	20 (18)
	Xinning	XN	26°38′	111°00′	18 (18)
Central	Yuelushan	YLS	28°18′	112°94′	20 (8)
	Yanling	YL	26°34′	113°50′	20 (18)
	Wenming	WM	25°29′	113°26′	22 (18)
	Xiushui	XS	28°55′	114°43′	22 (18)
Eastern	Yuanqian	YQ	26°48′	115°01′	24 (8)
	Ganjiangyuan	GJY	26°04′	116°19′	22 (18)
	Liancheng	LC	25°36′	116°37′	24 (8)
	Dongxiang	DX	28°03′	116°50′	22 (8)
	Dayelin chang	DY	28°49′	118°11′	16 (8)
	Gutian shan	GTS	29°15′	118°09′	18 (18)
	Qiandaohu	QDH	29°30′	118°55′	24 (8)
	Shangyuantou	SYT	28°13′	118°35′	24 (18)

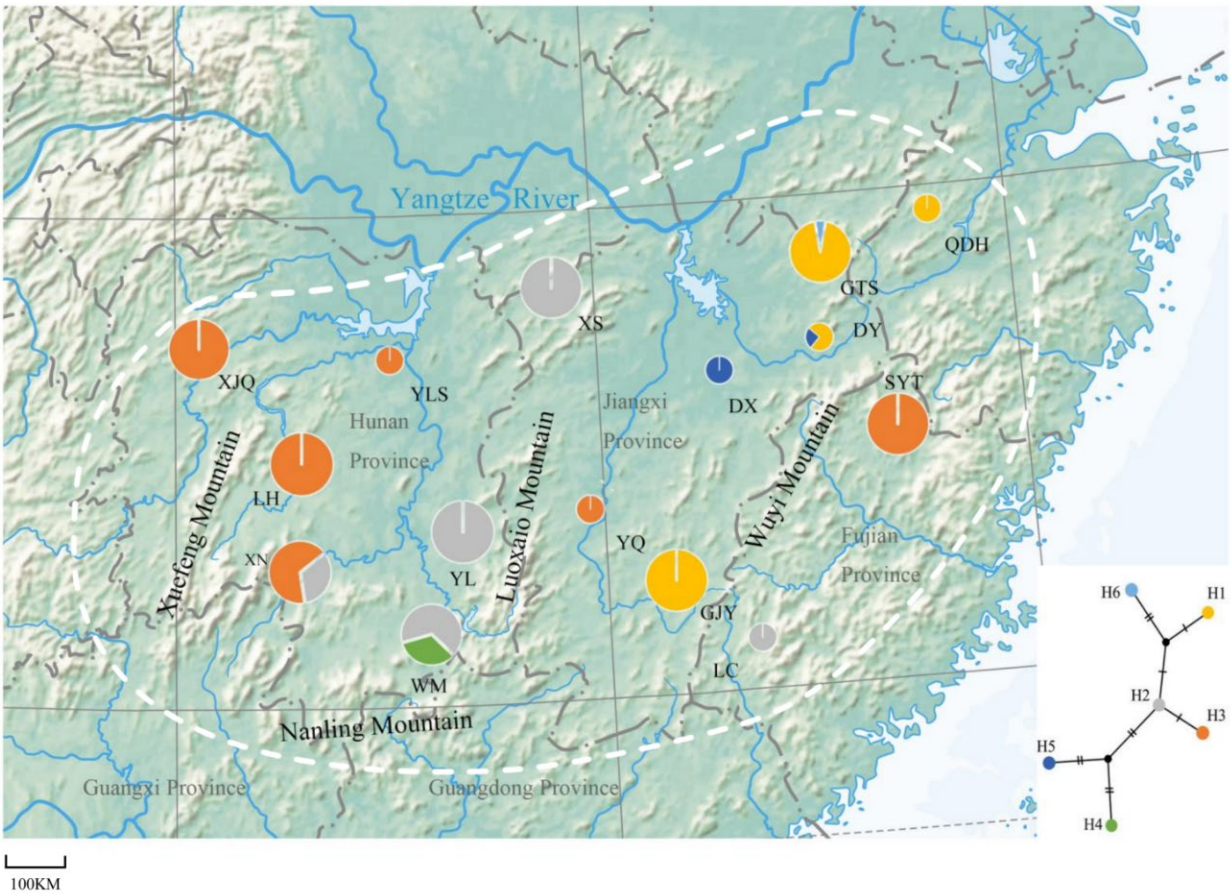


Figure 1. Locations and haplotypes distribution of sampled populations of *Castanopsis sclerophylla*. Circle size is proportional to the number of individuals sequenced. The dashed white line shows the current distribution range of *Castanopsis sclerophylla*. The inset at the lower right corner shows the relationships among haplotypes.

Chloroplast DNA sequences of trnM-trnV (GenBank accession numbers: MT635093-MT635125) and psbA-trnH (GenBank accession numbers: MT635060-MT635092) were

checked using Bioedit7.2.5 [23]. Multiple alignments were carried out with MEGA7 [24]. Nucleotide diversity (π) and haplotype diversity (H_d) were calculated with DnaSP v5 [25]. Population differentiation parameters (G_{ST} and N_{ST}) were estimated with PERMUT [26], and phylogeographic structure was inferred by testing for significant differences between G_{ST} and N_{ST} with 1000 permutations. A median-joining Network of haplotypes was constructed with Popart 1.7 [27].

The number of alleles detected (A), allelic richness (A_R), observed heterozygosity (H_O), gene diversity (H), gene diversity within populations (H_S), total gene diversity (H_T), inbreeding coefficient (F_{IS}), and genetic differentiation among populations (G_{ST}, F_{ST} and R_{ST}) were calculated for SSRs with FSTAT 2.9.4 [28]. Hardy–Weinberg equilibrium (HWE) was tested by allele randomizations with 10,000 permutations, and SSRs deviated from HWE were excluded from further analysis. Population genetic structure were inferred with a model-based Bayesian clustering approach implemented in STRUCTURE 2.3.4 [29] by assuming an admixture model and correlated allele frequencies among populations. Ten independent runs were conducted for each K value (from 1 to 15) with 50,000 burn-in steps followed by 500,000 Markov Chain Monte Carlo simulations. The optimal K value was determined by calculating the ΔK with Structure Harvester [30]. The average matrix of ancestry membership proportions was obtained over the 10 runs using CLUMPP v1.1.2 [31]. Analysis of molecular variance (AMOVA) was conducted with ARLEQUIN v3.5 [32] to partition genetic variation within populations, among populations, and among groups. According to the geographical distribution and population genetic structure, the 15 populations were clustered into Western (XJQ, LH, and XN), Central (YLS, YL, WM, and XS) and Eastern (YQ, GJY, LC, DX, DY, GTS, QDH, and SYT) group.

