Article

Genetic Diversity and Population Structure of an Arctic Tertiary Relict Tree Endemic to China (*Sassafras tzumu*) Revealed by Novel Nuclear Microsatellite (nSSR) Markers

Shuang Wang [1], Ying Wang [1], Jingbo Zhou [1], Pan Li [2], Hungwei Lin [1], Ye Peng [3], Lipeng Yu [4], Yunyan Zhang [1,*] and Zhongsheng Wang [1,*]

1   College of Life Sciences, Nanjing University, Nanjing 210023, China
2   Systematic & Evolutionary Botany and Biodiversity Group, MOE Laboratory of Biosystem Homeostasis and Protection, College of Life Sciences, Zhejiang University, Hangzhou 310058, China
3   College of Biology and the Environment, Nanjing Forestry University, Nanjing 210037, China
4   Mount Longwang Nature Reserve, Huzhou 313300, China
*   Correspondence: zyynju@smail.nju.edu.cn (Y.Z.); wangzs@nju.edu.cn (Z.W.); Tel.: +86-15261868978 (Y.Z.); +86-13770650868 (Z.W.)

Abstract: *Sassafras tzumu* (Hemsl.) Hemsl., as an Arctic Tertiary relict woody species, is an ecologically and economically important deciduous tree endemic to southern China. Nonetheless, the genetic resources and backgrounds of *S. tzumu* are still lacking and remain largely unclear. Here, we predicted 16,215 candidate polymorphic nuclear microsatellite (nSSR) loci from the assembled nucleus databases of six geographic-distant individuals of *S. tzumu* via CandiSSR. Among these nSSRs, the di- (75.53%) and tri-nucleotide (19.75%) repeats were the most abundant, and 27 new polymorphic SSRs were developed and characterized in 136 individuals from six natural populations of *S. tzumu*. The majority of the above 27 SSRs (24 loci, 88.89%) presented moderate polymorphism (mean PIC = 0.356), and the transferability of these markers in other *Sassafras* species was high (85.19%). A moderately low level of genetic diversity and a high variation (F_{ST} = 0.286) of six wild populations of *S. tzumu* were illuminated by 16 selected polymorphic nSSRs, with the average expected heterozygosity (H_E) of 0.430 at the species level and H_E ranging from 0.195 to 0.387 at the population level. Meanwhile, a bottleneck effect was shown in two populations. Consistent with the results of the principal coordinate analysis (PCoA) and phylogenetic trees, structure analysis optimally divided these six *S. tzumu* populations into two clusters, and the further strong population subdivision appeared from K = 2 to K = 5, which corresponded to two evolutionarily significant units (ESUs). Moreover, the significant correlation between genetic and geographic distance was tested by the Mantel test (r = 0.742, p = 0.006), clarifying the effect about isolation by distance (IBD), which could be possibly explained by the low gene flow (N_m = 0.625), a relatively high degree of inbreeding (F_{IS} = 0.166), a relatively large distribution, and mountainous barriers. Above all, our research not only enlarged the useful genetic resources for future studies of population genetics, molecular breeding, and germplasm management of *S. tzumu* and its siblings but also contributed to proposing scientific conservation strategies and schemes for the better preservation of *S. tzumu* and other *Sassafras* (Lauraceae) species.

Keywords: *Sassafras tzumu*; Arctic Tertiary relict plant; nuclear microsatellite; genetic diversity; population structure; population differentiation; ESU

1. Introduction

Sassafras J. Presl (Lauraceae) contains three extant species with a disjunct distribution between East Asia and North America: *S. tzumu* (Hemsl.) Hemsl. is located in the Yangtze River basin and its southern areas in continental China; *S. randaiense* (Hayata) Rehder is merely distributed in Central Taiwan; whereas *S. albidum* (Nutt.) Nees is found in the east of North America [1,2]. Previous studies pointed out that *Sassafras* is monophyletic,

and the species of North America (*S. albidum*) is sister to the clade of its two East Asian counterparts [3–5]. Morphologically, the above three sibling species mainly differ in their floral structure: *S. randaiense* distinguishes from *S. tzumu* in the number of anther chambers (two-locular versus four-locular), whereas we can tell *S. albidum* apart from its two Asian congeners in its six staminodes in the female flower and lacking the fourth staminal whorl [3,6].

S. tzumu, the focal species of our present study, is an Arctic Tertiary relict plant endemic to China, growing in both broad-leaved evergreen and mixed deciduous forests of subtropical China at an altitude range of 100–1900 m [7–9]. *S. tzumu* has strong tolerance for various types of soil, such as acid or sandy loam soil [7,8,10]. Additionally, *S. tzumu* has many important values in economy and ecology: its meticulous and durable timber is widely used for ship-building and manufacturing of fine furniture; its root and bark are used as medicine, with functions of activating blood circulation and dispersing blood stasis and treating muscle strain; its fruit, leaves, and root contain aromatic oil and safrole oleoresin; and it is also an ornamental landscape garden tree because of its peculiar shape and colorful leaves and the beautiful tree form [7,8,10,11]. In 2017, *S. tzumu* was listed in the Reference List of Major Cultivated Precious Tree Species in China by State Forestry Administration [12]. Meanwhile, some previous studies reported that *S. tzumu* is at risk due to their declining population size that resulted from overexploitation and other threats [7,8]. Therefore, it is crucial to conserve this precious species via enlarging its genetic resources and unraveling its genetic backgrounds of genetic diversity and population structure.

In recent years, with the rapid advancement of sequencing technology, developing molecular markers for the genetic survey of our interested species has become much easier and more effective [13]. Simple sequence repeat (SSR), also known as microsatellite, is a short repetitive DNA sequence consisting of a core motif (typically 2–6 bp) and a flanking sequence together [14]. SSR can be commonly spotted in both coding and noncoding regions of eukaryotic genomes [14], and it is widely used in the studies of population genetics, breeding programs, and phylogeography on account of its codominance, the abundance of genotypic polymorphism, high versatility, and variation detection rate [15–19]. Previous studies of *S. tzumu* primarily focused on germplasm investigation, taxonomy, cultivation, and population ecology [8,9,20–25], whereas a few studies have assessed the genetic diversity and population structure of this species [7,26–28]. For details, Ding et al. [26] firstly clarified the low-level genetic diversity (H_E = 0.280) of three natural populations of *S. tzumu* in Hubei Province [29] using an isozyme marker; Jiang et al. [27] found a high-level genetic diversity (H_E = 0.84) and a moderate genetic differentiation among five wild populations in Tianmu Mountain with five different altitudes via employing 13 polymorphic SSRs developed from *S. randaiense* [7] and *Cinnamomum camphora* [30]; Liu et al. [28] recently identified 11,862 SNP loci in 13 relatively agminated populations of *S. tzumu* and observed a relatively low-level genetic diversity (H_E = 0.3004) and a relatively high degree of population differentiation (F_{ST} = 0.118). The level of genetic diversity, status of population differentiation, and population structure on a relatively large scale were still controversial in *S. tzumu* in these investigations. Obviously, there is still a lack of some alternative molecular markers such as the nSSR marker directly mined from the genomic sequencing data of *S. tzumu* to augment its genetic resources and ascertain its genetic backgrounds.

The procedures of SSR development for a target species in a traditional way mainly consist of three steps: SSR discovery, primer design, and polymorphic survey in representative population or individuals. It is costly and extremely labor-intensive, time-consuming in the traditional experimental screening of the SSR polymorphic status and its subsequent applicability to genetic studies. Compared with the conventional methods or tools, CandiSSR is an easy-to-use, economic, and labor- and time-saving pipeline, which can friendly enable users to find putative polymorphic SSRs not only from the transcriptome datasets but also from multiple assembled genome sequences of a target species or genus along with several comprehensive assessments to help researchers focus more on subsequent genetic studies [19]. Hence, following Zhang et al.'s effective method of developing novel

SSR markers for the interested species [31], we firstly selected six individuals from six geographically distant populations of *S. tzumu* for low-coverage whole-genome short-gun sequencing in our present study to directly develop polymorphic SSR markers. We further investigated the genetic diversity and population structure of these six wild populations of *S. tzumu* via the polymorphic nSSR loci screened and developed from the nuclear genome data of *S. tzumu*. The results here will provide useful genetic resources for further studies of population genetics, molecular breeding, and germplasm management of *S. tzumu* and its congeners, as well as lay a solid basis for proposing scientific conservation strategies and schemes for their better preservation.

