

Article



# Population Study Reveals Genetic Variation and Introgression of Four Deciduous Oaks at the Junction between Taihang Mountain and Yanshan Mountain

Ziqi Pei <sup>1,2,†</sup>, Qinsong Yang <sup>1,2,†</sup>, Xining Chen <sup>1,2</sup>, Yu Zong <sup>3</sup>, Jinjin Li <sup>1,2</sup>, Xiong Yang <sup>1,2</sup>, Chenrui Huo <sup>1,2</sup>, Yixin Chen <sup>1,2</sup>, Na Luo <sup>1,2</sup>, Jialu Kang <sup>1,2</sup>, Xiaoqian Meng <sup>1,2</sup>, Yining Li <sup>1,2</sup>, Huirong Zhou <sup>1,2</sup>, Jiaxi Wang <sup>1,2</sup>, Yong Liu <sup>1,2</sup> and Guolei Li <sup>1,2,\*</sup>

- <sup>1</sup> Research Center of Deciduous Oaks, Beijing Forestry University, Beijing 100083, China
- <sup>2</sup> Key Laboratory for Silviculture and Conservation, Ministry of Education, Beijing Forestry University, Beijing 100083, China
- <sup>3</sup> College of Life and Chemistry Sciences, Zhejiang Normal University, Jinhua 321000, China
- \* Correspondence: glli226@163.com
- + These authors contributed equally to this work.

Abstract: Oaks (*Quercus* spp.) are considered model plants for studying plant evolution and natural gene introgression. Beijing area is at the junction between Taihang Mountain and Yanshan Mountain, and it is an overlapping distribution area of several deciduous oaks native to China. Interspecific hybridization often occurs in sympatric sibling species, resulting in blurred interspecific boundaries and hindering the development of breeding. To have better improvement and protection of these oaks, it is urgently necessary to evaluate the genetic diversity and population structure of these oak species. In this study, we collected eleven populations of four oak species (*Q. variabilis, Q. mongolica, Q. dentata* and *Q. aliena*) in the Beijing area. By using the polymorphic SSR markers, we analyzed the genetic variation of the collected 400 individuals, investigated the population structure, and found gene introgression events. *Q. variabilis* had a clearer genetic background as compared to the other three species. *Q. mongolica* had a more frequent gene introgression with *Q. dentata* and *Q. aliena*. Our data reveal the introgression and genetic diversity of the four oak species native to the Beijing area, laying the foundation for conducting an oak germplasm nursery with a clear genetic background for protection and further improvement.

Keywords: oak; population; genetic variation; SSR; DNA fingerprint

# 1. Introduction

Oaks (Quercus spp.) are widespread all over the world. Quercus is a large genus in Fagaceae and oaks have been cultivated for a long time, especially in Asia, Europe, and America, for their longevity and renewability [1]. There are over 450 Quercus species in the world and 51 in China. Oaks spread over south and north of China. In Beijing, the junction between Taihang Mountain and Yanshan Mountain, several deciduous oaks are natively distributed. Q. variabilis, Q. mogolica, Q. dentata, and Q. aliena are the most widely distributed oaks in north of China and have dominated the industrialization and utilization of ecological and economic plants. These four natively oaks are abundant, extensive, and overlapping at the junction between Taihang Mountain and Yanshan Mountain, and some of them are difficult to identify due to interspecific hybridization. This area is the north border for Q. variabilis and Q. aliena natural distribution and almost the south border for *Q. mongolica* natural distribution in China (Figure S1). Oak species have been formalized into two subgenera: subgenus Quercus, comprising Ponticae, Protobalanus, Virentes, Lobatae and Quercus; and subgenus Cerris, encompassing sections Ilex, Cerris, and Cyclobalanopsis [2-4]. Among the native oaks in Beijing area, Q. variabilis belongs to Cerris section East Asia Cerris with slender leaves, while Q. mogolica, Q. dentata, and Q. aliena belong to Quercus section Roburoids with oval and irregular leaves [4].



Citation: Pei, Z.; Yang, Q.; Chen, X.; Zong, Y.; Li, J.; Yang, X.; Huo, C.; Chen, Y.; Luo, N.; Kang, J.; et al. Population Study Reveals Genetic Variation and Introgression of Four Deciduous Oaks at the Junction between Taihang Mountain and Yanshan Mountain. *Forests* **2022**, *13*, 1647. https://doi.org/10.3390/ f13101647

Academic Editors: María-Dolores Rey, Jesus V. Jorrin Novo and María Ángeles Castillejo

Received: 7 September 2022 Accepted: 6 October 2022 Published: 8 October 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/).

Oaks being widely spread has been described as "the evolutionary success". One of the explanations of this success is their propensity for hybridization, contributing to adaptive introgression [5], which is increasingly recognized as an important process across diverse lineages of plant [6,7]. Studies have shown that interspecific hybridization and gene introgression frequently happen among species of deciduous oaks. However, the introgression makes it difficult to figure out the differences between two species with close relationship, leading to uncertainty when collecting the germplasm resources, especially for oak species. Studies on gene introgression among Quercus species have been carried out for decades. The hybridization and introgression have been clearly found in oak species widely distributed in Europe and North America, such as Q. rubra, Q. suber, Q. alba, and Q. robur [7–9]. A study in Japan also demonstrated the interspecific hybridization of *Q. mongolica* and *Q. dentata* [10]. Zeng et al. [11] found that *Q. mongolica* and *Q. liaotungensis* had a high degree of gene introgression in the Beijing area, leading to unclear species morphological characterization. Interspecific hybridization results in gene flow of each species, forming a complex genetic background. The origin of the individual could not be identified. However, the gene introgression among Q. variabilis, Q. mongolica, Q. dentata, and *Q. aliena* is very complex, and the genetic background of these oak species in the Beijing area is still unknown. The ambiguity of genetic background also leads to the inaccuracy of some existing studies in representing a particular oak species.