BOTTLENECK 1.2.02 [33] was used to detect the signal of recent reductions in effective population size. Heterozygosity excess within population was assessed by Wilcoxon signed-rank test with 10,000 simulations, and the expected heterozygosity under mutation-drift equilibrium was calculated using both stepwise mutation model (SMM) and two-phased mutation model (TPM). The latter model comprised 95% single-step and 5% multiple-step mutation. Eight hypothesized evolutionary scenarios (Figure A1) for the Western, Central, and Eastern groups were statistically assessed with DIYABC 2.1.0 [34] by using summary statistics including mean number of alleles, mean gene diversity, genetic differentiation coefficient, classification index, and shared allele distance among samples. One million simulations were performed for each scenario, and the 10% of simulated data sets closest to observed data were used to evaluate posterior probability and distributions of parameters with logistic regression.

In addition to our sampling locations, 119 sites where *C. sclerophylla* has been recorded were obtained from the Chinese Virtual Herbarium. The 19 climatic variables of each site were downloaded from Worldclim 1.4 [35], including different periods such as the Last-interglacial (LIG, 130 Ka BP), the Last Glacial Maximum (LGM, 22 Ka BP), the Mid Holocene (MH, 6 Ka BP), and the present (1996–1990). Ecological niche modelling was performed with MAXENT v3.3.3 [36]. The correlation between climatic variables were evaluated by calculating Pearson's correlation coefficient with SPSS v26.0 (IBM Corp. in Armonk, NY, USA). The most important set of uncorrelated climatic variables were determined with a correlation threshold of 0.70 and contribution threshold of 3%. The max temperature of warmest month (BIO05), temperature annual range (BIO07), mean temperature of driest quarter (BIO09), and precipitation of coldest quarter (BIO19) were finally kept for ecological niche modelling. The accuracy of the model was evaluated with the area under curve (AUC) of a receiver operating characteristic plot [37]. The potential distribution of *C. sclerophylla* in the past was inferred by projecting niche model onto the climate of different periods such as the LIG, LGM, and MH. According to Jenks Natural Breaks classification [38], the distribution areas were grouped into five classes: non-suitable, extremely low-suitable, low-suitable, middle-suitable, and high-suitable area.

3. Results

An alignment of 1171 bp of chloroplast DNA was obtained in which six haplotypes (H1-H6) and 16 variable sites were discriminated (Table 2). Variable sites included four insertion/deletion polymorphism and 12 SNPs. Haplotype diversity (H_d) and nucleotide diversity (π) were 0.706 and 0.00162, respectively. Permutation tests showed that there was a significant difference between G_{ST} and N_{ST} (N_{ST} = 0.909, G_{ST} = 0.861, p = 0.047), indicating presence of phylogeographic structure in *C. sclerophylla*. The three most common haplotypes were H3, H2, and H1, which mainly occurred in the western, central, and eastern regions (Figure 1), respectively. The evolutionary relationships of haplotypes were shown on a median-joining network (Figure 1), in which H2 was on the interior with multiple connections and had a close relationship to H1 and H3.

Table 2. Chloroplast DNA sequence variation and haplotypes revealed in this study.

Region Group	Population Code	Haplotype	Frequency	Variable Site															
				Psba-trnH								Trnm-trnV							
				1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
Western	XJQ	H3	18	-	T	A	A	-	TCTTTG	C	A	C	-	C	T	#	C	A	T
	LH	H3	18	-	T	A	A	-	TCTTTG	C	A	C	-	C	T	#	C	A	T
	XN	H2	6	-	T	A	A	-	TCTTTG	C	A	C	-	A	G	#	C	A	-
		H3	12	-	T	A	A	-	TCTTTG	C	A	C	-	C	T	#	C	A	T
Central	YLS	H3	8	-	T	A	A	-	TCTTTG	C	A	C	-	C	T	#	C	A	T
	YL	H2	18	-	T	A	A	-	TCTTTG	C	A	C	-	A	G	#	C	A	-
	WM	H2	12	-	T	A	A	-	TCTTTG	C	A	C	-	A	G	#	C	A	-
		H4	6	-	G	A	A	-	-	C	G	C	-	C	G	#	C	A	-
	XS	H2	18	-	T	A	A	-	TCTTTG	C	A	C	-	A	G	#	C	A	-
Eastern	YQ	H3	8	-	T	A	A	-	TCTTTG	C	A	C	-	C	T	#	C	A	T
	GJY	H1	18	TT	T	A	A	-	TCTTTG	T	A	C	-	C	G	#	A	A	-
	LC	H2	8	-	T	A	A	-	TCTTTG	C	A	C	-	A	G	#	C	A	-
	DX	H5	8	-	G	T	T	T	TCTTTG	C	A	A	-	C	T	#	A	A	-
	DY	H1	6	TT	T	A	A	-	TCTTTG	T	A	C	-	C	G	#	A	A	-
		H5	2	-	G	T	T	T	TCTTTG	C	A	A	-	C	T	#	A	A	-
	GTS	H1	17	TT	T	A	A	-	TCTTTG	T	A	C	-	C	G	#	A	A	-
		H6	1	-	T	A	A	-	TCTTTG	C	A	C	ATTTTGA	C	G	-	A	G	-
	QDH	H1	8	TT	T	A	A	-	TCTTTG	T	A	C	-	C	G	#	A	A	-
	SYT	H3	18	-	T	A	A	-	TCTTTG	C	A	C	-	C	T	#	C	A	T

-, deletion; #, insertion (AAATGTAAATGGACGCCCGGATTGGACCGAACCT).

