forests

MDPI

Article

Revealing the Genetic Structure and Differentiation in Endangered *Pinus bungeana* by Genome-Wide SNP Markers

Qian Tian [1,2], Yousry A. El-Kassaby [3] and Wei Li [1,2,*]

[1] National Engineering Laboratory of Tree Breeding, Key Laboratory of Genetics and Breeding in Forest Trees and Ornamental Plants of Ministry of Education, College of Biological Sciences and Technology, Beijing Forestry University, Beijing 100083, China; tq_ian@163.com
[2] The Tree and Ornamental Plant Breeding and Biotechnology Laboratory of National Forestry and Grassland Administration, College of Biological Sciences and Technology, Beijing Forestry University, Beijing 100083, China
[3] Department of Forest and Conservation Sciences, Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, BC V6T 1Z4, Canada; y.el-kassaby@ubc.ca
* Correspondence: bjfuliwei@bjfu.edu.cn

Abstract: Understanding the genetic variation and differentiation of natural populations is essential for their protection, specifically if the species status is endangered as with *Pinus bungeana*. Here, we used 346,840 high density and strong specificity SNP loci to carry out genetic analyses (i.e., genetic diversity, genetic structure, phylogeny, and geographical differentiation) on 52 *P. bungeana* individuals from 5 populations (4 natural and one artificial) representing the main regions of the species distribution in China. Genetic diversity assessment indicated a trend of genetic diversity gradual decrease from west to east across the species distribution areas. Population genetic structure, PCA and phylogenetic analyses consistently indicated that populations in the central and eastern regions were clustered together, while those from the western regions were separated. Mantel test values indicated the presence of geographic isolation among populations, an important factor contributing to the observed genetic differentiation. The maximum likelihood tree and potential migration events inferred from TreeMix analysis indicated the presence of historical genetic exchanges between the west of Qinling Mountains and the Lvliang Mountains populations. Based on the generated genetic information, in situ and ex situ conservation strategies for *P. bungeana* germplasm resources are proposed, these strategies could be valuable for the conservation, protection and genetic improvement of this endangered species.

Keywords: genetic diversity; geographical differentiation; genetic resources conservation; single nucleotide polymorphism

Citation: Tian, Q.; El-Kassaby, Y.A.; Li, W. Revealing the Genetic Structure and Differentiation in Endangered *Pinus bungeana* by Genome-Wide SNP Markers. *Forests* **2022**, *13*, 326. https://doi.org/10.3390/f13020326

Academic Editor: Evi Alizoti

Received: 1 December 2021
Accepted: 31 January 2022
Published: 17 February 2022

1. Introduction

Pinus bungeana Zucc. ex Endl., Pinaceae, is an economically and ecologically valuable conifer tree species native to central and western China [1]. The geographic position of its natural distribution range is between 30°52′ and 38°15′ N and from 104°15′ to 113°15′ E, and the altitude range is about 600–1850 m [2]. The species has appealing attributes (e.g., beautiful tree shape, dense crown, and light green or white irregular peeling bark) that made it distinctive among pine species and has gained high ornamental and urban design/gardening values. However, the impact of human disturbance and the characteristics of this species (e.g., slow growth, lack of competitive advantage for seedlings and difficulty of regeneration) have led to a sharp reduction in the number of natural populations of this species, and resulted in placing it on the Threatened Species List of China's Higher Plants, and its protection level is Endangered (EN) [3]. Among the existing natural populations of *P. bungeana*, only the populations in the southeast of Gansu province, the southwest of Shanxi province, and central and western regions of Shaanxi province are

concentrated, with relatively large quantity and high diversity [4]. The rest of the areas that the species is distributed in are sporadically covered by small numbers of individuals [5], especially the mountainous areas on the southern edge of the distribution areas, where *P. bungeana* populations have been difficult to find [6].

The creation of small fragmented populations was mainly due to the species' substantial harvest in recent years [7], thus reducing gene flow and increasing genetic erosion, ultimately affecting its adaptive potential needed for dealing with environmental contingencies and long-term survival [8]. Compared with other pines, the average genetic diversity of *P. bungeana* natural populations is relatively low, but the degree of differentiation among populations is relatively high [4,9,10]. Natural populations may produce unique or rare genes in the process of adapting to different ecological environments, which is generally of importance for conserving gene diversity of unknown value and potential use. Therefore, it is necessary to protect the genetic diversity of *P. bungeana* natural populations.