DNA molecular markers, especially SSR (Simple Sequence Repeat) markers, have been developed for many years and are widely used in recognition of species or cultivars [12–17]. In coconut, 74 polymorphic SSR markers were identified by genotyping-by-sequencing (GBS) analysis of 40 coconut accessions. Different traits could be divided into several groups, which was consistent with the amplicons of SSR markers [18]. Recently, SSR markers have been used in the identification of tree peony varieties, and the 32 phenotypes were divided into three main clusters by 31 SSR markers [19]. In oaks, although SSR markers have been used to characterize the populations worldwide for many years [20–23], genetic resources of native oaks can be found at the junction between Taihang Mountain and Yanshan Mountain, a specific region.

DNA fingerprints can be used to reveal and compare the relationship among individual oak within a species or different species, and can be effectively applied to genetic analysis and breeding [24,25]. We can generate the DNA fingerprint of an individual oak in those areas by using SSR markers to obtain exclusive identity information. In peas, DNA fingerprint and ISSR marker analysis indicated that peas distributed in two places had a tendency of differentiation [26]. Additionally, modern fingerprinting is crucial for varietal protection and germplasm characterization, even for protecting plant breeders' rights. Five microsatellite markers were used to distinguish different varieties of rice, ultimately solving the problem of mixing of two rice cultivars [27]. As there is no oak cultivar native to China, germplasm is very important for plant breeding. High-quality cultivars are the prerequisite for efficient breeding. In order to understand the genetic background and make use of resources for seedling breeding and grafting, it is beneficial to investigate the germplasm resources and obtain the genetic diversity of the four oak species.

Here, we collected 400 individuals of four oak species, including individuals both in natural and artificial forests at the junction between Taihang Mountain and Yanshan Mountain. By using nine SSR markers that were previously developed with high polymorphism, we analyzed the genetic variation among four oak species (*Q. variabilis*, *Q. aliena*, *Q. dentata*, and *Q. mongolica*) and the genetic diversity within one species from different region. By giving fingerprints of these oak individuals, we provided valuable data for further germplasm resource collecting and better improvement and protection of these native oaks. Oak individuals with clear genetic background provide an experimental basis for subsequent molecular breeding and other biological and genetic studies of oaks native to China.

## 3 of 13

# 2. Materials and Methods

#### 2.1. Sample Distribution and Plant Materials

Beijing is located in a region of widespread native oaks (Supplementary Figure S1) [28]. Briefly, in this study, we chose 6 representative regions for sampling in Beijing, which have abundant oak resources (see Supplementary Table S2 for details). The direct distance between any two of the sampling regions is at least 20 km. 10 ha quadrat was investigated in each region, in which oak species without serious pests and diseases were selected. A total of 400 samples were collected in this study, including 136, 74, 77, and 113 individuals from *Q. variabilis*, *Q. aliena*, *Q.dentata*, and *Q. mongolica*, respectively.

Sampling sites are mostly forest parks rich in forest resources. Except for the *Q. variabilis* and *Q. aliena* populations in Jiufeng (JF), other populations are natural populations. Two kinds of populations were selected according to the previous speculation that the genetic diversity was low in artificial populations and high in natural populations. These areas are the north border of the North China Plain and belong to the warm temperate semi-humid continental monsoon climate zone, with four distinct seasons and abundant rainfalls.

According to the earlier investigation on the spot, all individuals in good growth condition were labeled and numbered. After photographing and recording the position, we collected young leaves or dormant buds. All the samples were frozen in liquid nitrogen and stored at -80 °C until they were used for DNA extraction and isolation. Numbers of samples in each region are shown in Supplementary Table S2.

# 2.2. DNA Isolation and SSR Amplification

Genomic DNA of 400 samples were extracted by using Plant DNA Isolation Mini Kit (Vazyme, Nanjing, China). User manual and modified cetyltrimethyl ammonium bromide method was also used, as previously described [29,30]. The DNA was eluted into 50  $\mu$ L Tris-EDTA buffer as stock solution. Subsequently, the quality and the concentration of the genomic DNA were quantified using 1% agarose gel electrophoresis and a NanoDrop 2000 spectrophotometer (Thermo Scientific, Waltham, MA, USA), respectively. The genomic DNA was diluted and separated to 20 ng/ $\mu$ L aliquots and stored at -20 °C for polymerase chain reactions (PCR) amplification. Nine polymorphic SSR primer pairs were selected as previously reported [31], namely QrZAG96, QrZAG102, QrZAG112, QrZAG7, Qden03011, Qden03021, Qden03032, Qden05011, and Qden05031. The information of primers is listed in Supplementary Table S3.

The nine pairs of verified SSR primers with M13 (-21) tail at their 5' end were used for PCR amplification, together with M13 (-21) tailed fluorescent primers (M13-FAM, M13-ROX, M13-HEX) as previously described [30,32]. PCRs were conducted in 15  $\mu$ L reaction volumes containing 1  $\mu$ L of genomic DNA (20 ng/ $\mu$ L), 7.5  $\mu$ L of 2× Rapid Taq Master Mix (Vazyme), 0.4  $\mu$ L of fluorescent primer, 0.1  $\mu$ L of forward SSR primer, 0.5  $\mu$ L of reverse primer, and 5.5  $\mu$ L of ddH<sub>2</sub>O. PCR procedures were as follows: 95 °C (5 min), 32 cycles at 95 °C (30 s)/55–59 °C (30 s) (depending on the Tm value of SSR primers)/72 °C (30 s), followed by 8 cycles 95 °C (30 s)/53 °C (30 s)/72 °C (15 s) and a final extension at 72 °C for 5 min. PCR products were sent to Ruibo BioTech (Beijing, China) for capillary fluorescence electrophoresis detection to read the length of the PCR products.

# 2.3. Data Analysis

# 2.3.1. Characteristics and Genetic Diversity of SSR Primers

GeneMarker 2.2.0 software was used to read the polymorphic loci and obtain the genotypes of all samples. Linkage disequilibrium analysis was tested for all locus pairs in each population by randomization using Arlequin 3.1 [33]. Polymorphism parameters of 9 pairs of SSR markers were calculated by GenAlEx6.5 [34], respectively, by the number of alleles (Na), effective alleles (Ne), Shannon's information index (I), observed heterozygosity (Ho), expected heterozygosity (He), Fixation Index (F), and so on.