2. Results

2.1. Prediction, Verification, and Transferability of Newly Developed nSSR Loci of S. tzumu

A total of 16,215 potential polymorphic nSSR loci with 424 repeated types were identified by the effective pipeline CandiSSR. The most abundant motif was dinucleotide repeat (12,248; 75.53%) followed by trinucleotide repeat (3203; 19.75%), and the tetranucleotide, pentanucleotide, and hexanucleotide repeats accounted for 2.61%, 0.91%, and 1.19%, respectively (Figure 1). According to Zhang et al.'s principles of primer screening [31], we selected 48 candidate polymorphic nSSRs to test their amplification success and effectiveness using six individuals of *S. tzumu* and then detected their polymorphisms via 12 individuals from six representative geographically distant populations (two individuals were randomly chosen from each population). As a result, 35 nSSR loci (73.92%) were successful in amplification, among which 27 pairs of primers (56.25%) were polymorphic. The sequences of these above 27 polymorphic nSSRs were submitted to GenBank and annotated against the GenBank nonredundant database using BLASTX [32]. The detailed characteristics of these nSSRs are shown in Table 1 and Table S1.
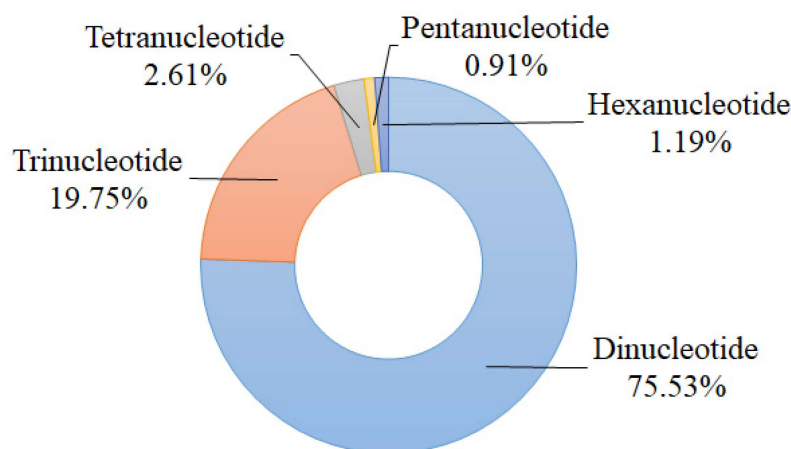
Figure 1. Statistics of types of candidate polymorphic nSSRs of *S. tzumu*.

For the assessment of cross-species transferability of these 27 newly developed nSSR loci in other two *Sassafras* species (*S. albidum* and *S. randaiense*), we found that 23 nSSR loci (85.19%) were successfully amplified in *S. albidum* and *S. randaiense*. Two nSSRs (N362 and N2716) failed to yield positive amplicon, and the other two nSSRs (N2689 and N7047) did not show clear readable allelic bands in these two species. Furthermore, 17 nSSRs (62.96%) presented the polymorphisms in 22 individuals of *S. albidum*, and 14 nSSRs (51.85%) successfully presented the polymorphisms (various genotypes in different individuals) in 10 individuals of *S. randaiense* (Tables S1 and S2).

Table 1. Characteristics of 27 newly developed nSSR loci of *S. tzumu* in this study.

Locus	Repeat Motif	Optimal Primer Temperature (°C)	Size Range (bp)	GenBank Accession Number	BLASTX Top Hit Description
N43	$(AAAG)_8$	59	157–169	OP094707	zinc finger protein [Cinnamomum micranthum f. kanehirae]
N168	$(AAC)_7$	56	134–146	OP094689	-
N362	$(AAG)_7$	55	185–206	OP094695	-
N711	$(AATAAA)_5$	59	187–223	OP094710	putative nuclease HARBI1 [Cinnamomum micranthum f. kanehirae]
N1037	$(ACAA)_5$	55	139–155	OP094684	hypothetical protein SLEP1_g25016 [Shorea leprosula]
N1047	$(ACAT)_5$	56	162–166	OP094703	-
N2689	$(AGA)_8$	52	188–197	OP094687	Alcohol dehydrogenase superfamily, zinc-type [Cinnamomum micranthum f. kanehirae]
N2716	$(AGAT)_5$	59	159–175	OP094697	-
N2799	$(AGGTGA)_5$	52	159–183	OP094690	-
N6336	$(ATC)_6$	59	178–187	OP094701	-
N6339	$(ATC)_6$	56	169–172	OP094694	-
N6493	$(ATGT)_6$	59	128–148	OP094692	-
N6882	$(CAC)_8$	55	167–188	OP094708	-
N6889	$(CACCTC)_5$	56	159–189	OP094705	-
N7047	$(CCCTCA)_5$	59	182–206	OP094702	-
N8706	$(CTC)_6$	55	114–129	OP094688	-
N8723	$(CTC)_9$	51	156–177	OP094685	zinc finger protein 1-like protein [Cinnamomum micranthum f. kanehirae]
N8865	$(CTT)_8$	55	191–203	OP094699	methyltransferase-like protein 13 isoform X1 [Cinnamomum micranthum f. kanehirae]
N10640	$(GAT)_6$	53	181–187	OP094691	N-carbamoylputrescine amidase [Glycine max]
N13148	$(TATAC)_5$	56	162–177	OP094704	-
N13155	$(TATG)_5$	58	179–191	OP094686	-
N13174	$(TATTTT)_6$	56	178–196	OP094706	-
N14929	$(TCT)_7$	57	187–202	OP094698	-
N15402	$(TGA)_6$	57	178–199	OP094709	-
N15940	$(TTC)_5$	55	82–100	OP094700	-
N16055	$(TTCT)_5$	55	161–169	OP094693	-
N16204	$(TTTTC)_5$	56	133–148	OP094696	-

2.2. Genetic Diversity of nSSRs of *S. tzumu*

As a result, a total of 125 reproducible alleles were screened in 27 nSSR primers set by 136 individuals of *S. tzumu* from six populations (Table S3). For each locus, the average of the number of alleles (N_A), effective number of alleles (N_E), observed heterozygosity (H_O), and expected heterozygosity (H_E) was 4.630, 1.750, 0.333, and 0.408, respectively. The polymorphism information content (*PIC*) of these loci varied from 0.071 (N1037) to 0.567 (N15402), and the average of *PIC*, H_S, H_T, and *I* was 0.356, 0.325, 0.413, and 0.742, respectively (Table 2). The proportion of nSSRs with moderate (0.25 < *PIC* < 0.5) or high polymorphism (*PIC* > 0.5) was 88.89%. These findings indicated that the majority of our developed nSSRs showed relatively high polymorphism in *S. tzumu*. The average inbreeding coefficient (F_{IS}) was 0.209, and the F_{IS} in 21 loci were all greater than zero (Table 2).

Table 2. Genetic characteristics of 27 polymorphic nSSRs in *S. tzumu*.

Locus	N_A [1]	N_E [2]	H_O [3]	H_E [4]	H_S [5]	H_T [6]	PIC[7]	I[8]	F_{IS} [9]
N43	4	1.858	0.537	0.464	0.446	0.464	0.384	0.759	−0.159
N168	4	1.950	0.412	0.489	0.406	0.487	0.402	0.792	0.158
N362	7	1.758	0.228	0.433	0.222	0.453	0.400	0.878	0.474
N711	7	1.447	0.250	0.310	0.288	0.318	0.291	0.676	0.194
N1037	5	1.077	0.051	0.072	0.064	0.067	0.071	0.201	0.285
N1047	2	1.826	0.676	0.454	0.420	0.452	0.350	0.645	−0.493
N2689	4	1.755	0.397	0.432	0.375	0.440	0.393	0.795	0.081
N2716	5	1.593	0.088	0.374	0.142	0.403	0.328	0.689	0.765
N2799	5	1.996	0.426	0.501	0.308	0.503	0.418	0.853	0.149
N6336	4	1.322	0.066	0.244	0.228	0.252	0.228	0.500	0.730
N6339	2	1.534	0.257	0.349	0.287	0.358	0.287	0.532	0.264
N6493	6	2.199	0.728	0.547	0.526	0.550	0.448	0.929	−0.332
N6882	6	1.797	0.265	0.445	0.381	0.440	0.426	0.985	0.406
N6889	5	2.189	0.551	0.545	0.438	0.538	0.440	0.891	−0.012
N7047	4	1.494	0.287	0.332	0.297	0.329	0.303	0.635	0.136
N8706	5	2.273	0.507	0.562	0.519	0.561	0.511	1.067	0.098
N8723	7	1.655	0.279	0.397	0.275	0.417	0.379	0.893	0.297
N8865	4	1.676	0.419	0.405	0.327	0.393	0.359	0.731	−0.035
N10640	3	2.301	0.199	0.567	0.223	0.568	0.469	0.905	0.651
N13148	5	1.547	0.331	0.355	0.340	0.372	0.327	0.692	0.068
N13155	3	1.956	0.449	0.491	0.349	0.493	0.373	0.700	0.086
N13174	4	1.533	0.265	0.349	0.316	0.367	0.320	0.682	0.242
N14929	4	1.859	0.434	0.464	0.358	0.436	0.365	0.703	0.065
N15402	7	2.532	0.353	0.607	0.470	0.613	0.567	1.246	0.420
N15940	6	1.391	0.294	0.282	0.255	0.310	0.266	0.611	−0.043
N16055	3	1.355	0.007	0.263	0.241	0.283	0.237	0.487	0.972
N16204	4	1.391	0.235	0.282	0.274	0.296	0.258	0.554	0.166
T-Mean [10]	4.630	1.750	0.333	0.408	0.325	0.413	0.356	0.742	0.209
E-Mean [11]	4.563	1.805	0.359	0.430	0.324	0.436	0.370	0.764	0.166

[1] Number of alleles; [2] Effective number of alleles; [3] Observed heterozygosity; [4] Expected heterozygosity; [5] Average genetic diversity within populations; [6] Total genetic diversity; [7] Polymorphism information content; [8] Shannon's information index; [9] inbreeding coefficients; [10] Mean of total loci; [11] Mean of loci after excluding the loci with excessive null allele frequencies or significant linkage disequilibrium.