Sixteen out of 32 SSR loci significantly deviated from the HWE ($p < 0.01$) due to the presence of null alleles and were excluded from further analysis. A total of 163 alleles were revealed at the 16 retained microsatellite loci (Table 3). The observed number of alleles per locus were from 2–21 with an average of 10. The observed heterozygosity (H_O) varied from 0.203 to 0.814. The within-population gene diversity (H_S) and the overall gene diversity (H_T) were 0.219–0.822 and 0.23–0.893, respectively. Over the 16 loci, the values of F_{IS} and F_{ST} were 0.016 and 0.085, respectively. At the population level, the average number of alleles (A) ranged from 3.9 to 6.1 (Table 4), allele richness (A_R) differed from 3.693 to 5.513, and genetic diversity (H) varied from 0.464 to 0.641. Population WM harbored the highest genetic diversity (A = 6.1, A_R = 5.513, and H = 0.614), while population LC showed the lowest genetic diversity (A = 3.9, A_R = 3.693, and H = 0.464). The inbreeding indices (F_{IS}) were from −0.064 to 0.117. The signal of recent reductions in effective population size was revealed in all 15 populations of *C. sclerophylla* under both TPM and SMM model ($p < 0.05$, Table 4).

Table 3. Genetic diversity of *Castanopsis sclerophylla* at the 16 retained microsatellite loci.

Locus	A	AR	HO	HS	HT	FIS	FST	RST	GST
CC30080	7	3.071	0.469	0.504	0.545	0.068	0.08	0.1	0.075
CC34976	13	9.577	0.796	0.782	0.867	−0.018	0.103	0.116	0.098
CC935	15	5.424	0.573	0.628	0.677	0.087	0.079	0.013	0.072
CC6538	12	6.615	0.68	0.678	0.73	−0.003	0.08	0.125	0.072
CcC02022	13	8.742	0.696	0.749	0.802	0.071	0.072	0.149	0.066
CC39174	8	3.665	0.531	0.51	0.582	−0.042	0.132	0.044	0.124
CsCAT14	17	10.331	0.768	0.765	0.825	−0.004	0.077	0.052	0.073
CC39198	2	2	0.496	0.485	0.5	−0.023	0.035	0.035	0.031
CC704−1	4	2.721	0.282	0.276	0.306	−0.022	0.101	0.059	0.099
CFA71	21	11.327	0.814	0.822	0.893	0.01	0.085	0.086	0.08
CC42621	6	3.007	0.203	0.219	0.23	0.075	0.052	0.06	0.048
CS24	10	6.988	0.705	0.698	0.741	−0.01	0.062	0.109	0.058
CC3754	6	3.83	0.401	0.447	0.541	0.104	0.182	0.229	0.173
CT161	6	3.312	0.339	0.342	0.365	−0.007	0.067	0.062	0.063
CS20	16	7.887	0.729	0.752	0.822	0.03	0.088	0.162	0.085
CT128	7	4.514	0.527	0.502	0.519	−0.05	0.033	0.052	0.032
Average	10	4.733	0.563	0.572	0.622	0.016	0.085	0.091	0.079

A: alleles observed; A_R: allelic richness for 16 diploid individuals; H_O: observed heterozygosity; H_S: gene diversity within populations; H_T: gene diversity in total population; F_{IS}: inbreeding index; F_{ST}: genetic differentiation among populations; R_{ST}: genetic differentiation among populations under a stepwise mutation model; and G_{ST}: the proportion of the total genetic diversity occurred among population.

Table 4. Genetic diversity parameter in 15 populations of *Castanopsis sclerophylla* at the 16 SSR loci.

Region Group	Population Code	A	A_R	H	F_{IS}	Wilcoxon Signed-Rank Test under TPM Model	Wilcoxon Signed-Rank Test under SMM Model
Western	XJQ	5.4	4.986	0.595	0.007	$p = 0.00011$	$p = 0.00011$
	LH	4.9	4.626	0.565	−0.04	$p = 0.00004$	$p = 0.00008$
	XN	4.8	4.62	0.609	0.094	$p = 0.00001$	$p = 0.00001$
Central	YLS	5.4	5.186	0.641	0.03	$p = 0.00001$	$p = 0.00001$
	YL	5.3	5.002	0.571	0.047	$p = 0.00005$	$p = 0.00042$
	WM	6.1	5.513	0.614	0.117	$p = 0.00655$	$p = 0.01677$
	XS	4.9	4.643	0.577	−0.064	$p = 0.00014$	$p = 0.00042$
Eastern	YQ	4.6	4.161	0.503	−0.035	$p = 0.00001$	$p = 0.00001$
	GJY	5.5	5.138	0.553	0.008	$p = 0.01070$	$p = 0.01932$
	LC	3.9	3.693	0.464	−0.032	$p = 0.00008$	$p = 0.00105$
	DX	5.6	5.25	0.617	0.014	$p = 0.00004$	$p = 0.00011$
	DY	4.6	4.625	0.572	0.098	$p = 0.00002$	$p = 0.00002$
	GTS	4.4	4.255	0.551	−0.033	$p = 0.00011$	$p = 0.00033$
	QDH	5.9	5.291	0.601	−0.009	$p = 0.00001$	$p = 0.00001$
	SYT	4.3	4.002	0.556	0.016	$p = 0.00004$	$p = 0.00053$
	Average	5.0	4.733	0.573	0.015	-	-

A: Average number of alleles, A_R: allele richness, H: genetic diversity, F_{IS}: inbreeding index. TPM: two-phased mutation model, and SMM: stepwise mutation model.

The optimal K-value was found to be 2 in STRUCTURE analysis, indicating that there were two genetic clusters in *C. sclerophylla*. Generally, population XJQ, LH, XN, YLS, YL, WM, and XS composed one cluster and other populations formed another (Figure 2), which suggest an eastern-to-western differentiation. According to geographical distribution, the former cluster could be further divided into a western group and a central group. Population XJQ, LH, and XN were located in the Xuefeng mountain range and made up the western group (W), YLS, YL, WM, and XS were situated in the Luoxiao mountain range and composed the central group (C), and the other eight populations were mainly seated in the Wuyi mountain range and formed the eastern group (E). Further genetic structuring

was observed in the eastern group when K = 3, in which GJY and LC were separate from the other populations. There was no significant difference in microsatellite diversity among the three groups, and the central group had highest microsatellite diversity (A_R = 5.086, H_O = 0.580, and H_S = 0.600). AMOVA analysis showed that genetic variation of *C. sclerophylla* mainly existed within populations (Table 5), and that there was significant differentiation among three groups.