#### 2.3.2. Genetic Differentiation among Oak Populations and Optimal Population Classification

The analysis of the population structure was assessed in STRUCTURE 2.3.4 [35] using a Bayesian clustering approach setting parameters with a burn-in period of 100,000 iterations and 100,000 MCMC iterations after burn-in. In this study, we used "admixture model" and "allele frequencies correlated" during the modeling process. This approach revealed genetic structure by assigning individuals or predefined groups to K clusters. Different K values that ranged from 1 to 10 were used to infer the number of clusters for 10 replicate runs.  $\Delta K$ ( $\Delta K = mean (|L''(K)|)/sd[L(K)]$ ) and lnPPK were shown in STRUCUTRE Harvester [36], using the results of STRUCTURE.

Different results could be produced in the same conditions by different 10 replicate cluster analyses of the same data. Thus, CLUMPP 1.1.2 [37] and R 4.2.0 were used to average these 10 replicate results generated from STRUCTURE data and visualize the outcomes separately. Proportion of gene pools distributed in 10 oaks populations were plotted on maps in the form of pie charts. Depending on the specific coordinates of the sampling sites, pie charts of different populations were plotted on the topographic map. The clustering tree of both individuals (NJ) and populations (UPGMA) were made using PowerMarker 3.25 [38]. Trees were then colored and modified using Adobe Illustrator CC 2018.

#### 2.3.3. Genetic Diversity of 11 Oak Population and PCoA Analysis

Na, Ne, Ho, He, and fixation index (F), Fst, and gene flow (Nm) among 11 populations were calculated by GenAlEx6.5. The software was used to conduct the PCoA of the four oak species and analyze the length range of the amplified alleles among different species. According to the length range of the amplified alleles, the histogram was shown using SigmaPlot 12.5.

## 3. Results

# 3.1. Polymorphism of the Chosen SSR Primers

As previous studies have validated that the nine pairs of SSR markers have high polymorphism and stability and identify oak species [39,40], we chose these SSR markers to identify the genetic background of the four oak species (*Q. variabilis, Q. mongolica, Q. dentata,* and *Q. aliena*). Meanwhile, the results of LD analysis showed that nine locus were in linkage equilibrium. First, we collected 400 individual samples (including 136 of *Q. variabilis,* 113 of *Q. mongolica,* 77 of *Q. dentata,* and 74 of *Q. aliena*) from six regions of Beijing (see Supplementary Table S2 for details). We used these 400 samples to validate the polymorphism of the nine SSR markers. The nine SSR markers amplified 215 alleles in total, with an average of 23.9 alleles per locus (Supplementary Table S1). The observed heterozygosity (Ho) ranged from 0.45 to 0.82, with an average of 0.61. For all the observed loci, the Ho value was lower than expected (He), which ranged from 0.81 to 0.96 with an average of 0.87 (Supplementary Table S1).

# 3.2. Genetic Variation within Oak Species and Populations

We further analyzed the genetic variation of populations of different oak species from different regions by using SSR amplification results. Overall, *Q. variabilis* has the lowest Ho and He value among the four species, while *Q. mongolica* has the highest Ho and He value (Table 1). In *Q. variabilis*, the number of alleles per locus (Na) ranged from 6.56 to 10.89 among different populations in different regions, and the average Na was 8.61. *Q. variabilis* populations on Dayang Mountain (DY-Qv) and Zhoukoudian (ZK-Qv) have lower Ho compared to those on Jiufeng (JF-Qv) and Shangfang Mountain (SF-Qv), with the highest inbreeding coefficient (F = 0.26 and 0.31).

Species	Populations <sup>1</sup>	Na	Ne	Ho	He	F
Q. variabilis	JF-Qv	6.89	4.19	0.59	0.66	0.04
	ZK-Qv	10.11	4.92	0.50	0.71	0.31
	SF-Qv	6.56	4.59	0.56	0.67	0.18
	DY-Qv	10.89	4.64	0.49	0.69	0.26
	Qv-mean	8.61	4.59	0.54	0.68	0.20
Q. aliena	JF-Qa	10.11	5.90	0.59	0.73	0.19
	SF-Qa	12.11	6.49	0.61	0.76	0.21
	Qa-mean	11.11	6.19	0.60	0.75	0.20
Q. dentata	SF-Qd	6.22	4.77	0.61	0.76	0.18
	ZK-Qd	3.33	2.78	0.52	0.59	0.07
	DY-Qd	13.11	6.19	0.69	0.81	0.14
	Qd-mean	7.56	4.58	0.61	0.72	0.13
Q. mongolica	YM-Qm	16.89	8.28	0.67	0.86	0.22
	BH-Qm	12.67	8.15	0.74	0.86	0.14
	Qm-mean	14.78	8.21	0.70	0.86	0.18

Table 1. Genetic variation of different oak populations.

<sup>1</sup> The name of population is made up of regions and species, e.g., JF-Qv means *Q. variabilis* samples in Jiufeng.

In the SF-Qa population, Na, Ne, Ho, and He values were higher than those in JF-Qa population, while they have a similar sample amount (35 and 39 in SF and JF, respectively, Supplementary Table S2). The inbreeding coefficient of SF-Qa population is lower than that in JF-Qa population (Table 1), suggesting the higher genetic diversity of *Q. aliena* populations in Shangfang Mountain where natural populations are distributed. Similar patterns have been found in populations of *Q. variabilis*. Therefore, this result indicated that the genetic diversity of the artificial population was lower than that of the natural population.

For *Q. mongolica*, BH-Qm population has a higher Ho value and much lower inbreeding coefficient as compared to YM-Qm population, indicating that *Q. mongolica* in Baihua Mountain has abundant genetic diversity.