Assessing and understanding the biological and ecological factors affecting the distribution of genetic diversity among and within a species' populations over time and space are essential for the successful implementation of management and conservation strategies, particularly for those most vulnerable [11]. Formulating sound germplasm conservation strategies, the selection of in situ conservation sites, and the sampling strategies for ex situ conservation collections require intimate knowledge of species genetic structure for use as baseline references [12,13].

Molecular markers are effective in studying species' genetic diversity and populations structure and differentiation. *P. bungeana* populations' genetic diversity and differentiation have been studied using Amplified Fragment Length Polymorphisms (AFLP) [14], microsatellites (simple sequence repeats: SSR) [10,15], and Start Codon Targeted (SCoT) [16] polymorphisms; however, these studies used a limited number of markers with a sparse genomic density; thus, they did not reveal the studied populations' comprehensive genetic history. Single Nucleotide Polymorphism (SNPs), are the most abundant form of variation at the DNA level, and can comprehensively reflect individuals' and populations' genetic variation [17]. In particular, several population structure studies have indicated that SNPs have a higher resolving power than SSRs, especially when small sample sizes were used or when the overall level of genetic differentiation is low [18–21]. Du et al. [22] used SNPs to investigate the genetic structure and phylogenetic relationships of *Crataegus* (a genus of several hundred species of shrubs and trees) and suggested an evolutionary model for the Chinese main species. Similarly, Rodger et al. [23] assessed patterns of genetic structure and differentiation of the natural populations of an endangered daisy (*Rutidosis leptorrhynchoides*) using SNP data, and defined populations for their conservation management planning. *P. bungeana* has a complex and huge genome. The genome-wide SNPs data can be an effective tool for studying the genetics of this species. Yet, *P. bungeana* population genetic variation revealed by SNP markers has not been investigated.

Here, we used a large number of SNP loci developed based on SLAF-seq to carry out genetic analysis on *P. bungeana* from 5 populations representing the species different regions in China. We estimated the extent and distribution of genetic diversity among these small scattered populations, determined their phylogenetic relationship and geographical differentiation, and inferred the potential historical migration events among populations. According to the genetic information obtained from the analysis, we have formulated an effective conservation strategy, which could also be used for the protection and genetic improvement of this endangered species.

2. Materials and Methods

2.1. Plant Materials and DNA Extraction

A total of 52 *P. bungeana* individuals were selected from five populations, including four natural populations located in three provinces (Gansu, Shaanxi and Shanxi): Maiji Mountain, Gansu (MJS), Lantian, Shaanxi (LT), Wuzi Mountain, Shaanxi (WZS), Baiwa Mountain, Shanxi (BWS), and an additional artificial introduction population, Jiufeng,

Beijing (JF) was also included in the present study (Figure 1). These four representative natural populations were selected from the core natural distribution areas of *P. bungeana* in China, according to previous investigations and studies [4]. The JF population was artificially introduced in the 1950s to prevent soil erosion. At present, trees grow well in this population, but the specific provenance origin remains to be investigated. The specific location, ecological conditions and sample size of the sampled populations are shown in Table S1. To avoid sampling related individuals, the distances between sampled trees were >10 m. The relatively low sample size per population is due to the fact that the sampled populations are altitudinally marginal, located on mountain peaks or cliffs, a fact that renders their exhaustive sampling extremely risky and dangerous. For this reason, the accessible individuals, fulfilling the distance requirements between them, were sampled. From every studied tree, fresh needles were collected, immediately immersed in liquid nitrogen, and stored at −80 °C until further use. Genomic DNA was extracted by the CTAB method [24].

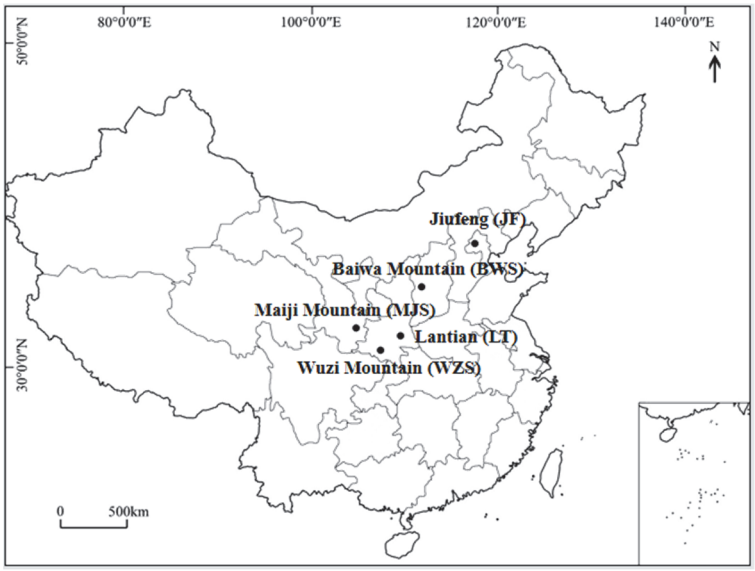

Figure 1. Geographical distribution of the studied *P. bungeana* populations.