For analyzing the genetic diversity of six *S. tzumu* populations, 11 loci (N43, N168, N711, N1037, N2689, N2799, N6336, N6882, N8706, N13148, and N16055) were excluded from subsequent calculation because of the excessive null allele frequency (greater than 0.08) [33] or significant linkage disequilibrium (less than 0.05) [34] according to the estimated results we obtained (Tables S4–S7). The averages of H_O, H_E, H_S, H_T, PIC, I, and F_{IS} were changed to 0.359, 0.430, 0.324, 0.436, 0.370, 0.764, and 0.166 from 0.333, 0.408, 0.325, 0.413, 0.356, 0.742, and 0.209, respectively (Table 2). Overall, no significant changes were shown.

At the population level, the expected heterozygosity (H_E) ranged from 0.195 (LL) to 0.387 (BYS), and the observed heterozygosity (H_O) ranged from 0.244 (LL) to 0.464 (JGS). The highest values of allele richness (A_R), number of effective allele (A_E), average number of allele (A), number of private allele (A_P), Shannon's information index (I), and expected heterozygosity (H_E) were all detected in the BYS population (2.891, 1.800, 3.125, 7, 0.674, and 0.387), and the lowest were in the LL population (2.016, 1.382, 2.125, 1, 0.341, and 0.195) (Table 3). Moreover, a genetic bottleneck effect was observed in population YX and JGS (Table S8).

2.3. Population Structure and Differentiation in S. tzumu

High genetic differentiation (F_{ST} = 0.286) existed among six populations in *S. tzumu* [35]. The results of the AMOVA revealed that the main variation of *S. tzumu* was within the populations (71.60%), and 28.40% of the total variation was attributed to differences among

the six populations (Table 4). The pairwise population genetic differentiation index (F_{ST}) and gene flow (N_m) among populations were also estimated in Table 5, from which we can find that all the F_{ST} values were significantly greater than 0.25 between population LL and the other five populations with the range of 0.356 (BYS) to 0.536 (JGS); and the values of N_m were all less than 1 among LL and the other five populations. The results of the Mantel test of the IBD effect revealed a significant correlation ($r = 0.742$, $p = 0.006$) between the geographical and genetic distance among the six populations (Figure S1, Table S9).

Table 3. Genetic parameters of six populations in *S. tzumu*.

Population Code	N [1]	A [2]	A_E [3]	A_R [4]	A_P [5]	H_O [6]	H_E [7]	μH_E [8]	I [9]
BYS	21	3.125	1.800	2.891	7	0.366	0.387	0.396	0.674
LL	20	2.125	1.382	2.016	1	0.244	0.195	0.200	0.341
HNS	22	2.750	1.643	2.575	5	0.318	0.355	0.363	0.599
YX	26	2.438	1.648	2.258	1	0.392	0.322	0.329	0.526
JZF	33	2.563	1.555	2.213	3	0.381	0.309	0.313	0.502
JGS	14	2.125	1.595	2.125	2	0.464	0.334	0.346	0.518
Mean/Total	136	2.521	1.604	2.346	19	0.361	0.317	0.325	0.527

[1] Number of individuals; [2] Average number of alleles; [3] Effective number of alleles; [4] Allelic richness (average of per locus in each population) based on minimum sample size of population with 14 diploid individuals; [5] Number of private alleles; [6] Observed heterozygosity; [7] Expected heterozygosity; [8] Unbiased expected heterozygosity; [9] Shannon's information index.

Table 4. Result of AMOVA for *S. tzumu*.

Source of Variation	d.f.	Sum of Squares	Variance Components	Percentage of Variation
Among populations	5	242.774	1.027 Va	28.40
Within populations	266	689.178	2.591 Vb	71.60
Total	271	931.952	3.618	

Table 5. Pairwise population genetic differentiation index (F_{ST}) (above diagonal) with all significant population differentiation where $p < 0.01$ and number of migrants per generation (N_m) (below diagonal) among populations of *S. tzumu*.

	BYS	LL	HNS	YX	JZF	JGS
BYS	-	0.356	0.147	0.234	0.184	0.205
LL	0.452	-	0.443	0.530	0.514	0.536
HNS	1.448	0.314	-	0.224	0.185	0.182
YX	0.820	0.221	0.868	-	0.127	0.162
JZF	1.109	0.236	1.102	1.712	-	0.063
JGS	0.970	0.216	1.120	1.292	3.740	-
Total F_{ST}	0.286					
Total N_m	0.625					

STRUCTURE yielded the highest likelihood when individuals were grouped into two clusters ($K = 2$) with the maximum ΔK (Figure 2). However, the maximum of ln$p$ (D) was exhibited with $K = 5$ (Figure 2), and the further strong population subdivision appeared from $K = 2$ to $K = 5$ (Figure 2). LL, the westernmost population, was assigned to one cluster (the red color in Figure 2), and the other five populations were assigned to another cluster (the green color in Figure 2) when running STRUCTURE with $K = 2$, although the population BYS appeared mixed into two clusters. When $K = 3$, populations BYS and HNS (the southernmost population) were formed as a separate cluster (the blue color in Figure 2). With $K = 4$, populations JZF and JGS (the northernmost population) were assigned to a separate cluster (the yellow color in Figure 2), whereas population YX (the easternmost population) was separated and showed some degree of mixed ancestry (the green color in

Figure 2). When *K* = 5, the population BYS separated into another cluster (the purple color in Figure 2).
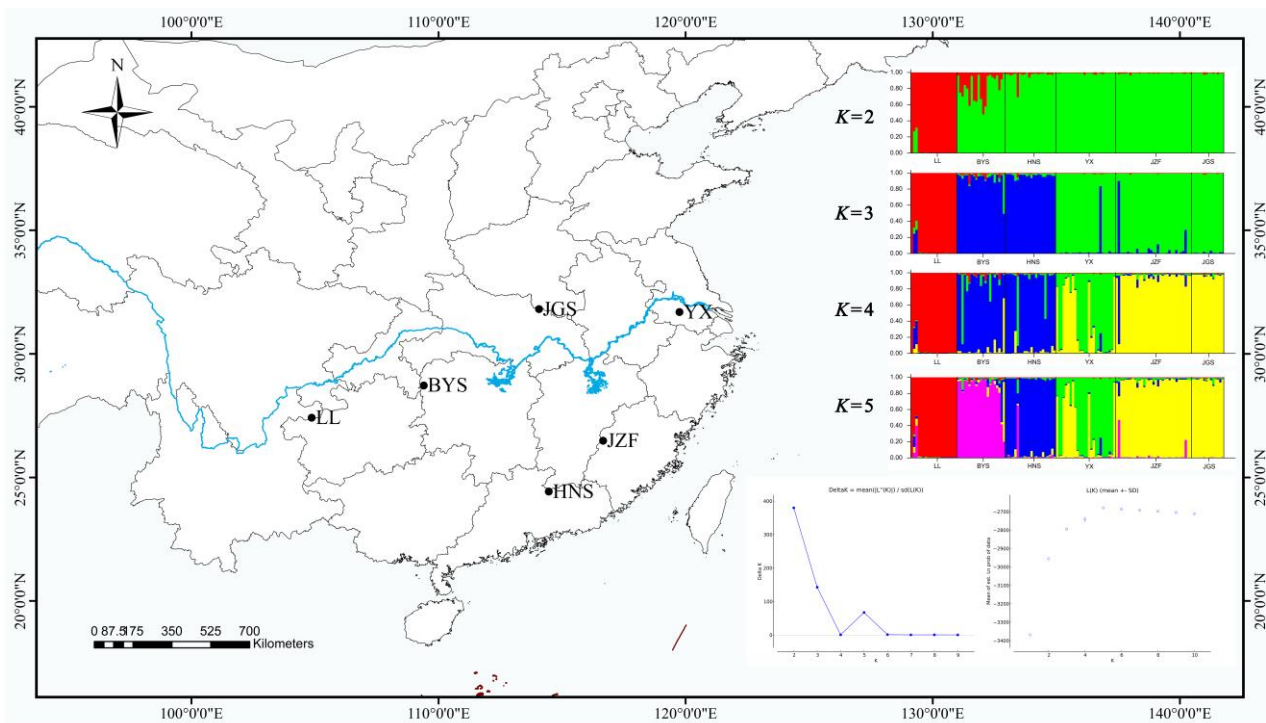
Figure 2. Geographic distribution of six populations of *S. tzumu* in China and population genetic structure among six populations of *S. tzumu* estimated using STRUCTURE based on nSSR with values of *K* from 2 to 5.