Next, we calculated the genetic differentiation coefficient (Fst) and gene flow (Nm) among populations within the same oak species in different region. The results showed that *Q. dentata* had the highest genetic differentiation coefficient and lowest gene flow among populations from Shangfang Mountain, Dayang Mountain, and Zhoukoudian (Table 2). *Q. mongolica* had the lowest Fst (0.01), indicating that only 1% variation happened between YM-Qm population and BH-Qm population.

Species	Regions	Fst <sup>1</sup>	Nm
Q. variabilis	JF SF DY ZK	0.04	6.16
Q. aliena	JF SF	0.02	12.25
Q. dentata	SF DY ZK	0.08	2.80
Q. mongolica	YM BH	0.01	17.61

Table 2. Genetic variation among populations within the same oak species in different regions.

<sup>1</sup> Fst and Nm are the parameters of genetic variation. Fst: genetic differentiation coefficient. Nm: gene flow.

#### 3.3. Population Structure and Gene Introgression

A neighbor-joining tree was conducted to show the genetic relationship among all the observed individuals. Most of the individuals of *Q. variabilis* had the longest genetic distance with the other three species. Notably, a few individuals of *Q. variabilis* were grouped with *Q. dentata* and *Q. mongolica* (Figure 1), indicating that possible gene introgression happened among species. Meanwhile, *Q. dentata*, *Q. aliena*, and *Q. mongolica* had more individuals grouped with each other, which suggested that gene introgression happened more frequently among these three species, confirming that these three species had a closer relationship as compared to *Q. variabilis*.



**Figure 1.** Neighbor-joining tree of 400 oak individuals from four species. Different colors indicate different species. The cluster of *Q. variabilis* (green), *Q. aliena* (blue), *Q. dentata* (cyan), and *Q. mongolica* (orange) is shown in the circle NJ tree. Length of lines show the genetic distance. The scale bar represents 0.1 genetic distance.

The population structure of 400 oak individuals was also confirmed by principal coordinates analysis (PCoA). We firstly used all the samples to do the PcoA analysis. The results showed that the first coordinate separated *Q. variabilis* and the other three species, while the other three species were grouped together (Figure 2a). In addition, PCoA analysis by using the samples for each species in different regions could not generate the significant group (Supplementary Figure S2). The results show that there was no distinct genetic differentiation among the different populations from different regions in the Beijing area for each oak species.

The 400 oak individuals were evaluated for population stratification. The assumed number of clusters was set from K = 1 to 10. The Bayesian analysis by STRUCUTRE Harvester showed that the individuals were suitable to be divided into four groups (K = 4) (Supplementary Figure S3). The STRUCTURE results showed that *Q. variabilis* (green) has a relatively clear genetic background (Figure 2b). Further, four native oaks species were evaluated for population stratification, respectively. The assumed number of clusters

was set from K = 2 to 5 [41]. The STRUCTURE results showed that each species has a low degree of differentiation among different sampling region, especially *Q. dentata* and *Q.mongolica* (Supplementary Figure S4), with high Nm values (Table 2). This result was consistent with PCoA analysis and phylogenetic tree (Figures 1, 2 and S2). *Q. mongolica* had abundant genetic diversities and gene introgression individuals from *Q. aliena* and *Q. dentata*, indicating that *Q. mongolica*, had a closer genetic relationship with *Q. aliena* and *Q. dentata* as compared to *Q. variabilis*, as we could also see the population structure when K = 2 and K = 3 (Supplementary Figure S5).



**Figure 2.** PCoA and population structure for all the 400 individuals. (**a**) PCoA groups of four oak species are calculated and plotted by GenAlEx6.5. Different colors indicate different species, and each point represents an individual. (**b**) Population structure based on Bayesian clustering approaches with K = 4 by STRUCTURE software. Each bar represents a single individual and different colors indicate different subpopulations. The individual species are shown at the top of the bars.

The regions where we collected samples are located on the edge of the North China Plain (Figure 3), and they are near the borders of natural distribution of *Q. mongolica* and *Q. aliena* (Supplementary Figures S1 and S6). The pie charts indicated a proportion of gene pools distributed in 10 oaks populations, of which *Q. mongolica* is the most complex.

We then used the unweighted pair-group method and the arithmetic (UPGMA) algorithm to build the phylogenetic tree of populations of four species from different regions in the Beijing area. Populations from one species were grouped together and *Q. mongolica* had a closer genetic relationship with *Q. aliena* than that with *Q. dentata* (Figure 4). Notably, JF-Qv population, which was an artificial population, had a further genetic distance with other three *Q. variabilis* populations, indicating that the individuals might be transplanted from another area outside Beijing.



**Figure 3.** Population distribution map and genetic background of all samples. Different colors indicate the different genetic clusters in STRUCTURE analysis.



**Figure 4.** The UPGMA analysis of populations from different regions. Eleven populations mentioned above are evaluated. The scale bar represents 0.05 genetic variation. The abbreviations of the population names are listed in Table S2.

## 3.4. DNA Fingerprints of Oak Species in Beijing Area

As there was no oak germplasm nursery with clear genetic background in China, we provided the fingerprints of the observed 400 oak individuals. The DNA fingerprints of these 400 individuals were represented with numbers that showed the allele length and letters that indicated the amplified SSR alleles (Supplementary Table S4). The length of amplified alleles varied among different species. Qden03021 had a large range of amplicons in *Q. variabilis* as compared to the other species, while Qden05031 had a large range of amplicons in *Q. mongolica* (Supplementary Figure S7), suggesting the disparities in polymorphism of different alleles in different oak species.

# 4. Discussion

Oaks are valuable in ecologic, economic, and research fields. They have proved to be a successfully evolved species in the world, with extensive radiation and expansion [5], and are ideal materials for us to study gene introgression and natural hybridization of woody plants [7,21,42]. Hipp et al. [3] used genotyping-by-sequencing and SNP markers to generate large-scale data, demonstrating frequent gene diversity and introgression of oak species in *Quercus* genus. We conducted a fine-scale study to reveal the genetic relationship of oak species at the junction between Taihang Mountain and Yanshan Mountain, taking a close look at the genetic diversity within and among species, and also observed the gene introgression among these oak species. We collected 400 individuals of oak species in the area which is the north border of North China Plain. In this area, several native oak species, including *Q. variabilis*, *Q. aliena*, *Q. dentata*, and *Q. mongolica* are naturally distributed. As there is no improved cultivar for any of these oak species native to China, it is meaningful to detect the genetic variation of these individuals, preparing for germplasm nursery using accessions with a clear genetic background.