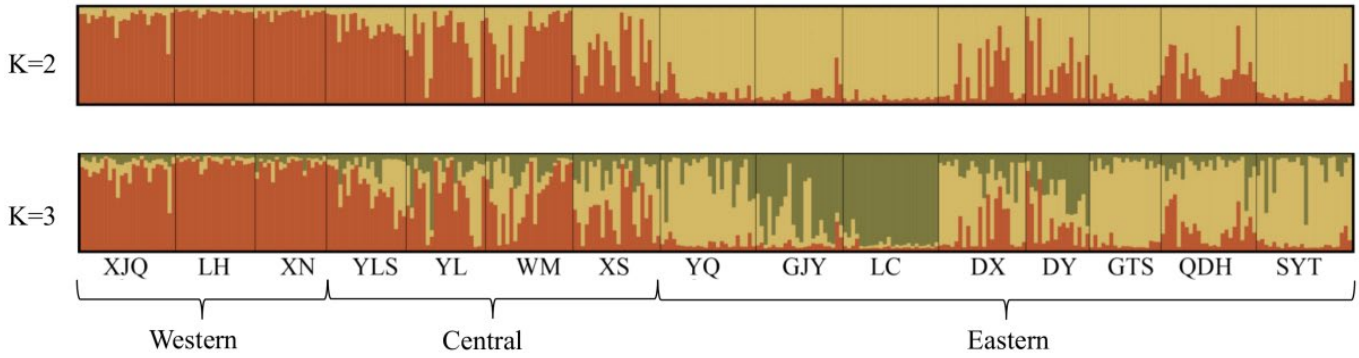


Figure 2. Bayesian clustering analysis of *Castanopsis sclerophylla* based on nuclear microsatellite markers.

Table 5. Analysis of molecular variance of *Castanopsis sclerophylla* based on nuclear SSRs.

Source of Variation	Degree of Freedom	Sum of Squares	Variance Components	Percentage Variation	*p*-Value
		As a whole			
Among populations	14	316.589	0.423	8.48	$p < 0.000$
Within populations	625	2854.255	4.567	91.52	
		Three groups			
Among groups	2	81.696	0.111	2.21	$p < 0.000$
Among populations Within groups	12	234.893	0.352	7.00	$p < 0.000$
Within populations	625	2854.255	4.567	90.79	$p < 0.000$

The posterior probabilities of eight hypothesized evolutionary scenarios for the western, central, and eastern groups of *Castanopsis sclerophylla* are shown in Table 6. Scenario 6 had the highest posterior probability of 0.3404 with 95% confidence interval (CI) from 0.3277 to 0.3531, which showed no overlap with the 95% CI of posterior probability of the other seven scenarios. Scenario 6 suggested that an eastern-to-western differentiation occurred in *C. sclerophylla* at about 5950 generations ago (Figure 3, Table 7), which was followed by genetic admixture at about 1460 generations ago due to secondary contact, as we could see that the effective population size of western and eastern groups increased at this time. By considering generation time of 25 years for *Castanopsis* species [39], the differentiation between the western and eastern group occurred at about 148.7 Ka (243.5–44.2 Ka), and the genetic admixture happened at about 36.5 Ka (94.5–3.1 Ka). The admixture rate from the western group (0.515) was slightly higher than that from the eastern group (0.485).

Table 6. Posterior probabilities of eight hypothesized evolutionary scenarios for the western, central, and eastern groups of *Castanopsis sclerophylla*.

Evolutionary Scenario	Posterior Probability (95% Confidence Interval)
1	0.0294 (0.0172–0.0416)
2	0.0119 (0.0000–0.0248)
3	0.0263 (0.0143–0.0384)
4	0.0170 (0.0042–0.0297)
5	0.1006 (0.0898–0.1114)
6	0.3404 (0.3277–0.3531)
7	0.2873 (0.2740–0.3007)
8	0.1872 (0.1703–0.2040)

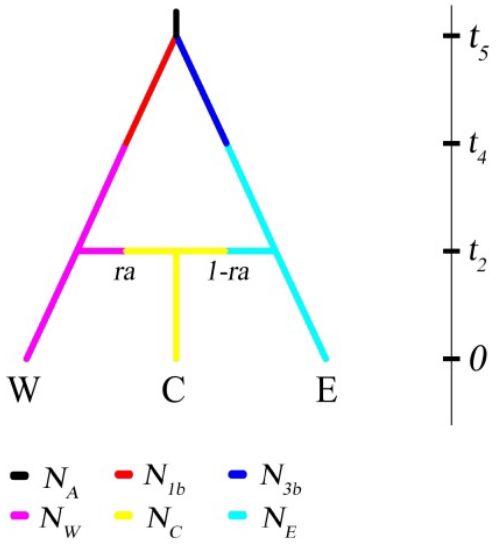


Figure 3. The evolutionary scenario with highest posterior probability for the western (W), central (C), and eastern (E) groups of *Castanopsis sclerophylla* based on scenario 6. The ancestral population with effective population size of N_A split into the western and eastern groups at time t_5 with effective population size of N_{1b} and N_{3b}, and their effective population sizes increased at time t_4 to N_W and N_E. Genetic admixture occurred at time t_2 with admixture rate of r_a, and the central group has the effective population size of N_C.

Table 7. Posterior distribution of the parameters in the hypothesized scenario 6 for the western, central, and eastern groups of *Castanopsis sclerophylla*.