Consistent with the result of STRUCTURE when *K* = 2, the topology of UPGMA and NJ trees all indicated that the six populations of *S. tzumu* could be grouped into two clades: population LL and the other remaining five populations (BYS, HNS, YX, JZF, and JGS) (Figure 3). Moreover, the first two components of the principal coordinate analysis (PCoA) were generated based on the mean genotypic genetic distance among the six populations of *S. tzumu*, accounting for 73.31% (Axis 1) and 11.58% (Axis 2) of the total variance. It presented that the JZF and JGS populations had the most similar value, and the other four populations formed their own positions (Figure 4). The above results also accorded with the results of phylogenic and genetic cluster analysis.
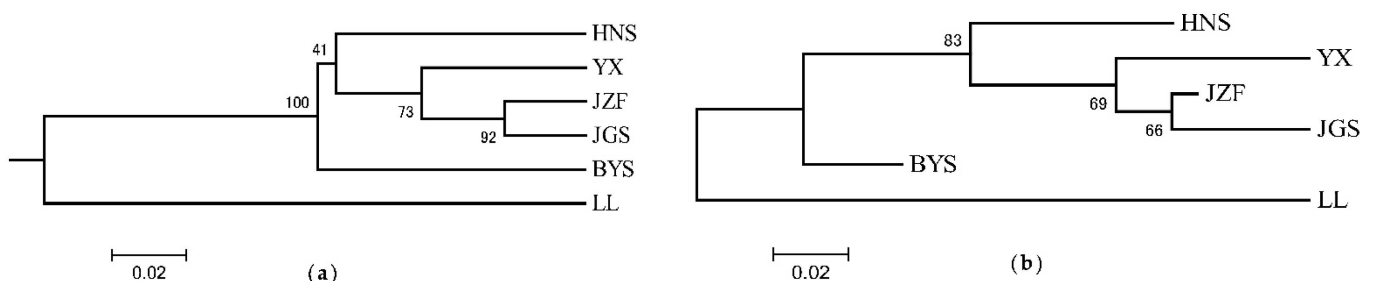
Figure 3. Phylogeny relationships among six populations of *S. tzumu* by unweighted pair-group method (UPGMA) (**a**) and neighbor-joining (NJ) method (**b**) based on D_A distance.
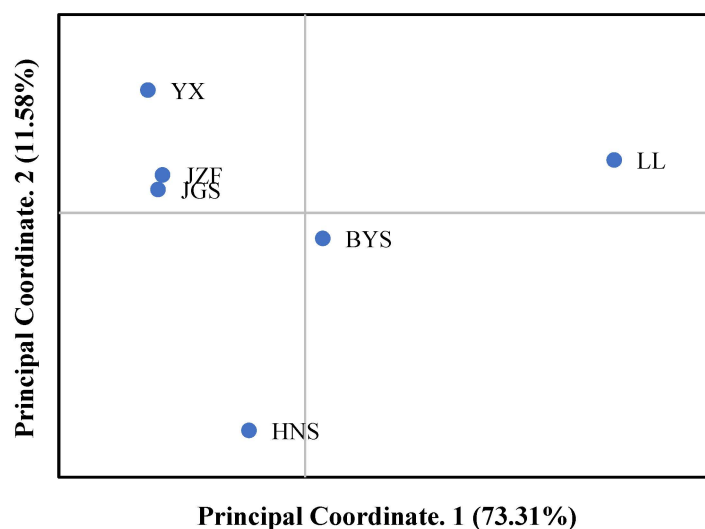
Figure 4. Principal coordinate analysis (PCoA) among six populations based on mean genotypic genetic distance.

3. Discussion

3.1. Characterization, Polymorphism, and Transferability of Newly Developed nSSRs of S. tzumu

In this study, following Zhang et al.'s effective method and under the help of the pipeline of CandiSSR [31], we successfully predicted plenty of putative polymorphic nSSR loci from the sequenced nuclear genome's contigs and scaffolds of six geographically distant individuals of *S. tzumu*. Compared with the traditional way of developing SSR markers, the above procedures are believed to be much more time- and energy-saving.

For the details of our candidate polymorphic nSSR loci for *S. tzumu*, we found that dinucleotide repeats were the most abundant type (75.53%) followed by the trinucleotide repeats (19.75%) (Figure 1). This distribution was commonly found in other tree species, such as *Dalbergia hupeana* (dinucleotide repeats account for 69.67%, whereas trinucleotide repeats account for 27.03%); and *Parrotia subaequalis* (80.54%, 16.77%) [36,37]. Worth mentioning is the fact that the trinucleotide repeat is deemed to be one of the most abundant repetition motif types for tree or fruit species because they hold less chance to the frameshifts in the coding sequences of species [36,38]. Meanwhile, a relatively low SSR density was observed in the coding sequences of species because the high mutation rate of SSR may affect the gene function, regulation of gene transcription, translation, and metabolic activities [38]. Indeed, the obtained 27 polymorphic nSSRs were annotated against the GenBank nonredundant database, among which, seven nSSR loci were correspondingly linked to transcription and other vital activities (Table 1). In addition, for the assessment of the polymorphism of these 27 novel nSSRs, the majority of the developed nSSRs (88.89%) showed high or moderate polymorphism (*PIC* > 0.25), which indicated that these nSSRs have great power to illuminate the population genetic diversity and variation of *S. tzumu*.

Furthermore, many previous studies indicated that SSR markers have high cross-species transferability among closely related species in plants [7,36,37,39]. For the genus *Sassafras* here, 20 nSSRs from *S. randaiense* has been successfully transferrable to *S. tzumu* using one individual [7]; in our study, the tests of interspecies transferability of the 27 nSSRs from *S. tzumu* to the other two *Sassafras* species yielded a transfer success rate of 85.19% (23 loci), and the proportion of polymorphic primers for *S. albidum* and *S. randaiense* was relatively high (62.96% and 51.85%, respectively) (Tables S1 and S2). It revealed a relatively high cross-species transferability of these nSSR markers, which will have great assistance in the studies of population genetics, phylogeography, phylogenetics, molecular breeding, and resource management of *S. tzumu* and the other Sassafras (Lauraceae) species.

3.2. Population Genetic Diversity of S. tzumu

Extensive literature pointed out that plant genetic diversity could be affected by various internal and external factors; for internal factors, they include the breeding system, genetic drift (bottleneck effect as an exceptional case), natural selection, gene mutation, and gene flow; whereas the environmental change and human interference are supposed to be the external factors [40–50]. The expected heterozygosity (H_E) and Shannon's information index (I) are commonly used to assess population genetic diversity [36]. In this study, H_E and I were estimated to be 0.317 and 0.527, respectively, at the population level (Table 3), which were higher than the three populations in Hubei Province in China of *S. tzumu* ($H_E = 0.280$) uncovered by isozyme analysis [26], but significantly lower than the five populations in Tianmu Mountain along with different altitudes ($H_E = 0.802$) revealed by SSR markers [27]. We assumed that these differences might be due to different methods and sample distributions [51]. At the species level, our results revealed an intermediate level of genetic diversity ($H_E = 0.430$, $I = 0.764$) (Table 2), which was consistent with the study of Liu ($H_E = 0.3004$) [28]. Meanwhile, this genetic diversity calculated by our study with more extensive sampling covering the distribution range of *S. tzumu* was lower than the average H_E of long-lived perennial plants ($H_E = 0.68$), but slightly higher than the average of endemic plants ($H_E = 0.42$) [41]. Additionally, compared with the other species of Lauraceae these share the similar distribution range of *S. tzumu*, the genetic diversity of *S. tzumu* in this study was generally lower than *Machilus pauhoi* ($H_E = 0.751$) [52], *Lindera glauca* ($H_E = 0.59$, $I = 1.17$) [53], *Cinnamomum camphora* ($H_E = 0.785$) [54], and *Persea americana* ($H_E = 0.6927$) [55] but a little higher than those of some endangered species, such as *Litsea szemaois* ($I = 0.3826 \pm 0.2333$) [56], *Cinnamomun chago* ($H_E = 0.340$) [57], and *Phoebe bournei* ($H_E = 0.43$, $I = 0.72$).

Considering that *S. tzumu* is a species with entomophily, and the main factor affecting genetic diversity is the breeding system [27,35,40,44,58,59], therefore, *S. tzumu* has a lower gene flow than plants with anemophily [40]. Furthermore, the characteristics of pollen abortion, short time to maintain pollen viability, a low proportion of normal-growing embryo sacs, and low seed setting percentage of *S. tzumu* [9] also lead to a lower level of gene flow [40] ($N_m = 0.625$) (Table 5). Moreover, the relatively high level of inbreeding ($F_{IS} = 0.166$) and heterozygote deficit were detected in *S. tzumu*, which also resulted in its low genetic diversity. Furthermore, two populations (YX and JGS) were examined experiencing a recent genetic bottleneck effect [60]; we assumed that human interference and habitat fragmentation were the significant affect factors for this result according to the field investigation. As indicated by a previous study, the total distribution areas of *S. tzumu* have migrated southward and decreased by 15.06% from the Last Glacial Maximum (LGM) to present [61]. LL, as the westernmost population at the edge of the distribution areas, was located in the distribution area with contraction from the LGM to present, which corresponded to the result of the lowest genetic diversity of LL among the six populations (Table 3). Some species in partially overlapping distribution areas with *S. tzumu* also had low genetic diversity due to the above factors, such as *Salix taishanensis* ($I = 0.3505$) [62], *Magnolia sinostellata* ($H_E = 0.27312$) [63], *Michelia chapensis* ($H_E = 0.3255$) [64], and *Sinomanglietia glauca* ($H_E = 0.423$) [65].