Previous studies have screened some SSR markers with high polymorphism for oak species [31,43]. We used these SSR markers and further confirmed the polymorphism of the picked 9 markers. For the individuals of the four oak species in Beijing area, these SSR markers exhibit high polymorphism, and 215 alleles were detected in total, 23.9 on average (Supplementary Table S1), much higher than the studies in Q. petraea [12] and similar to the studies on *Q. mongolica* and *Q. liaotungensis*, which collected 1166 individuals [11]. This result indicates that these SSR markers are efficient for oak species identification [43]. Range size shows some relationship to heterozygosity and allelic richness [44], in that Q. mongolica had the highest heterozygosity (He = 0.86) and allelic richness (Na = 14.78), suggesting that *Q. mongolica* had the most widespread distribution among the four species. Naturally, the phylogenetically close relationship between Q. dentata and Q. aliena with Q. mongolica could be well explained. Whether they originated and evolved in Q. mongolica remains unconfirmed. Both total alleles and expected heterozygosity (He) were lower than the studies in *Q. variabilis* on a national scale [44], speculating that it might due to artificial population. It is interesting that the inbreeding coefficient (F) of the artificially planted Q. variabilis population (JF-Qv) was 0.04 (Table 1), suggesting that the individuals probably came from random mating populations. Meanwhile, the UPGMA analysis showed that JF-Qv artificial population had a phylogenetically distant relationship with other Q. variabilis populations in this area (Figure 4), indicating that the trees in Jiufeng might be transplanted from other areas outside Beijing. Thus, the sampled individuals might come from a different provenance. In contrast, the other artificial population, Q. aliena population in Jiufeng (JF-Qa), had a strong gene flow (Nm = 12.25) with the natural population in Shangfang Mountain (SF-Qa) (Table 2), suggesting that some individuals in JF-Qa population might be transplanted from Shangfang Mountain. Gene flow determines the genetic structure and survival potential of future populations of a species [45], and the previous study reported a lower gene flow (Nm = 3.648) among the population of Q. variabilis than our study [46], which objectively reflected that Q. variabilis distributed at the junction between Taihang Mountain and Yanshan Mountain had more survival potential. In addition, SF-Qa natural population has more abundant genetic diversities than JF-Qa artificial population (Table 1), further proving that JF-Qa might be generated from the same provenance.

The phylogenetic tree, combining UPGMA analysis and STRUCTURE analysis, showed that *Q. variabilis* evolved separately from the other three species (Figures 1, 2 and 4). This finding is consistent with previous studies using SNP markers and limited individuals within one species [3]. As Hipp et al. [3] reported, *Q. mongolica, Q. aliena*, and *Q. dentata* belong to *Quercus* section *Roburoids*, while *Q. variabilis* belongs to *Cerris* section *East Asia Cerris*, together with *Q. accutisima*, and *Q. dentata* evolves separately from *Q. mongolica* and *Q. aliena*. The introgression of *Q. variabilis* and *Q. accutisima* has already been confirmed by using high-quality genomic resources [47]. Our results also demonstrate that *Q. mongolica* and *Q. aliena* have a closer genetic relationship as compared to *Q. dentata* 

(Figures 1, 2 and 4). The gene flow between BH-Qm and YM-Qm reached 17.61, much higher than the other three species. Considering the geographical distance between Baihua Mountain and Yunmeng Mountain, the two Q. mongolica populations might have evolved together and separated recently. As many previous studies have demonstrated that gene introgression happens frequently in Quercus genus [7,9,21,48], it could not be ignored when analyzing the oak population in Beijing area. The STRUCTURE analysis and the phylogenetic tree showed that gene introgression individuals were grouped with corresponding species. Q. mongolica has the most individuals that have gene introgression with Q. dentata and Q. aliena (Figures 1 and 2). Q. variabilis has the least gene introgression individuals, suggesting that Q. variabilis has a phylogenetically distant relationship with the other three species. A possible contributor is the lower successful rate of interspecific hybridization with the other three species in natural conditions. This result provides guidance for distant hybridization breeding of *Q. variabilis*, which should take the hybridization compatibility into consideration. In addition, as previously reported, *Q. variabilis* has a distinct flowering time compared to the other three species [49], which might also create less chances for interspecific hybridization [50]. Previous studies on Q. dentata and Q. aliena showed that the male flowers of *Q. dentata* and the female flowers of *Q. aliena* in this area had a certain overlap [51], suggesting that these two oaks had a frequent outcrossing.

The junction between Taihang Mountain and Yanshan Mountain is not only near the north border of the natural distribution of Q. variabilis and Q. aliena but also near the south border of the natural distribution of *Q. mongolica*. *Q. mongolica* distributed on the mountains with an altitude over 600 m. The north part of this distribution area is high while the south part is low, topographically, with a warm temperate sub-humid continental monsoon climate. Considering the pollen dispersal of oak species and the northwest wind direction in spring at the junction between Taihang Mountain and Yanshan Mountain, it could be explained that southwest of Beijing had abundant species diversities of deciduous oaks, which may also explain the warmer climate (Figure 3). Fine-scale spatial genetic structure of the oak species could be further studied in the Beijing area to further identify the interspecific hybrids and their adaption to the environment [40,52,53]. During our investigation on the spot, we observed that the leaf shape of Q. dentata was similar to *Q. mongolica* (Supplementary Figure S8) but quite different from the leaf of *Q. variabilis*. Further studies could use SNP markers to link the morphological traits or environmental adaptions (e.g., temperature, drought, salt, etc.), and the genotype and more oak species native to China should be identified by molecular markers. With a clearly identified genetic background, we could use these genetic resources for breeding.