2.2. SLAF Sequencing and Development of SNP Markers

According to the genome size and GC content of *P. bungeana*, the *Pinus taeda* genome (https://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/404/065/GCA_000404065.3_Ptaeda2.0/, accessed on 1 September 2019) was used as the reference genome for enzyme digestion prediction. Genomic DNA was digested with restriction enzymes (EcoRV-HF® and ScaI-HF®), and Nipponbare DNA was used as a control to assess the normal rate of enzyme digestion. The Specific-Locus Amplified Fragment (SLAF) library was constructed and SLAFs prepared for paired-end sequencing on the Illumina High-Seq 2500 sequencing platform (Illumina, Inc.; San Diego, CA, USA) at Beijing Biomarker Technologies Corporation. The raw reads were clustered based on similarity above 90%. The SLAF tags were defined as the group with the most samples. The sequence with the highest depth in each SLAF tag was used as a reference, and GATK [25] and SAMTOOLS [26] were employed for SNP calling. Vcftools v0.1.13 [27] was employed to filter all SNP loci according to completeness >80% and minor allele frequency (MAF) > 5%, and obtain high-consistent SNP loci.

2.3. Population Genetic Analyses

The genetic diversity of each population was evaluated using the selected SNP loci. Genetic diversity parameters, including observed and expected heterozygosity, nucleotide diversity, and inbreeding coefficient were calculated using software Vcftools v0.1.13 [27]. Population structure was analyzed by Admixture software [28], and clustering was con-

ducted by assuming that the cluster number (K value) ranged from 1 to 10. The clustering results were then cross-verified, and the optimal clustering number was determined according to the minimum value of the cross-validation error rate. Based on the genetic distance matrix, the phylogenetic tree was constructed by MEGA X [29], using the Kimura 2-parameter model, 1000 bootstrap replicates, and neighbor-joining method. Principal component analysis (PCA) was performed using EIGENSOFT [30] for the visualization of the genetic structure at multivariate space.

The shortest linear distance between populations was measured in Google Earth. We used the ade4 package in R to conduct a Mantel test (correlation between genetic and geographic distances) first only on 4 natural populations and then all 5 populations (with JF added). TreeMix 1.13 [31] was used to construct maximum-likelihood trees based on the population allele frequency covariance matrix and added migration edges to the phylogenetic tree to improve the fit, and inferred the migration events between populations and the directionality of gene flow.

3. Results

3.1. Development of Polymorphic SLAF Tags and Selection of SNP Markers

Through SLAF-seq, a total of 824.1 Mb reads were obtained among the 52 samples, and the average of Q30 reached 97.10% and average of GC reached 37.05%. A total of 23,597,049 SLAF tags with an average sequencing depth of 32.66× were obtained, and a total of 1,291,290 SNP markers were developed. The number of SNP markers in each sample ranged from 258,199 to 973,925. The SNP integrity and heterozygosity of each sample ranged from 19.99 to 75.42% and 8.13 to 23.28%, respectively. All SNPs were filtered according to the criteria of deletion rate <20% and minor allele frequency (MAF) > 5%. After filtering invalid SNPs, 346,840 high-quality SNP markers were retained for genetic analyses.

3.2. Genetic Diversity Analysis

The nucleotide diversity (π) of the five populations ranged from 0.2427 ± 0.2123 (JF) to 0.2842 ± 0.2110 (MJS). The observed heterozygosity (*Ho*) ranged from 0.3537 ± 0.0589 (JF) to 0.4245 ± 0.0367 (WZS); the expected heterozygosity (*He*) ranged from 0.3587 ± 0.0001 (LT) to 0.3914 ± 0.0014 (MJS) (Table 1). With the exception of WZS and BWS, the observed heterozygosity of the other three populations was lower than the expected heterozygosity (*Ho* < *He*). The inbreeding coefficients (F_{IS}) of WZS and BWS were negative, ranging from −0.0969 to −0.0697, indicating the presence of an excess of heterozygotes; while F_{IS} for JF, MJS and LT were all positive (range: 0.0063 to 0.0632), indicating that there was a low degree of inbreeding effect in these three populations, resulting in more homozygotes and less heterozygotes. The overall assessment indicates that the genetic diversity of MJS and WZS were relatively high, and that of JF population was comparatively low.