3.3. Population Structure and Differentiation of S. tzumu

The population genetic structure has an inseparable relationship with the biological evolution process, which included natural selection in heterogeneous environments, genetic drift, and gene flow [66,67]. Structure analysis in this study optimally divided six selected *S. tzumu* populations into two clusters: LL, the westernmost population in the Yunnan–Guizhou Plateau, formed a separate cluster, whereas the other five populations (BYS, HNS, YX, JZF, and JGS) were assigned to another cluster (Figure 2). Moreover, there are many mountains in subtropical China [68] based on the study about the distribution of the mountain ranges and peaks in China [69], and the significant effect of isolation by distance (IBD) was detected by the Mantel test ($r = 0.742$, $p = 0.006$) in this study (Figure S1), which

corresponded to the relatively large geographic distance and barriers of multiple mountain ranges of the six populations of *S. tzumu*, and these two factors also played facilitating roles in promoting genetic differentiation and hindering the gene flow. A previous study also revealed the effect of IBD and the barrier of the Meiling Mountain and Lu Mountain in *S. tzumu* [28]. In the present study, we identified that the Wuling Mountain, Dalou Mountain, and Yunnan–Guizhou Plateau played major barrier roles among LL and the other five populations. Similar mechanism of population structures has been reported in other species, such as *Cerasus dielsiana* (Wuling Mountain) [70]; *Tetracentron sinense* (Qinling Mountain, Hengduan Mountain, and Dalou Mountain) [35]; *Sinomanglietia glauca* (Wuling Mountain) [65]; and *Phoebe zhennan* (Wu Mountain and Wuling Mountain) [71].

When it comes to the level of population differentiation, our study unveiled a relatively high genetic differentiation (F_{ST} = 0.286) that existed among the six populations of *S. tzumu*. Compared with other forest trees of Lauraceae, it was much higher than *Cinnamomun camphora* (F_{ST} = 0.109) [54], *Cinnamomun chago* (F_{ST} = 0.2198) [57], and *Cryptocarya chinensis* (F_{ST} = 0.141) [72] but lower than *Lindera glauca* (F_{ST} = 0.362) [73] and *Machilus pauhoi* (F_{ST} = 0.3366) [74], and had a similar value with *Phoebe bournei* (F_{ST} = 0.287) [75]. In addition, the AMOVA indicated that 28.40% (Table 4) of the total variation was distributed among the populations, and most of the genetic variation was distributed within the populations. This finding agreed that long-lived, outcrossing, and/or late successional taxa retain most of their genetic variability within populations with the analyses based on microsatellite DNA loci [41]. Apart from the geographic distance, Liu pointed out that the environmental distance had a more significant effect on the genetic distance than the geographic distance (IBE = 1.54 IBD), especially with the influence of the mean annual temperature [28], and the amount of precipitation was also considered to be the main effect factor on the suitability degree of the distribution areas of *S. tzumu* [61]. These findings supported our views and were consistent with the biological character of habitat preference with warmth and wetness of *S. tzumu* [28]. Therefore, both the geographic distance and climate difference played important roles for the intraspecific divergence of *S. tzumu*.

3.4. Conservation Implications for S. tzumu

Above all, a moderately low level of genetic diversity is believed to relatively lack the ability to adapt to various habitats and changing climate [28,66], and increase the risk of being endangered for a species [7,8]. Here, in our study, *S. tzumu* presented an intermediate level of genetic diversity in its distribution range. Hence, it is urgent and crucial to protect this Arctic Tertiary relict endemic species in China. As revealed by our study herein, population LL (the westernmost population) and the other five populations (BYS, HNS, YX, JZF, and JGS) of *S. tzumu* formed two detached clusters, which could be considered as two ESUs for its long-term preservation of genetic resources. What is more, all populations should be protected in situ based on the low genetic diversity among populations as resources (money, manpower, etc.) permit.

4. Materials and Methods

4.1. Plant Materials and DNA Extraction

The fresh young leaves of six individuals were collected from six geographic representative populations in southern China (Figure 2, Table 6) to represent the various distributions of *S. tzumu*. To evaluate and validate the polymorphism of the nSSRs developed from our sequencing data, the leaf material of two individuals in *S. tzumu* was collected from each of the six populations from across the distribution range of the species in southern China. For each population, more than 14 individuals were sampled at least 10 m apart to minimize the likelihood of sampling clones, resulting in a total of 136 individuals (Table 6). The polymorphic nSSRs were further tested for transferability to two related species in *Sassafras* genus, *S. albidum* and *S. randaiense* (Table S10). The leaf materials of 22 individuals of *S. albidum* were collected in 22 different regions in the USA and 10 individuals of *S. randaiense* in Taiwan, China (Table S10). The collected leaf materials were

dried and stored in silica gel at 4 °C for further DNA extraction. The representative voucher specimens of the *Sassafras* species were deposited in the Herbarium of Nanjing University Table 6 and Table S10). Total genomic DNA was extracted from leaf tissue using Plant DNA-zol Reagent (Invitrogen Life Technologies), following the manufacturer's protocol. DNA quality was examined on a 0.8% agarose gel stained with 1× GelRed (Biotium Inc., Fremont, CA, USA), and the concentration was checked using a Nano-Drop Lite spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA).

Table 6. Information of plant materials of *S. tzumu* collected in this study.

Population Code	Collection Locality	Geographic Coordinates	Number of Individuals	Voucher Specimens
BYS	Baiyun Mountain Natural Reserve, Hunan Province, China	28.7301° N, 109.3955° E	21	ZJB201903
LL	Leli Village Forestry Farm, Yunnan Province, China	27.4202° N, 104.8601° E	20	ZJB201904
HNS	Huangniushi Natural Reserve, Guangdong Province, China	24.4294° N, 114.4556° E	22	ZJB201905
YX	Yixing National Forest Park, Jiangsu Province, China	31.6986° N, 119.7449° E	26	ZJB201906
JZF	Junzifeng Peak Natural Reserve, Fujian Province, China	26.4863° N, 116.6559° E	33	ZJB201907
JGS	Jigong Mountain Natural Reserve, Henan Province, China	31.8185° N, 114.0583° E	14	ZJB201908

4.2. NSSR Development, Polymorphism, and Transferability Assessment

One individual was selected from each of the 6 populations to construct six next-generation sequencing genomic paired-end DNA libraries using the low-coverage shotgun sequencing in the HiSeq XTen platform (Illumina Inc., San Diego, CA, USA) at BGI (China). For each individual, about 7 Gb of raw data was obtained with paired-end 150 bp read length. Following Zhang et al.'s effective method [31], we filtered the chloroplast- and mitochondria-containing contigs and used CandiSSR [19] to identify potentially polymorphic nSSR primer pairs from the remaining scaffolds of the assembled nuclear genome database with a individual (from the BYS population) as the reference. Then, the specific potential polymorphic nSSRs were selected based on the above method [31]. The total selected candidate nSSRs were used for preliminary screening, and the corresponding primers were synthesized by Sangon Biotech (Shanghai, China).

Six randomly selected individuals from six populations were used to initially screen the amplification success and effectiveness of the primer pairs and verify the optimal primer temperature by PCR amplifications and 1% agarose gel stained with 1× GelRed (Biotium Inc., Fremont, CA, USA). The polymorphism of each locus was then performed by selecting 12 individuals from 6 representative geographically distant populations (two individuals were chosen from each population at random) (Table 6). Cross-species amplification was then conducted in two more *Sassafras* species (Table S10) to test the transferability of the polymorphic nSSRs above. PCR amplifications were performed on a GeneAmp9700 DNA Thermal Cycler (Perkin-Elmer, Waltham, MA, USA) following the standard protocol of the AmpliTaq Gold 360 Master PCR kit (Thermofisher Biotech Company, Applied Biosystems, Foster City, CA, USA) in a final volume of 15 µL, which contained 1 µL of template DNA, 7.5 µL of AmpliTaq Gold 360 Master Mix (Thermofisher Biotech Company, Applied Biosystems, Foster City, CA, USA), 5.5 µL of deionized water, 0.5 µL of reverse primers (10µM), and 0.5 µL of forward primers (10µM), with the 5′ end labeled with a fluorescent dye (FAM, HEX, TAMRA, or ROX) (Table S1). Moreover, the PCR procedure was as follows: 5 min initial denaturation at 95 °C, 35 cycles of 45 s at 95 °C, 30 s annealing at an optimal primer temperature (Table S1) and 30 s synthesis at 72 °C, followed by a final 10 min extension step at 72 °C and a 4 °C maintained temperature. The PCR products were detected by an ABI 3730XL capillary electrophoresis analyzer (Applied Biosystems,

Foster City, CA, USA) with GeneScan-500LIZ as an internal reference (Applied Biosystems). Fragments were genotyped for their presence or absence at each locus, and the allele sizes were scored by using GeneMarker v2.2.0 (SoftGenetics, State College, PA, USA) and manually checked twice to reduce genotyping errors.