DNA fingerprints could give the putative cultivar an identity, which is very important for intellectual property protection. Here, we provided all the DNA fingerprints of 400 samples generated from SSR data (Supplementary Table S4). Further study could be done to get the link between DNA fingerprints and their morphological traits by using more SSR markers or SNP markers like studies in some other species [18,19], making it clearer to focus on some specific trait breeding.

In conclusion, we detected the genetic variation of natural and artificial populations of four oak species native to China and confirmed the genetic relationship and gene introgression of *Q. variabilis*, *Q. mongolica*, *Q. aliena*, and *Q. dentata*. Our work has laid a foundation for collecting oak germplasm and provides valuable data for further protection and improvement. In addition, we screened some individuals with a clear and simple genetic background for whole genome sequencing and cross-breeding.

**Supplementary Materials:** The following supporting information can be downloaded at: https: //www.mdpi.com/article/10.3390/f13101647/s1, Figure S1: Distribution map of *Q. variabilis* (a), *Q. aliena* (b), *Q. dentata* (c) and *Q. mongolica* (d) respectively in China. The red squares represent Beijing area. This figure is modified from Wang et al. [1]. Figure S2: PCoA analysis of different populations within the same species. Different color and shape of the dots indicate the different populations. Figure S3:  $\Delta K$  and lnPPK values are showed with blue dot. The values are calculated according to *K* = 1 to 10. Figure S4: STRUCTURE analysis of 4 native oaks when *K* = 2 to 5. *Q. variabilis* (a), *Q. aliena* (b), *Q. dentata* (c) and *Q. mongolica* (d) respectively. Figure S5: STRUCTURE analysis of populations when K = 2 and K = 3. Dark blue background shows *Q. variabilis* cluster. Figure S6: The sampling regions of four native oaks. *Q. variabilis* (a), *Q. aliena* (b), *Q. dentata* (c) and *Q. mongolica* (d) respectively. Figure S7: Length variety of the amplicons of each allele in different oak species. Green, blue, cyan and orange bars indicate length variety of each allele in *Q. variabilis* (a), *Q. aliena*, *Q. aliena*, *Q. dentata* and *Q. mongolica* respectively. Figure S8: Leaf surface and back of *Q. variabilis* (a), *Q. aliena* (b), *Q. dentata* (c) and *Q. mongolica* (d) respectively. Table S1: Main genetic parameters for the chosen 9 SSR primers. Table S2: Sample information. Table S3: Information of the 9 SSR primers. Table S4: DNA fingerprints of 400 samples.

**Author Contributions:** Q.Y. and G.L. designed the experiments. Z.P. and Q.Y. conducted most of the experiment, analyzed the data. Z.P. and Q.Y. wrote the manuscript. Z.P., X.C. and J.L. conducted DNA extraction and SSR PCR experiment. Y.Z. analyzed the data and revised the manuscript. Z.P., Q.Y., X.C., J.L., X.Y., C.H., Y.C., N.L., J.K., X.M., Y.L. (Yining Li), H.Z. and J.W. collected the samples. Y.L. (Yong Liu) and G.L. supervised the project and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Fundamental Research Funds for the Central Universities (2021ZY11), the China Postdoctoral Science Foundation (2021M700451).

**Data Availability Statement:** The data presented in this study are available on reasonable request from the corresponding author.

Acknowledgments: We thank Yan Wang for providing useful suggestions on SSR data analysis and manuscript preparing. We thank Kundong He for English language improvement.

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

# References

- Plomion, C.; Aury, J.-M.; Amselem, J.; Leroy, T.; Murat, F.; Duplessis, S.; Faye, S.; Francillonne, N.; Labadie, K.; Le Provost, G.; et al. Oak genome reveals facets of long lifespan. *Nat. Plants* 2018, *4*, 440–452. [CrossRef]
- 2. Liu, X.; Chang, E.; Liu, J.; Jiang, Z. Comparative analysis of the complete chloroplast genomes of six white oaks with high ecological amplitude in China. *J. For. Res.* **2021**, *32*, 2203–2218. [CrossRef]
- Hipp, A.L.; Manos, P.S.; Hahn, M.; Avishai, M.; Bodénès, C.; Cavender-Bares, J.; Crowl, A.A.; Deng, M.; Denk, T.; Fitz-Gibbon, S.; et al. Genomic landscape of the global oak phylogeny. *New Phytol.* 2020, 226, 1198–1212. [CrossRef] [PubMed]
- 4. Hubert, F.; Grimm, G.W.; Jousselin, E.; Berry, V.; Franc, A.; Kremer, A. Multiple nuclear genes stabilize the phylogenetic backbone of the genus *Quercus. Syst. Biodivers.* **2014**, *12*, 405–423. [CrossRef]
- 5. Kremer, A.; Hipp, A.L. Oaks: An evolutionary success story. New Phytol. 2020, 226, 987–1011. [CrossRef] [PubMed]
- Mir, C.; Toumi, L.; Jarne, P.; Sarda, V.; Di Giusto, F.; Lumaret, R. Endemic North African *Quercus afares* Pomel originates from hybridisation between two genetically very distant oak species (*Q. suber* L. and *Q. canariensis* Willd.): Evidence from nuclear and cytoplasmic markers. *Heredity* 2006, 96, 175–184. [CrossRef]
- Crowl, A.A.; Manos, P.S.; McVay, J.D.; Lemmon, A.R.; Lemmon, E.M.; Hipp, A.L. Uncovering the genomic signature of ancient introgression between white oak lineages (*Quercus*). New Phytol. 2020, 226, 1158–1170. [CrossRef] [PubMed]
- Löpez-Aljorna, A.; Bueno, M.A.; Aguinagalde, I.; Martín, J.P. Fingerprinting and genetic variability in cork oak (*Quercus suber* L.) elite trees using ISSR and SSR markers. *Ann. For. Sci.* 2007, 64, 773–779. [CrossRef]
- 9. Moran, E.V.; Willis, J.; Clark, J.S. Genetic evidence for hybridization in red oaks (*Quercus* sect. *Lobatae*, Fagaceae). *Am. J. Bot.* **2012**, 99, 92–100. [CrossRef]
- Nagamitsu, T.; Uchiyama, K.; Izuno, A.; Shimizu, H.; Nakanishi, A. Environment-dependent introgression from Quercus dentata to a coastal ecotype of *Quercus mongolica* var. *crispula* in northern Japan. *New Phytol.* 2020, 226, 1018–1028. [CrossRef] [PubMed]
- 11. Zeng, Y.-F.; Liao, W.-J.; Petit, R.J.; Zhang, D.-Y. Geographic variation in the structure of oak hybrid zones provides insights into the dynamics of speciation. *Mol. Ecol.* 2011, 20, 4995–5011. [CrossRef] [PubMed]
- 12. Lupini, A.; Aci, M.M.; Mauceri, A.; Luzzi, G.; Bagnato, S.; Menguzzato, G.; Mercati, F.; Sunseri, F. Genetic diversity in old populations of sessile oak from Calabria assessed by nuclear and chloroplast SSR. *J. Mt. Sci.* **2019**, *16*, 1111–1120. [CrossRef]
- Zurn, J.D.; Nyberg, A.; Montanari, S.; Postman, J.; Neale, D.; Bassil, N. A new SSR fingerprinting set and its comparison to existing SSR- and SNP-based genotyping platforms to manage *Pyrus germplasm* resources. *Tree Genet. Genomes* 2020, *16*, 1614–2942. [CrossRef]
- 14. Yücedağ, C.; Müller, M.; Gailing, O. Morphological and genetic variation in natural populations of *Quercus vulcanica* and *Q. frainetto. Plant Syst. Evol.* **2021**, 307, 8. [CrossRef]
- 15. Du, Q.; Pan, W.; Xu, B.; Li, B.; Zhang, D. Polymorphic simple sequence repeat (SSR) loci within cellulose synthase (*PtoCesA*) genes are associated with growth and wood properties in *Populus tomentosa*. *New Phytol.* **2013**, 197, 763–776. [CrossRef] [PubMed]