Population Historical Parameters	Mean	Median	Q025	Q075
Effective population size				
N_A	42,200	37,900	8950	93,400
N_W	60,300	59,200	24,100	97,100
N_C	76,300	79,900	35,100	98,700
N_E	85,800	87,700	61,800	99,300
N_{1b}	22,100	19,700	3050	55,100
N_{3b}	25,100	22,600	3700	61,000
Time (generation)				
t_2	1460	1240	124	3780
t_4	4450	4260	958	9700
t_5	5950	5950	1770	9740
Genetic admixture rate				
r_a	0.515	0.520	0.117	0.898

N_A, N_W, N_C, and N_E: the effective population size of the ancestral, western, central, and eastern groups, respectively; N_{1b}, N_{2b}: the effective population sizes of western and eastern group at time of t_4; t_2, t_4, t_5: period of different evolutionary events; and r_a: genetic admixture rate.

The precipitation of coldest quarter (BIO19) and the max temperature of warmest month (BIO05) had the highest contributions to the niche modelling of *C. sclerophylla*. The area under the receiver operating characteristic curve was 0.923, and the predicted present distribution range was highly consistent with the actual distribution range, indicating that the ecological niche model had high reliability. *Castanopsis sclerophylla* mainly occurred in Guangdong and Guangxi provinces during LIG (Figure 4a), but it could be found in more northern regions such as some parts of Hunan and Jiangxi provinces at this time, and the high-suitable area showed some discontinuity in the eastern-to-western direction. Distribution of this species contracted further south during LGM (Figure 4b) and the high-suitable area became more limited and fragmented. *Castanopsis sclerophylla* expanded northward during MH (Figure 4c) and occupied the much larger area than that during LIM.

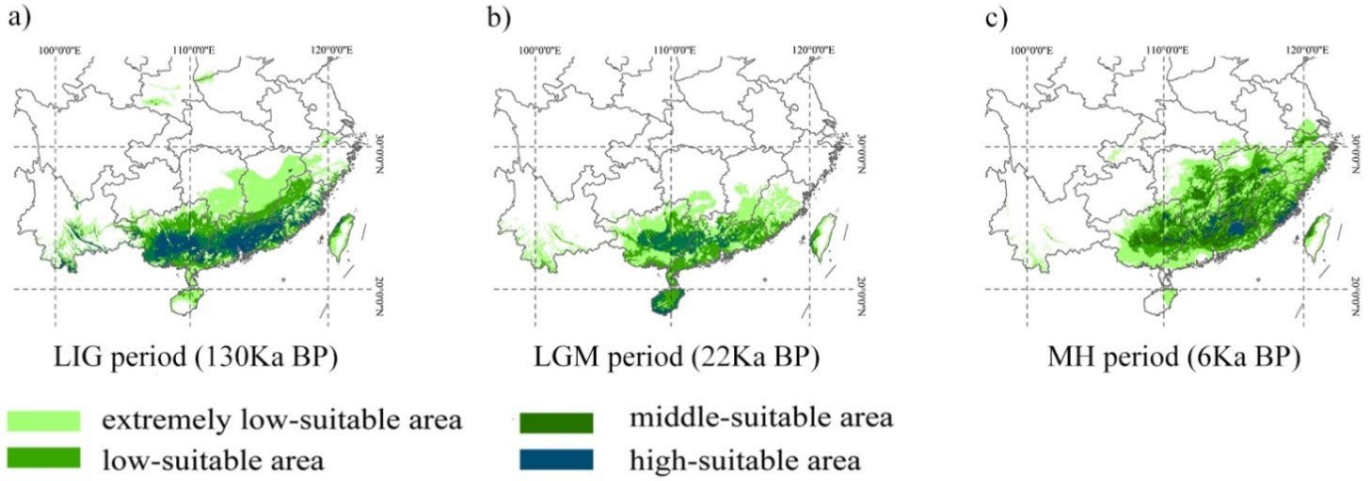


Figure 4. Potential distribution range of *Castanopsis sclerophylla* during the different periods, (**a**) the Last-interglacial (LIG, 130 Ka BP), (**b**) the Last Glacial Maximum (LGM, 22 Ka BP), (**c**) the Mid Holocene (MH, 6 Ka BP).

4. Discussion

Studying population structure has made an essential contribution to understanding evolutionary processes and demographic history in forest ecology research [40]. In this study, significant genetic differentiation among the western, central, and eastern groups of *C. sclerophylla* has been evidenced. Haplotype H2 was most likely to be the ancestral haplotype since it was interior on the network with multiple connections and had the highest frequencies [41]. Haplotype H3 and H1 also had high frequencies and were closely related to H2. Haplotypes H3, H2, and H1 mainly occurred in the western, central, and eastern populations, respectively, indicating population divergence arose along the Xuefeng, Luoxiao, and Wuyi mountain ranges. Bayesian clustering analysis based on nuclear SSRs suggested genetic eastern-to-western differentiation in *C. sclerophylla*, which was very similar to the population genetic structure of a widespread subtropical tree species *Castanopsis eyrei* [6]. DIYABC simulations demonstrated that the genetic diversity pattern of *C. sclerophylla* could be explained by geographic isolation, followed by secondary contact. The morphotectonic events in the middle-late Pleistocene, in particular the continuous uplift of the Xuefeng, Luoxiao, and Wuyi mountain ranges following the Himalayan orogeny [42,43], was supposed to promote the eastern-to-western differentiation in *C. sclerophylla*. The three southwest-to-northeast mountain ranges effectively reduced gene exchange in the longitude direction, and resulted in remarkable phylogeographic structure in *C. sclerophylla*. The climate became cold and dry in China from middle Pleistocene [42], and considerable climate changes would lead to population shrinkage and discontinuity, and intensify the eastern-to-western differentiation of *C. sclerophylla*. The BOTTLENECK analysis showed that all 15 populations of *C. sclerophylla* had experienced recent bottlenecks, which agreed

with the assumption that the species became discontinuous and separated into more isolated populations during middle Pleistocene.

DIYABC simulations suggested that genetic admixture between the western and eastern groups of *C. sclerophylla* was caused by secondary contact at about 36.5 Ka BP. The climate of China during 40–30ka B. *p*. was especially warm and humid, and the temperature was about 4 °C higher than that at the present and the precipitation was also higher than now [44,45]. Suitable climatic conditions facilitate *C. sclerophylla* to expand, and the enlarged distribution range provided an opportunity for the secondary contact of the west and east populations. Lying in the midst of a longitudinal depression between China's western highlands and the coastal ranges of Fujian province is the Chiang-nan hilly region, where lowlands serve as a putative dispersal corridor for many plant species, increasing the population connectivity among the Xuefeng, Luoxiao, and Wuyi mountain ranges and even the Zhejiang hilly region [46,47]. This explains why haplotype H3 was found in SYT and YQ.