Table 1. Genetic diversity statistics among different the 5 *P. bungeana* populations.

Populations	N	*Ho* ± SD	*He* ± SD	π ± SD	F_{is}
JF	10	0.3537 ± 0.0589	0.3655 ± 0.0001	0.2427 ± 0.2123	0.0323
MJS	10	0.3663 ± 0.1207	0.3914 ± 0.0014	0.2842 ± 0.2110	0.0632
LT	12	0.3564 ± 0.0449	0.3587 ± 0.0001	0.2829 ± 0.1964	0.0063
WZS	10	0.4245 ± 0.0367	0.3870 ± 0.0002	0.2838 ± 0.2099	−0.0969
BWS	10	0.3938 ± 0.0239	0.3681 ± 0.0001	0.2548 ± 0.2105	−0.0697

N, sample size; H_o, observed heterozygosity; H_e, expected heterozygosity; π, nucleotide diversity; F_{is}, inbreeding coefficient; SD, standard deviation.

3.3. Population Structure Analysis

The optimal clustering number was determined to be 3 (K = 3), as it has the lowest cross-validation error rate (Figure 2a), indicating that the sampled populations could be categorized into three groups (Figure 2b). Group 1 contained 32 individuals from JF, BWS and LT, indicating that these three populations probably originated from the same ancestor.

Groups 2 and 3 were exclusively composed of individuals representing MJS and WZS populations, respectively.

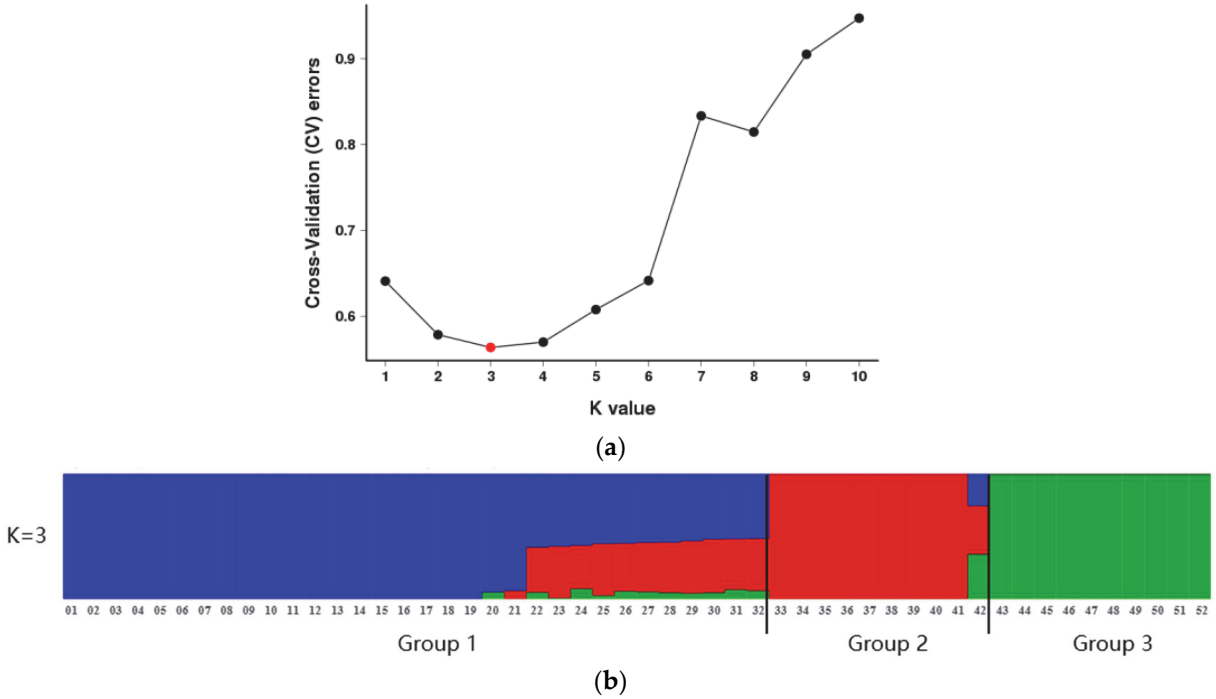

Figure 2. *P. bungeana* genetic structure. (**a**) Estimation of the optimal number of clusters, defined by the Evanno method; (**b**) The proportion of ancestry for each individual (K = 3). Each vertical bar corresponds to an individual, and different colors represent inferred genetic clusters. Segments of each vertical line show extent of admixture in an individual.