- Du, Q.; Wang, B.; Wei, Z.; Zhang, D.; Li, B. Genetic Diversity and Population Structure of Chinese White Poplar (*Populus tomentosa*) Revealed by SSR Markers. J. Hered. 2012, 103, 853–862. [CrossRef] [PubMed]
- 17. Wei, Z.; Du, Q.; Zhang, J.; Li, B.; Zhang, D. Genetic Diversity and Population Structure in Chinese Indigenous Poplar (*Populus simonii*) Populations Using Microsatellite Markers. *Plant Mol. Biol. Rep.* **2012**, *31*, 620–632. [CrossRef]
- Riangwong, K.; Wanchana, S.; Aesomnuk, W.; Saensuk, C.; Nubankoh, P.; Ruanjaichon, V.; Kraithong, T.; Toojinda, T.; Vanavichit, A.; Arikit, S. Mining and validation of novel genotyping-by-sequencing (GBS)-based simple sequence repeats (SSRs) and their application for the estimation of the genetic diversity and population structure of coconuts (*Cocos nucifera* L.) in Thailand. *Hortic. Res.* 2020, 7, 156. [CrossRef]
- 19. Luo, L.; Yang, Y.; Zhao, H.; Leng, P.; Hu, Z.; Wu, J.; Zhang, K. Development of EST-SSR markers and association analysis of floral scent in tree peony. *Sci. Hortic.* **2021**, *289*, 110409. [CrossRef]
- 20. Ramos-Ortiz, S.; Oyama, K.; Rodríguez-Correa, H.; González-Rodríguez, A. Geographic structure of genetic and phenotypic variation in the hybrid zone between *Quercus affinis* and *Q. laurina* in Mexico. *Plant Species Biol.* **2015**, *31*, 219–232. [CrossRef]
- An, M.; Deng, M.; Zheng, S.-S.; Jiang, X.-L.; Song, Y.-G. Introgression Threatens the Genetic Diversity of *Quercus austrocochinchinensis* (Fagaceae), an Endangered Oak: A Case Inferred by Molecular Markers. *Front. Plant Sci.* 2017, *8*, 229. [CrossRef] [PubMed]
- Yang, J.; Vázquez, L.; Chen, X.; Li, H.; Zhang, H.; Liu, Z.-L.; Zhao, G. Development of Chloroplast and Nuclear DNA Markers for Chinese Oaks (*Quercus* Subgenus *Quercus*) and Assessment of Their Utility as DNA Barcodes. *Front. Plant Sci.* 2017, *8*, 816. [CrossRef]
- 23. Burkardt, K.; Pettenkofer, T.; Ammer, C.; Gailing, O.; Leinemann, L.; Seidel, D.; Vor, T. Influence of heterozygosity and competition on morphological tree characteristics of *Quercus rubra* L.: A new single-tree based approach. *New For.* **2020**, *52*, 679–695. [CrossRef]
- Nybom, H. Applications of DNA Fingerprinting in Plant Breeding. In DNA Fingerprinting: Approaches and Applications; Birkhäuser: Basel, Switzerland, 1991; pp. 294–311.
- 25. Nybom, H.; Weising, K.; Rotter, B. DNA fingerprinting in botany: Past, present, future. *Investig. Genet.* **2014**, *5*, 1. [CrossRef] [PubMed]
- Stavridou, E.; Lagiotis, G.; Karapetsi, L.; Osathanunkul, M.; Madesis, P. DNA Fingerprinting and Species Identification Uncovers the Genetic Diversity of Katsouni Pea in the Greek Islands Amorgos and Schinoussa. *Plants* 2020, *9*, 479. [CrossRef]
- 27. Ganopoulos, I.; Argiriou, A.; Tsaftaris, A. Adulterations in Basmati rice detected quantitatively by combined use of microsatellite and fragrance typing with High Resolution Melting (HRM) analysis. *Food Chem.* **2011**, *129*, 652–659. [CrossRef]
- 28. Wang, L.M.; Ren, X.W.; Liu, Y.Q. The geographical distribution of deciduous oak in China. J. Beijing For. Univ. 1985, 2, 57–69.
- Yue, X.-Y.; Liu, G.-Q.; Zong, Y.; Teng, Y.-W.; Cai, D.-Y. Development of genic SSR markers from transcriptome sequencing of pear buds. J. Zhejiang Univ. Sci. B 2014, 15, 303–312. [CrossRef]
- Wang, Y.; Jia, H.-M.; Shen, Y.-T.; Zhao, H.-B.; Yang, Q.-S.; Zhu, C.-Q.; Sun, D.-L.; Wang, G.-Y.; Zhou, C.-C.; Jiao, Y.; et al. Construction of an anchoring SSR marker genetic linkage map and detection of a sex-linked region in two dioecious populations of red bayberry. *Hortic. Res.* 2020, 7, 53. [CrossRef]
- Kampfer, S.; Lexer, C.; Glössl, J.; Steinkellner, H. Characterization of (GA)n Microsatellite Loci from *Quercus Robur*. *Hereditas* 1998, 129, 183–186. [CrossRef]
- 32. Schuelke, M. An economic method for the fluorescent labeling of PCR fragments. Nat. Biotechnol. 2000, 18, 233–234. [CrossRef]
- 33. Excoffier, L.; Laval, G.; Schneider, S. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol. Bioinform.* **2005**, *1*, 47–50. [CrossRef]
- 34. Peakall, R.; Smouse, P.E. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research— An update. *Bioinformatics* **2012**, *28*, 2537–2539. [CrossRef]
- Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* 2000, 155, 945–959. [CrossRef] [PubMed]
- 36. Earl, D.A.; VonHoldt, B.M. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **2011**, *4*, 359–361. [CrossRef]
- 37. 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]
- 38. Liu, K.; Muse, S.V. PowerMarker: An integrated analysis environment for genetic marker analysis. *Bioinformatics* 2005, 21, 2128–2129. [CrossRef]
- 39. Wang, Y. Interspecific Hybridization of Quercus Dentata, *Q. mongolica-Q. wutaishansea* Based on SSR Markers. Master's Thesis, Shandong University, Jinan, China, 2012.
- Gao, J.P.; Qin, L.; Cao, Q.Q.; Fang, K.F.; Tian, Y.L. Genetic structure of populations and introgression of three *Quercus* species in mountainous area of Beijing. *J. Beijing For. Univ.* 2020, 42, 58–67. [CrossRef]
- 41. Neophytou, C.; Gärtner, S.M.; Vargas-Gaete, R.; Michiels, H.-G. Genetic variation of Central European oaks: Shaped by evolutionary factors and human intervention? *Tree Genet. Genomes* **2015**, *11*, 79. [CrossRef]
- 42. Curtu, A.L.; Gailing, O.; Finkeldey, R. Patterns of contemporary hybridization inferred from paternity analysis in a four-oakspecies forest. *BMC Evol. Biol.* 2009, *9*, 284–289. [CrossRef]
- 43. Guo, B. Construction of SSR fingerprint and research of genetic structure in relative *Quercus* species. *J. Beijing For. Univ.* **2018**, 40, 10–18. [CrossRef]