The ecological niche modelling showed that the distribution of *C. sclerophylla* contracted southward during the LGM period (about 22 Ka BP), which would explain why the ancient haplotype H1 was found in the southern end of the Xuefeng, Luoxiao, and Wuyi mountain ranges. The southern limit of these three mountain ranges is the Nanling mountain range, which spans from west to east for more than 1000 km in subtropical China and plays dual roles as a glacial refugium and a dispersal corridor [48]. The distribution range expanded northward during the MH (about 6 Ka BP). Significant negative correlation between genetic diversity (A_R) and latitude was revealed in central populations (result not shown), which is consistent with the suggestion that the Luoxiao mountain range was one potential recolonization route for plants in subtropical China [5]. The flowering period of *C. sclerophylla* is from April to May, the southwest monsoon in summer and the southwest-to-northeast alignment of mountains, hills, and valleys alternating in Chiang-nan hilly region facilitated pollen dispersal northward. The central group of *C. sclerophylla* possessed higher genetic diversity than the western and eastern groups, and genetic admixture caused by secondary contact could be one reason for the higher genetic diversity in the central region. The Xuefeng, Luoxiao, and Wuyi mountain ranges should be considered as separate conservation units because they possessed distinct chloroplast haplotypes, and the Luoxiao mountain range should be considered as the key region for conservation of *C. sclerophylla* due to it containing the highest diversity of SSR alleles.

5. Conclusions

In this study, we revealed the genetic structure and demographic history of *C. sclerophylla*. During the Last-interglacial, eastern-to-western differentiation of *C. sclerophylla* was promoted by isolation due to the uplift of the Xuefeng, Luoxiao, and Wuyi mountain ranges, then genetic admixture occurred as a result of secondary contact between differentiated groups when the climate became warm and humid at about 36.5 Ka BP. Niche modelling and SSR allelic diversity pattern indicated post-glacial expansion of *C. sclerophylla* from the southern end of the Luoxiao mountain range. The Luoxiao mountain range possesses a higher level of genetic diversity than other regions and should be considered as a key conservation unit of *C. sclerophylla*. These results also provided important implications for understanding the dynamics of evergreen broad-leaved forest in subtropical China.

Author Contributions: Conceptualization, Y.S., X.Q. and Z.Z.; methodology, R.C., X.Z. and Y.S.; software, R.C.; validation, Y.S., S.C. and X.C.; formal analysis, R.C., X.Z., S.C. and X.C.; investigation, R.C., X.Z. and Y.S.; resources, Y.S., X.Q. and Z.Z.; data curation, R.C. and S.C.; writing—original draft preparation, S.C.; writing—review and editing, Y.S.; visualization, S.C.; supervision, Y.S.; project administration, Y.S.; and funding acquisition, Y.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the National Natural Science Foundation of China (No. 31770698) and the Special Fund for Talents of South China Agricultural University.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

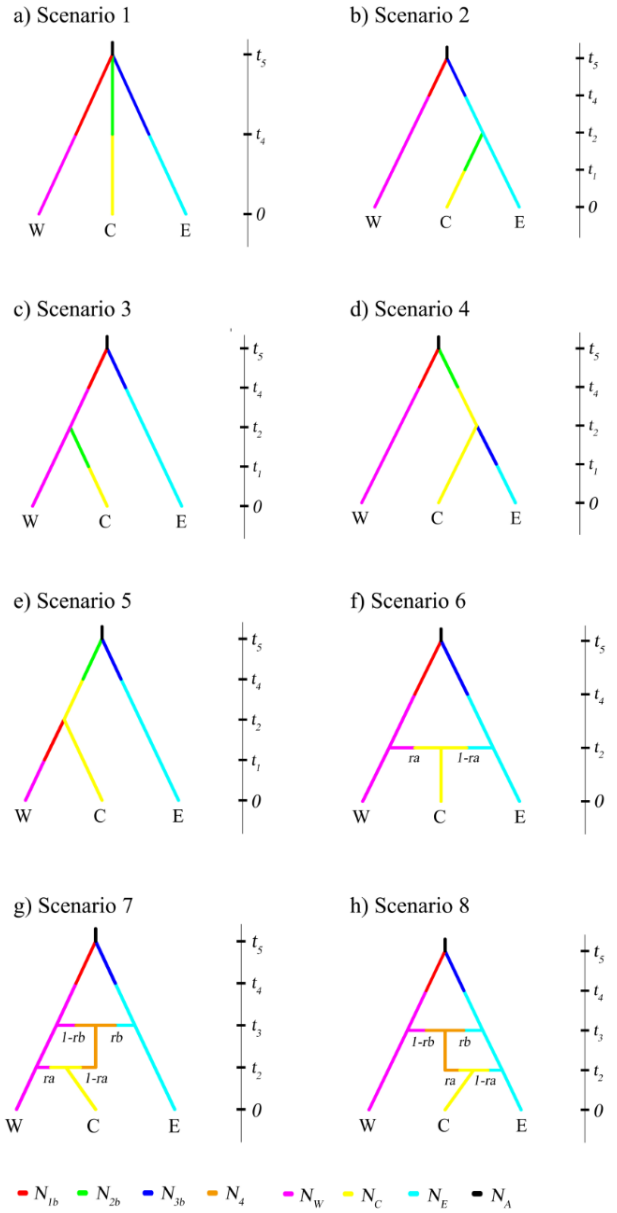


Figure A1. Eight hypothesized evolutionary scenarios for the western, central, and eastern groups of *Castanopsis sclerophylla*. N_A, N_W, N_C, N_E, and N_4: effective population size of ancestral population (A), the western (W), central (C), and eastern (E) group, and an unsampled population. N_{1b}, N_{2b}, and N_{3b}: size reduction of the western and eastern group. t_1, t_2, t_3, t_4, and t_5: times of different evolutionary events. r_a, r_b: admixture rate at different times. (**a**) The ancestral population splits into three groups, and their effective population sizes increase later; (**b**) The ancestral population splits into the western and eastern groups, then the central group diverges from the eastern group; (**c**) The ancestral population splits into the western and eastern groups, then the central group diverges from the western group; (**d**) The ancestral population splits into the western and central groups, then the eastern group diverges from the central group; (**e**) The ancestral population splits into the central and eastern groups, then the western group diverges from the central group; (**f**) The ancestral population splits into the western and eastern groups, the central group is derived from an admixture event; (**g**) The ancestral population splits into the western and eastern groups, the central group is derived from recurrent admixture with the western group; (**h**) The ancestral population firstly splits into the western and eastern groups, the central group is derived from recurrent admixture with the eastern group.