4.3. Assessment of Genetic Diversity of S. tzumu

For each nSSR locus, the polymorphism information content (PIC), number of alleles (N_A), expected heterozygosity (H_E), and observed heterozygosity (H_O) were calculated by CERVUS v3.0.7 [76]; the total genetic diversity (H_T) and average genetic diversity within populations (H_S) were estimated by FSTAT v2.9.4 [77]; the inbreeding coefficient (F_{IS}) was calculated by INEst v2.2 [78]; and the Shannon's information index (I) and effective number of alleles (N_E) were counted by Popgene v1.31 [79]. Additionally, deviation from the Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium (LD) [80] for each locus was tested using a Markov chain (dememorization: 10,000; batches: 20, iterations per batch: 5000) in GENEPOP v4.7.5 [81]. The null allele frequency of each locus in each population was estimated based on the expectation maximization algorithm in FREENA [82].

For each population of *S. tzumu*, the allelic richness (A_R) was estimated by FSTAT v2.9.4 [76], and the following genetic diversity parameters were calculated by GenALEx v6.502 (Canberra, Australia) [83]: the average number of alleles (A), effective number of alleles (A_E), number of private alleles (A_P), observed heterozygosity (H_O), expected heterozygosity (H_E), unbiased expected heterozygosity (μH_E), and Shannon's information index (I). Furthermore, the Wilcoxon signed-rank test and Z-test were used to estimate the excess of heterozygosity under the infinite model (IAM), strict stepwise mutation model (SMM), and two-phase model (TPM) by INEst v2.2 [77] to perform a bottleneck test [84] for each population.

4.4. Population Structure and Differentiation of S. tzumu

The total genetic differentiation coefficient (F_{ST}) and pairwise population genetic differentiation index (F_{ST}) were calculated by Genetix v4.05 program [85], and an estimate of N_m (the number of migrants per generation/gene flow) among the populations was computed using the formula $N_m = 0.25(1 - F_{ST})/F_{ST}$ [86]. In addition, the analysis of molecular variance (AMOVA) was performed to partition the total genetic variance among populations using ARLEQUIN v3.5.2.2 [87]. Moreover, GenALEx v6.502 (Canberra, Australia) [83] was used to compute the principal coordinate analysis (PCoA) and perform the Mantel test to estimate the correlation between Nei's unbiased genetic distance and geographic distance (9999 permutations) to reveal the influence of the isolation-by-distance (IBD) model. Moreover, the study of the phylogenetic relationship among the populations of *S. tzumu* was performed by POPTREE [88] based on shared allele distances among the populations (D_A genetic distance) to build the UPGMA and NJ tree [89] using unweighted pair-group method (UPGMA) [90] and neighbor-joining method (NJ) [91].

Bayesian methods were utilized to reveal the optimal number of genetic clusters (K) for the studied populations, and these Bayesian analyses were implemented using STRUCTURE v2.3.4 [92], with ten independent runs for each K value that ranges from 1 to 10 that were performed to determine the maximum value of the posterior probability of the data [ln$p$(D)] with 100,000 MCMC (Markov Chain Monte Carlo) repetitions and 100,000 burn-in periods. We assumed that the studied populations conformed to the admixture model and independent allele frequencies. The obtained results from STRUCTURE were further interpreted by STRUCTURE HARVESTER [80], which implemented Evanno's ΔK method [93] to estimate the parameter (ΔK) that corresponded to the change of ln$p$(D) between consecutive K values for the calculation of the optimal K. The most likely number of cluster (K) was identified as the one that maximized ln$p$(D) and/or ΔK.

5. Conclusions

In this study, we first identified 16,215 potential polymorphism nSSR loci based on the assembled nuclear genome database of six individuals of *S. tzumu*. The di- (75.53%) and trinucleotide (19.75%) repeats were the most abundant. A total of 27 novel nSSRs with polymorphism were utilized to illuminate the genetic diversity of 136 individuals in six populations of *S. tzumu*, resulting most markers (88.89%) showing high or moderate polymorphism (*PIC* = 0.356). The majority of nSSRs (85.19%) could be successfully transferred to the other two *Sassafras* species, and some of them showed polymorphism. The population genetic results revealed a moderately low genetic diversity of *S. tzumu* with H_E ranging from 0.195 (LL) to 0.387 (BYS) at the population level, and the average H_E was 0.430 at the species level. The bottleneck effect appeared in some populations (YX and JGS). The population differentiation was high (F_{ST} = 0.286), and the majority of variations (71.60%) were within the populations. Consistent with the results of the principal coordinate analysis (PCoA) and phylogenetic trees, structure analysis optimally divided these six *S. tzumu* populations into two clusters, and the further strong population subdivision appeared from K = 2 to K = 5, which corresponded to two evolutionarily significant units (ESUs). Moreover, the significant correlation between genetic and geographic distance was tested by the Mantel test (r = 0.742, p = 0.006), clarifying the effect about isolation by distance (IBD), which could be possibly explained by the low gene flow (N_m = 0.625), a relatively high degree of inbreeding (F_{IS} = 0.166), a relatively large distribution, and mountainous barriers. Taken together, our research gives a deeper understanding of the population genetic diversity and structure of *S. tzumu* in subtropical China and will contribute to the conservation and utilization of *S. tzumu* and the other *Sassafras* (Lauraceae) species.

References

1. Li, S.H.; Li, X.W.; Li, J.; Huang, P.H.; Wei, F.N.; Cui, H.B.; van der Werff, H. Lauraceae. In *Flora of China*; Science Press: Beijing, China, 1984; Volume 7.
2. Little, E.L. *National Audubon Society Field Guide to North American Trees*; Chanticleer: New York, NY, USA, 1998.
3. Nie, Z.L.; Wen, J.; Sun, H. Phylogeny and biogeography of *Sassafras* (Lauraceae) disjunct between eastern Asia and eastern North America. *Plant Syst. Evol.* **2007**, *267*, 191–203. [CrossRef]

4. Chanderbali, A.S.; van der Werff, H.; Renner, S.S. Phylogeny and historical biogeography of Lauraceae: Evidence from the chloroplast and nuclear genomes. *Ann. Mo. Bot. Gard.* **2001**, *88*, 104–134. [CrossRef]

5. Rohde, R.; Rudolph, B.; Ruthe, K.; Lorea-Hernandez, F.G.; de Moraes, P.L.R.; Li, J.; Rohwer, J.G. Neither Phoebe nor Cinnamomum—The tetrasporangiate species of *Aiouea* (Lauraceae). *Taxon* **2017**, *66*, 1085–1111. [CrossRef]

6. Rehder, A. The American and Asiatic species of *Sassafras*. *J. Arnold Arbor.* **1920**, *1*, 242–245. [CrossRef]

7. Chung, K.F.; Lin, T.T.; Tsai, Y.S.; Lin, S.T.; Peng, C.I. Isolation and Characterization of Microsatellite Loci in *Sassafras Randaiense* (Lauraceae). *Am. J. Bot.* **2011**, *98*, E326–E329. [CrossRef]

8. Zhang, K.L.; Zhang, Y.; Jia, D.W.; Tao, J. Species Distribution Modeling of Sassafras Tzumu and Implications for Forest Management. *Sustainability* **2020**, *12*, 4132. [CrossRef]

9. Liu, N.N.; Liu, J.; Zhang, M.Y.; Wang, L.; Yang, S.L.; Liao, W.B. Characteristics of *Pinus taiwanensis + Sassafras tzumu* community in Taoyuandong National Nature Reserve of Hu'nan Province. *J. Plant Resour. Environ.* **2017**, *26*, 84–92.