- 44. Spence, E.; Fant, J.; Gailing, O.; Griffith, M.; Havens, K.; Hipp, A.; Kadav, P.; Kramer, A.; Thompson, P.; Toppila, R.; et al. Comparing Genetic Diversity in Three Threatened Oaks. *Forests* **2021**, *12*, 561. [CrossRef]
- Hamrick, J.L.; Godt, M.J.W.; Sherman-Broyles, S.L. Factors influencing levels of genetic diversity in woody plant species. In Population Genetics of Forest Trees; Springer: Berlin, Germany, 1992; pp. 95–124.
- Shi, X.; Wen, Q.; Cao, M.; Guo, X.; Xu, L.A. Genetic Diversity and Structure of Natural *Quercus variabilis* Population in China as Revealed by Microsatellites Markers. *Forests* 2017, *8*, 495. [CrossRef]
- 47. Fu, R.; Zhu, Y.; Liu, Y.; Feng, Y.; Lu, R.-S.; Li, Y.; Li, P.; Kremer, A.; Lascoux, M.; Chen, J. Genome-wide analyses of introgression between two sympatric Asian oak species. *Nat. Ecol. Evol.* **2022**, *6*, 924–935. [CrossRef] [PubMed]
- Manos, P.; Hipp, A. An Updated Infrageneric Classification of the North American Oaks (*Quercus* Subgenus *Quercus*): Review of the Contribution of Phylogenomic Data to Biogeography and Species Diversity. *Forests* 2021, 12, 786. [CrossRef]
- 49. Liu, P. Studies on Hybridization and Seed Dispersal of Three *Quercus* Species in Beijing Area. Master's thesis, Beijing Forestry University, Beijing, China, 2020.
- Schermer, E.; Bel-Venner, M.; Gaillard, J.; Dray, S.; Boulanger, V.; Le Roncé, I.; Oliver, G.; Chuine, I.; Delzon, S.; Venner, S. Flower phenology as a disruptor of the fruiting dynamics in temperate oak species. *New Phytol.* 2020, 225, 1181–1192. [CrossRef]
- Liu, Y.; Li, Y.; Song, J.; Zhang, R.; Yan, Y.; Wang, Y.; Du, F.K. Geometric morphometric analyses of leaf shapes in two sympatric Chinese oaks: Quercus dentata Thunberg and *Quercus aliena* Blume (Fagaceae). *Ann. For. Sci.* 2018, 75, 90. [CrossRef]
- 52. Carabaña, M.V.; Gonzalez-Martinez, S.C.; Hardy, O.; Gil, L. Fine-scale spatial genetic structure in mixed oak stands with different levels of hybridization. *Mol. Ecol.* 2007, *16*, 1207–1219. [CrossRef]
- Curtu, A.; Craciunesc, I.; Enescu, C.; Vidalis, A.; Sofletea, N. Fine-scale spatial genetic structure in a multi-oak-species (*Quercus* spp.) forest. *Iforest-Biogeosci. For.* 2014, *8*, 324–332. [CrossRef]