References

1. Wu, Z.Y. *Vegetation of China*; Science Press: Beijing, China, 1980.
2. Fan, D.; Hu, W.; Li, B.; Morris, A.B.; Zheng, M.; Soltis, D.E.; Soltis, P.S.; Zhang, Z. Idiosyncratic responses of evergreen broad-leaved forest constituents in China to the Late Quaternary climate changes. *Sci. Rep.* **2016**, *6*, 31044. [CrossRef] [PubMed]
3. Zheng, Z.; Yuan, B.Y.; Petit-Maire, N. Palaeoenvironments in China during the Last Glacial Maximum and the Holocene Optimum. *Episodes* **1998**, *21*, 152–158.
4. Ni, J.; Yu, G.; Harrison, S.P.; Prentice, I.C. Palaeovegetation in China during the late Quaternary: Biome reconstructions based on a global scheme of plant functional types. *Palaeogeogr. Palaeocl.* **2010**, *289*, 44–61. [CrossRef]
5. Tian, S.; Lei, S.Q.; Hu, W.; Deng, L.L.; Li, B.; Meng, Q.L.; Soltis, D.E.; Soltis, P.S.; Fan, D.M.; Zhang, Z.Y. Repeated range expansions and inter-/postglacial recolonization routes of *Sargentodoxa cuneate* (Oliv.) Rehd. et Wils. (Lardizabalaceae) in subtropical China revealed by chloroplast phylogeography. *Mol. Phylogenetics Evol.* **2015**, *85*, 238–246. [CrossRef]
6. Shi, M.M.; Michalski, S.G.; Welk, E.; Chen, X.Y.; Durka, W. Phylogeography of a widespread Asian subtropical tree: Genetic east-west differentiation and climate envelope modelling suggest multiple glacial refugia. *J. Biogeogr.* **2014**, *41*, 1710–1720. [CrossRef]
7. Ye, J.W.; Li, D.Z.; Hampe, A. Differential Quaternary dynamics of evergreen broadleaved forests in subtropical China revealed by phylogeography of *Lindera aggregata* (Lauraceae). *J. Biogeogr.* **2019**, *46*, 1112–1123. [CrossRef]
8. Xu, J.; Deng, M.; Jiang, X.L.; Westwood, M.; Song, Y.G.; Turkington, R. Phylogeography of *Quercus glauca* (Fagaceae), a dominant tree of East Asian subtropical evergreen forests, based on three chloroplast DNA interspace sequences. *Tree Genet. Genomes* **2014**, *11*, 805. [CrossRef]
9. Lei, M.; Wang, Q.; Wu, Z.J.; Lopez-Pujol, J.; Li, D.Z.; Zhang, Z.Y. Molecular phylogeography of *Fagus engleriana* (Fagaceae) in subtropical China: Limited admixture among multiple refugia. *Tree Genet. Genomes* **2012**, *8*, 1203–1212. [CrossRef]
10. Ding, Y.; Wan, J.; Liu, C.; Shi, X.; Xia, X.; Prakash, S.; Zhang, X. Retrogradation properties and in vitro digestibility of wild starch from *Castanopsis sclerophylla*. *Food Hydrocoll.* **2020**, *103*, 105693. [CrossRef]
11. Young, A.; Boyle, T.; Brown, T. The population genetic consequences of habitat fragmentation for plants. *Trends Ecol. Evol.* **1996**, *11*, 413–418. [CrossRef]
12. Wang, R.; Compton, S.G.; Chen, X.Y. Fragmentation can increase spatial genetic structure without decreasing pollen-mediated gene flow in a wind-pollinated tree. *Mol. Ecol.* **2011**, *20*, 4421–4432. [CrossRef]
13. Shi, M.M.; Chen, X.Y. Leading-edge populations do not show low genetic diversity or high differentiation in a wind-pollinated tree. *Popul. Ecol.* **2012**, *54*, 591–600. [CrossRef]
14. Wang, R.; Compton, S.G.; Shi, Y.S.; Chen, X.Y. Fragmentation reduces regional-scale spatial genetic structure in a wind-pollinated tree because genetic barriers are removed. *Ecol. Evol.* **2012**, *2*, 2250–2261. [CrossRef]
15. Gupta, P.K.; Balyan, H.S.; Sharma, P.C.; Ramesh, B. Microsatellites in plants: A new class of molecular markers. *Curr. Sci.* **1996**, *70*, 45–54.
16. Ye, J.; Zhang, Y.; Wang, X. Phylogeographic history of broad-leaved forest plants in subtropical China. *Acta Ecol. Sin.* **2017**, *17*, 5894–5904.
17. Wolfe, K.H.; Li, W.H.; Sharp, P.M. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast and nuclear DNAs. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 9054–9058. [CrossRef]
18. Garrick, R.C.; Bonatelli, I.A.S.; Hyseni, C.; Morales, A.; Pelletier, T.A.; Perez, M.F.; Rice, E.; Satler, J.D.; Symula, R.E.; Thome, M.T.C. The evolution of phylogeographic data sets. *Mol. Ecol.* **2015**, *24*, 1164–1171. [CrossRef]
19. Kress, W.J.; Wurdack, K.J.; Zimmer, E.A.; Weigt, L.A.; Janzen, D.H. Use of DNA barcodes to identify flowering plants. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 8369–8374. [CrossRef]
20. Cheng, Y.P.; Hwang, S.Y.; Lin, T.P. Potential refugia in Taiwan revealed by the phylogeographic study of *Castanopsis carlesii* Hayata (Fagaceae). *Mol. Ecol.* **2005**, *14*, 2075–2085. [CrossRef]
21. Shi, Y.S.; Zhang, J.; Jiang, K.; Cui, M.Y.; Li, Y.Y. Development and characterization of polymorphic microsatellite markers in *Castanopsis sclerophylla* (Fagaceae). *Am. J. Bot.* **2011**, *98*, e19–e21. [CrossRef]
22. Li, C.; Sun, Y. Transferability analysis of EST-SSR markers of *Castanea mollissima* to *Castanopsis fargesii*. *Guihaia* **2012**, *32*, 293–297. (In Chinese)
23. Alzohairy, A.M. BioEdit: An important software for molecular biology. *GREF Bull. Biosci.* **2011**, *2*, 60–61.
24. Kumar, S.; Tamura, K.; Nei, M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief. Bioinform.* **2004**, *5*, 153–163. [CrossRef] [PubMed]
25. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **2009**, *25*, 1451–1452. [CrossRef]
26. Pons, Q.; Petit, R.J. Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics* **1996**, *144*, 1237–1245. [CrossRef]
27. Clement, M.J.; Posada, D.; Crandall, K.A. TCS: A computer program to estimate gene genealogies. *Mol. Ecol.* **2000**, *9*, 1657–1659. [CrossRef]
28. Goudet, J. *FSTAT, a Program to Estimate and Test Gene Diversities and Fixation Indices*; Version 2.9.3; Lausanne University: Lausanne, Switzerland, 2001.

29. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959. [CrossRef]
30. Earl, D.A.; Vonholdt, B.M. Structure Harvester: A website and program for visualizing Structure output and implementing the Evanno method. *Conserv. Genet. Resour.* **2012**, *4*, 359–361. [CrossRef]
31. Jakobsson, M.; Rosenberg, N.A. Clumpp: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **2007**, *23*, 1801–1806. [CrossRef]
32. Excoffier, L.; Laval, G.; Schneider, S. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evol. Bioinform. Online* **2005**, *1*, 47–50. [CrossRef]
33. Piry, S.; Luikart, G.; Cornuet, J.M. Bottleneck: A computer program for detecting recent reductions in the effective population size using allele frequency data. *J. Hered.* **1999**, *90*, 502–503. [CrossRef]
34. Cornuet, J.M.; Pudlo, P.; Veyssier, J.; Dehne-Garcia, A.; Gautier, M.; Leblois, R.; Marin, J.M.; Estoup, A. DIYABC v2.0: A software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics* **2014**, *30*, 1187–1189. [CrossRef]
35. Hijmans, R.J.; Cameron, S.E.; Parra, J.L.; Jones, P.G.; Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **2010**, *25*, 1965–1978. [CrossRef]
36. Phillips, S.J.; Anderson, R.P.; Schapire, R.E. Maximum entropy modeling of species geographic distributions. *Ecol. Model.* **2006**, *190*, 231–259. [CrossRef]
37. Swets, J.A. Measuring the accuracy of diagnostic systems. *Science* **1988**, *240*, 1285–1293. [CrossRef]
38. Creley, C.M.; Muchlinski, A.E.; Shilling, F.M. An ecological niche model to predict range expansion of the eastern gray squirrel in California. *Bull. South. Calif. Acad. Sci.* **2019**, *118*, 58–70. [CrossRef]
39. Liu, R.L.; Wang, L.; Du, T.Z. Features of population ecological quantity field of *Castanopsis fabri* community in Jinggang Mountains. *Sci. Silvae Sin.* **2008**, *44*, 1–7. (In Chinese)
40. Maldonado, C.; Mora-Poblete, F.; Echeverria, C.; Baettig, R.; Torres-Díaz, C.; Contreras-Soto, R.I.; Heidari, P.; Lobos, G.A.; do Amaral Júnior, A.T. A Neural Network-Based Spectral Approach for the Assignment of Individual Trees to Genetically Differentiated Subpopulations. *Remote Sens.* **2022**, *14*, 2898. [CrossRef]
41. Freeland, J.R. *Molecular Ecology*; John Wiley & Sons Ltd.: Chichester, UK, 2005.
42. Yang, H.J.; Xin, X.; Li, G.S. The causal mechanism of Quaternary environmental changes in China. *Quat. Sci.* **1989**, *2*, 97–111. (In Chinese)
43. Luo, D.; Xiao, Y.F.; Zhan, L.; Zhang, L. The tectogenesis control of Danxia landform in southeast China. *J. East China Inst. Technol.* **2010**, *33*, 147–153. (In Chinese)
44. Shi, Y.F.; Yu, G. Warm-humid climate and transgressions during 40~30 ka B.P. and their potential mechanisms. *Quat. Sci.* **2003**, *23*, 1–11. (In Chinese)
45. Shi, Y.F.; Zhao, J.D. The special warm-humid climate and environment in China during 40~30 ka BP: Discovery and review. *J. Glaciol. Geocryol.* **2009**, *31*, 1–10. (In Chinese)
46. Fan, D.; Sun, Z.; Li, B.; Kou, Y.; Hodel, R.G.J.; Jin, Z.; Zhang, Z. Dispersal corridors for plant species in the Poyang Lake Basin of southeast China identified by integration of phylogeographic and geospatial data. *Ecol. Evol.* **2017**, *7*, 5140–5148. [CrossRef]
47. Li, Y.; Zhang, X.; Fang, Y. Landscape features and climatic forces shape the genetic structure and evolutionary history of an oak species (*Quercus chenii*) in east China. *Front. Plant Sci.* **2019**, *10*, 1060. [CrossRef]
48. Tian, S.; Kou, Y.; Zhang, Z.; Yuan, L.; Li, D.; Lopez-Pujol, J.; Fan, D.; Zhang, Z. Phylogeography of *Eomecon chionantha* in subtropical China: The dual roles of the Nanling Mountains as a glacial refugium and a dispersal corridor. *BMC Evol. Biol.* **2018**, *18*, 20. [CrossRef]