10. Wang, X.; Yang, S.G.; Yu, F.; Ji, C.F.; Long, C.L.; Jiang, C.M. Research progress of *Sassafras tzumu*. *S. China For. Sci.* **2015**, *43*, 29–33. [CrossRef]

11. Zhou, X.P. Comparative analysis of the effect of wood mixed with wood-load afforestation. *For. By-Prod. Spec. China* **2022**, *4*, 22–24. [CrossRef]

12. Chen, H.Z.; Liu, J.; Jiang, J.M.; Sun, Y.; Dong, X.; Fu, A.P. A study on flowering and fruit development of *Sassafras tzumu*. *For. Res.* **2020**, *33*, 148–154. [CrossRef]

13. Zhang, Y.; He, Q.; Zhou, X.; Zheng, S.; Wang, Y.; Li, P.; Wang, Y. Genetic diversity and population structure of 93 rice cultivars (lines) (*Oryza sativa* Xian group) in Qinba in China by 3 types of genetic markers. *BMC Genom.* **2022**, *23*, 550. [CrossRef]

14. Li, C.; Li, J.P.; Chai, B.F. Development of microsatellites DNA marker and Its application. *J. Zhengzhou Inst. Aeronaut. Ind. Manag. Soc. Sci. Ed.* **2004**, *23*, 199–200. [CrossRef]

15. Powell, W.; Machray, G.C.; Provan, J. Polymorphism revealed by simple sequence repeats. *Trends Plant Sci.* **1996**, *1*, 215–222. [CrossRef]

16. Tautz, D. Hypervariability of Simple Sequences as a General Source for Polymorphic DNA Markers. *Nucleic Acids Res.* **1989**, *17*, 6463–6471. [CrossRef] [PubMed]

17. Varshney, R.K.; Graner, A.; Sorrells, M.E. Genic microsatellite markers in plants: Features and applications. *Trends Biotechnol.* **2005**, *23*, 48–55. [CrossRef]

18. Wen, J.H. Study on genetic diversity and construction of core collections about Vitis amurensis Rupr. germplasm resources by SSR markers. Ph.D. Thesis, Jilin Agricultural University, Changchun, China, 2011.

19. Xia, E.H.; Yao, Q.Y.; Zhang, H.B.; Jiang, J.J.; Zhang, L.P.; Gao, L.Z. CandiSSR: An Efficient Pipeline used for Identifying Candidate Polymorphic SSRs Based on Multiple Assembled Sequences. *Front. Plant Sci.* **2016**, *6*, 1171. [CrossRef]

20. Wang, X.; Yu, F.; Ji, C.F.; Long, C.L.; Yang, G.Y. Anatomical study on pollen development of *Sassafras tzumu* (Hemsl.) Hemsl. *J. Gansu Agric. Univ.* **2014**, *49*, 116–119. [CrossRef]

21. Yang, Z.; Tan, C.; Wei, Y.-M.; Rohwer, J.G.; Liu, B.; Yang, Y. Floral morphology and phenology of *Sassafras tzumu* (Lauraceae). *BMC Plant Biol.* **2022**, *22*, 1–12. [CrossRef]

22. Sun, J.J.; Jiang, B.; Zhu, J.R.; Wu, D.T.; Ye, N.N.; Qiu, H.J.; Yuan, W.G.; Wu, C.P.; Huang, Y.J.; Jiao, J.J.; et al. Prediction of potential suitable habitat and dominate environmental factors of *Sassafras tzumu* in Zhejiang with MaxEnt Model. *J. Northeast. For. Univ.* **2020**, *48*, 1–6. [CrossRef]

23. Sun, J.J.; Jiang, B.; Wu, C.P.; Yuan, W.G.; Zhu, J.R.; Huang, Y.J.; Jiao, J.J.; Shen, A.H. Study on the habitat and niche of *Sassafras tzumu* (Hemsl.) Hemsl. in Zhejiang Province. *Acta Ecol. Sin.* **2019**, *39*, 884–894.

24. Long, X.F. Effects of Different Mixing Methods on the Growth of *Pinus massoniana* and *Sassafras tsumu*. *J. Anhui Agric. Sci.* **2019**, *47*, 140–141.

25. Bi, G.Y.; Jiang, X.H.; Jiang, S.F. Growth and silvicultural technology of *Sassafras tzumu*. *Pract. For. Technol.* **2003**, *12*, 17. [CrossRef]

26. Ding, X.F.; Chen, H.L.; Cao, J.; Koono, K.; Ubukata, M.; Okamura, M. Preliminary study on genetic st ructure of three natural populations in *Sassafrass Tzumu*. *Hubei For. Sci. Technol.* **2006**, *005*, 1–2.

27. Jiang, A.P.; Jiang, J.M.; Liu, J. Genetic diversity and genetic structure in *Sassafras tsumu* populations along altitudinal gradients in Tianmushan Mountain, China. *Chin. J. Appl. Ecol.* **2016**, *27*, 1829–1836. [CrossRef]

28. Liu, X. Landscape genetics of *Sassafras tzumu* based on the Genotyping-by-Sequencing (GBS). Master's Thesis, Nanchang University, Nanchang, China, 2019.

29. Sylva, E.B.O.C. *China Sylva: The First Volume Lauraceae*; China Forestry Press: Beijing, China, 1982.

30. Kameyama, Y. Development of Microsatellite Markers for *Cinnamomum Camphora* (Lauraceae). *Am. J. Bot.* **2012**, *99*, E1–E3. [CrossRef]

31. Zhang, Y.Y. The conservation genetics research of *Parrotia subaequalis* (Hamamelidaceae), an endangered palaeoendemic plant in China. Master's Thesis, Nanjing University, Nanjing, China, 2018.

32. Altschul, S.F.; Madden, T.L.; Schaffer, A.A.; Zhang, J.H.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [CrossRef]

33. Wen, Y.F.; Uchiyama, K.; Han, W.J.; Ueno, S.; Xie, W.D.; Xu, G.B.; Tsumura, Y. Null alleles in microsatellite markers. *Biodivers. Sci.* **2013**, *21*, 117–126.

34. Wang, Z.F. *Molecular Ecology and Fundamentals of Data Analysis*; Science Press: Beijing, China, 2016.

35. Tian, Z.Q. The development of SSR markers and genetic structure analysis of an endangered plant *Tetracentron sinense* Oliv. Master's Thesis, China West Normal University, Nanchong, China, 2020.

36. Li, C.H.; Zheng, Y.Q.; Liu, Y.; Lin, F.R.; Huang, P. Development of Genomic SSR for the Subtropical Hardwood Tree Dalbergia hupeana and Assessment of Their Transferability to Other Related Species. *Forests* **2021**, *12*, 804. [CrossRef]

37. Zhang, Y.Y.; Shi, E.; Yang, Z.P.; Geng, Q.F.; Qiu, Y.X.; Wang, Z.S. Development and Application of Genomic Resources in an Endangered Palaeoendemic Tree, *Parrotia subaequalis* (Hamamelidaceae) From Eastern China. *Front. Plant Sci.* **2018**, *9*, 246. [CrossRef]

38. Song, X.; Li, N.; Guo, Y.; Bai, Y.; Wu, T.; Yu, T.; Feng, S.; Zhang, Y.; Wang, Z.; Liu, Z. Comprehensive identification and characterization of simple sequence repeats based on the whole-genome sequences of 14 forest and fruit trees. *For. Res.* **2021**, *1*, 1–10. [CrossRef]

39. Zhang, M.Y.; Zhang, Y.Y.; Wang, G.H.; Zhou, J.B.; Tian, Y.J.; Geng, Q.F.; Wang, Z.S. Development and characterization of 20 novel EST-SSR markers for *Pteroceltis tatarinowii*, a relict tree in China. *Appl. Plant Sci.* **2020**, *8*, e11320. [CrossRef] [PubMed]

40. Wen, Y.F.; Han, W.J.; Wu, S. Plant genetic diversity and its influencing factors. *J. Cent. South Univ. For. Technol.* **2010**, *30*, 80–87. [CrossRef]

41. Nybom, H. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Mol. Ecol.* **2004**, *13*, 1143–1155. [CrossRef]

42. Slatkin, M. Gene Flow and the Geographic Structure of Natural-Populations. *Science* **1987**, *236*, 787–792. [CrossRef] [PubMed]

43. Lande, R. Genetics and Demography in Biological Conservation. *Science* **1988**, *241*, 1455–1460. [CrossRef] [PubMed]

44. Wang, H.X.; Hu, Z.A. Plant breeding system genetic structure and conservation of genetic diversity. *Chin. Biodivers.* **1996**, *4*, 92–96.

45. Nevo, E. Evolution of genome-phenome diversity under environmental stress. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 6233–6240. [CrossRef]

46. Frankham, R.; Ballou, J.D.; Briscoe, D.A. *Introduction to Conservation Genetics*; Cambridge University Press: Cambridge University, Cambridge, UK, 2002.

47. Morjan, C.L.; Rieseberg, L.H. How species evolve collectively: Implications of gene flow and selection for the spread of advantageous alleles. *Mol. Ecol.* **2004**, *13*, 1341–1356. [CrossRef]

48. Petit, R.J.; Duminil, J.; Fineschi, S.; Hampe, A.; Salvini, D.; Vendramin, G.G. Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Mol. Ecol.* **2005**, *14*, 689–701. [CrossRef]

49. Freeland, J.R. *Molecular Ecology*; John Wiley & Sons Ltd.: West Sussex, UK, 2005.

50. Xu, G.B. *Plant Population Genetics*; Science Press: Beijing, China, 2009.

51. Vasilyeva, Y.; Chertov, N.; Nechaeva, Y.; Sboeva, Y.; Pystogova, N.; Boronnikova, S.; Kalendar, R. Genetic Structure, Differentiation and Originality of *Pinus sylvestris* L. Populations in the East of the East European Plain. *Forests* **2021**, *12*, 999. [CrossRef]

52. Zhu, Q.; Li, P.; Zhou, P.; Zhang, J.J.; Que, Q.M.; Hui, W.K.; Chen, X.Y. Optimization of SSR-PCR reaction system and study on genetic diversity of natural populations of *Machilus pauhoi*. *For. Res.* **2019**, *32*, 70–78. [CrossRef]

53. Xiong, B.; Zhang, L.M.; Ha, D.L.; Ju, Y.X.; Zhang, Z.X. Construction of core germplasm of wild *Lindera glauca* based on SSR molecular marker. *Mol. Plant Breed.* **2018**, *16*, 6380–6389. [CrossRef]

54. Liu, S. Population genetic structure of *Cinnamomum camphora* in China. Master's Thesis, Jiangxi Agricultural University, Nanchang, China, 2019.

55. Dong, M.C.; Yue, J.Q.; Li, J.X.; Yang, F.; Zhou, D.G.; Wang, S.H.; Long, C.R.; Guo, L.N.; Gao, J.Y.; Yang, E.C. Genetic diversity analysis of germplasm resources of *Persea americana* Mill. based on SSR fluorescent marker. *Mol. Plant Breed.* **2020**, *18*, 5403–5410. [CrossRef]

56. Chen, J.Q.; Ci, X.Q.; Li, Q.M.; Li, J. Genetic diversity of *Litsea szemaois*, an endangered species endemic to China, detected by inter-simple sequence repeat (ISSR). *Biodivers. Sci.* **2006**, *14*, 410–420. [CrossRef]

57. Zhang, X. Species germplasm conservation of the wild plants with extremely small populations—*Cinnamomum chago*. Ph.D. Thesis, Yunnan University, Kunming, China, 2019.

58. Zhang, X. Reproductive Biology and Conservation Genetics of *Rhododendron sinofalconeri*. Master's Thesis, Chinese Academy of Forestry, Beijing, China, 2020.

59. Pu, T.L.; Han, X.Q.; Luo, H.Y.; Deng, H.S.; Zou, M.L.; Jin, J.; Xia, Z.Q.; Wang, W.Q. Analysis of genetic of Moringa oleifera population based on SNP markers. *Chin. J. Trop. Crops* **2022**, *49*, 1–12.

60. Cornuet, J.M.; Luikart, G. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **1996**, *144*, 2001–2014. [CrossRef]

61. Liu, X.; Gong, X.; Chen, S.S.; Guan, B.C. Simulation of the distribution pattern of *Sassafras tzumu* and changes in habitat based on ArcGIS and MaxEnt. *Plant Sci. J.* **2018**, *36*, 320–326.

62. Liu, H.Y. Genetic Diversity and Population Ecology of *Salix taishanensis*. Master's Thesis, Shandong Agricultural University, Tai'an, China, 2020.

63. Yao, Z.X. The conservation genetics research of *Magnolia sinostellata*. Master's Thesis, South China Agricultural University, Guangzhou, China, 2019.

64. Jiang, J.M.; Teng, H.J.; Yuan, J.L.; Luan, Q.F.; Tan, Z.F. Genetic diversity of *Michelia chapensis* dandy populations. *For. Res.* **2005**, *18*, 109–113.

65. Xiong, M.; Tian, S.; Zhang, Z.R.; Fan, D.M.; Zhang, Z.Y. Population genetic structure and conservation units of *Sinomanglietia glauca* (Magnoliaceae). *Biodivers. Sci.* **2014**, *22*, 476–484.

66. Zhou, P.Y.; Hui, L.X.; Huang, S.J.; Ni, Z.X.; Yu, F.X.; Xu, L.A. Study on the Genetic Structure Based on Geographic Populations of the Endangered Tree Species: *Liriodendron chinense*. *Forests* **2021**, *12*, 917. [CrossRef]

67. Chen, X.Y.; Lu, H.P.; Shen, L.; Li, Y.Y. Identifying populations for priority conservation of important species. *Biodivers. Sci.* **2002**, *10*, 332–338.

68. Peng, B.Z.; Chen, F. Progress in the study of mountain vertical zonation in China. *Sci. Geogr. Sin.* **1999**, *19*, 303–308.

69. Wang, X.P.; Wang, Z.H.; Fang, J.Y. Mountain ranges and peaks in China. *Biodivers. Sci.* **2004**, *12*, 206–212.

70. Zhu, H. Phylogenetic position and population biogeography of *Cerasus dielsiana* (Rosaceae). Ph.D. Thesis, Nanjing Forest University, Nanjing, China, 2020.

71. Gao, J.H.; Zhang, W.; Li, J.Y.; Long, H.L.; He, W.; Li, X.Q. Amplified fragment length polymorphism analysis of the population structure and genetic diversity of *Phoebe zhennan* (Lauraceae), a native species to China. *Biochem. Syst. Ecol.* **2016**, *64*, 149–155. [CrossRef]

72. Gao, S.H.; Wang, Z.F.; Zhang, J.L.; Tian, S.N. Genetic diversity of *Cryptocarya chinensis* life stages in heishiding Guangdong Province. *J. Plant Genet. Resour.* **2005**, *6*, 300–303. [CrossRef]

73. Xiong, B.; Zhang, L.M.; Dong, S.B.; Zhang, Z.X. Population genetic structure and variability in *Lindera glauca* (Lauraceae) indicates low levels of genetic diversity and skewed sex ratios in natural populations in mainland China. *PeerJ* **2020**, *8*, e8304. [CrossRef]

74. Zhou, P.; Lin, W.; Zhu, Q.; Zhou, X.B.; Wu, L.Y.; Chen, X.Y. Genetic diversity of *Machilus pauhoi* assessed by SRAP markers. *J. Beijing For. Univ.* **2016**, *38*, 16–24. [CrossRef]

75. Huang, Y.Q. Genetic diversity analysis and core germplasm construction of *Phoebe bournei* based on SSR molecular markers. Master's Thesis, Chinese Academy of Forestry, Beijing, China, 2019.

76. Kalinowski, S.T.; Taper, M.L.; Marshall, T.C. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* **2007**, *16*, 1099–1106. [CrossRef]

77. Goudet, J. FSTAT (Version 2.9.3): A Program to Estimate and Test Gene Diversities and Fixation Indices. Available online: https://www2.unil.ch/popgen/softwares/fstat.htm (accessed on 21 July 2022).

78. Chybicki, I.J. INEST 2.2. The User Manual. Available online: https://www.ukw.edu.pl/pracownicy/strona/igo_chybicki/software_ukw (accessed on 1 August 2022).

79. Yeh, F.; Yang, R.; Boyle, T.; Ye, Z.; Mao, J.; Francis, C.Y.; Timothy, B.B.; Judy, X.M. POPGENE, Version 1.31. The User-Friendly Shareware for Population Genetic Analysis. Available online: http://www.ualberta.ca/~{}fyeh/index.htm (accessed on 3 September 2022).

80. Earl, D.A.; Vonholdt, B.M.; 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]

81. Rousset, F. Genepop'007: A complete re-implementation of the genepop software for Windows and Linux. *Mol. Ecol. Resour.* **2008**, *8*, 103–106. [CrossRef]

82. Chapuis, M.P.; Estoup, A. Microsatellite null alleles and estimation of population differentiation. *Mol. Biol. Evol.* **2007**, *24*, 621–631. [CrossRef] [PubMed]

83. Peakall, R.; Smouse, P.E. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Notes* **2006**, *6*, 288–295. [CrossRef]

84. Garza, J.C.; Williamson, E.G. Detection of reduction in population size using data from microsatellite loci. *Mol. Ecol.* **2001**, *10*, 305–318. [CrossRef] [PubMed]

85. Belkhir, K.; Borsa, P.; Chikhi, L.; Raufaste, N.; Bonhomme, F. GENETIX4. 05, logiciel sous Windows TM pour la génétiquedes populations. *Lab. Génome Popul. Interact. CNRS UMR* **2004**, *5000*, 1996–2004.

86. Wright, S. *Evolution and the Genetics of Populations*; University of Chicago Press: Chicago, IL, USA, 1984; Volume 1.

87. Excoffier, L.; Lischer, H.E.L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **2010**, *10*, 564–567. [CrossRef]

88. Takezaki, N.; Nei, M.; Tamura, K. POPTREE2: Software for Constructing Population Trees from Allele Frequency Data and Computing Other Population Statistics with Windows Interface. *Mol. Biol. Evol.* **2010**, *27*, 747–752. [CrossRef]

89. Nei, M.; Chesser, R.K. Estimation of Fixation Indexes and Gene Diversities. *Ann. Hum. Genet.* **1983**, *47*, 253–259. [CrossRef]

90. Sneath, P.H.A.; Sokal, R.R. *Numerical Taxonomy: The Principles and Practice of Numerical Classification*; Freeman: San Francisco, CA, USA, 1973.

91. Saitou, N.; Nei, M. The Neighbor-Joining Method—A New Method for Reconstructing Phylogenetic Trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425. [CrossRef]

92. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **2000**, *155*, 945–959. [CrossRef]

93. Evanno, G.; Regnaut, S.; Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **2005**, *14*, 2611–2620. [CrossRef]