3.4. Phylogenetic Relationship and Principal Component Analysis

In order to further analyze the genetic relationship among individuals of different geographic populations, the constructed phylogenetic tree demonstrated that individuals within each region clustered together, indicating that the studied populations were relatively independent and had minimal gene exchange (Figure 3). The PCA plot allowed the visualization of the genetic structure at multivariate space, it showed that the first three principal components explained 18.14, 10.30, and 5.78% of the genetic variation, respectively (Figure 4), and its clustering result was similar to the phylogenetic tree. JF, BWS and LT were closely related and shared one cluster, supporting the assignments of genetic structure made by the population admixture analyses (Figure 2). Additionally, individuals from MJS and WZS grouped into two separate clusters, again supporting the population admixture analysis (Figure 2).

3.5. Geographic Differentiation

Mantel tests were conducted to evaluate population isolation-by-distance and demonstrated the existence of significant correlation between genetic and geographic distances among both the 4 natural populations ($r = 0.6$; $p < 0.05$) and all 5 populations ($r = 0.47$; $p < 0.05$). These findings further highlight the genetic separation among the studied populations determined by the previous analyses.

To investigate the historical split and admixture among the studied five populations, the TreeMix analysis to infer a maximum-likelihood tree and potential migration events was conducted (Figure 5). The five populations were clustered into two primary branches on which JF and BWS were clustered together, which were then clustered with LT into one branch, while WZS and MJS clustered into another branch. These results were similar to those of the neighbor joining phylogenetic tree, PCA and admixture analyses. Three gene

flow events were inferred among the branch tips of the TreeMix phylogeny with the following directions: (1) from WZS to JF (weight = 0.07), (2) from BWS to MJS (weight = 0.39), and (3) from the common ancestor of MJS and WZS to BWS (weight = 0.13).

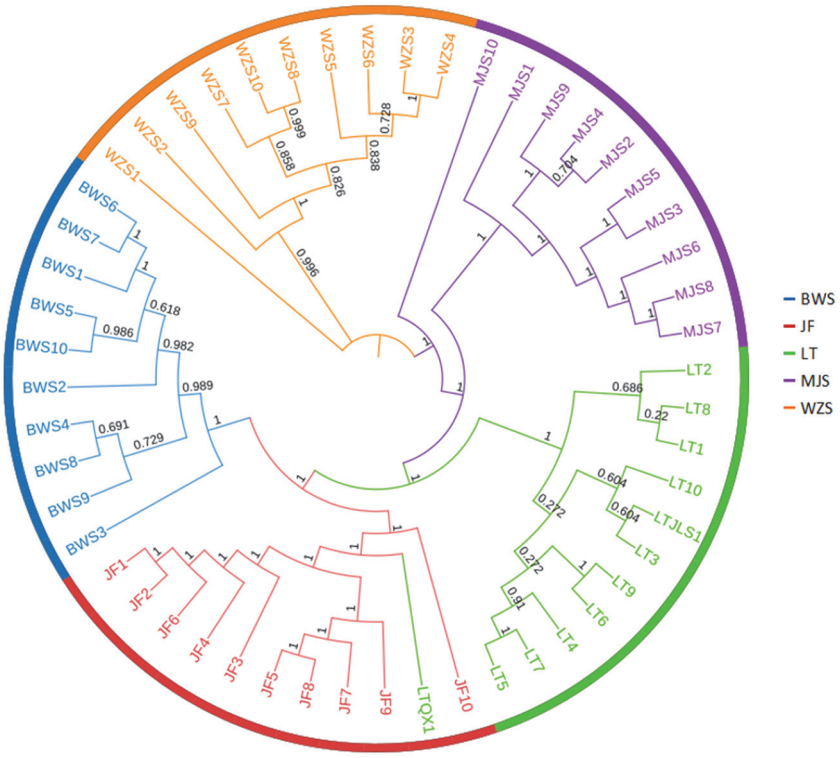

Figure 3. Phylogenetic tree of the 52 *P. bungeana* individuals using SNP data.

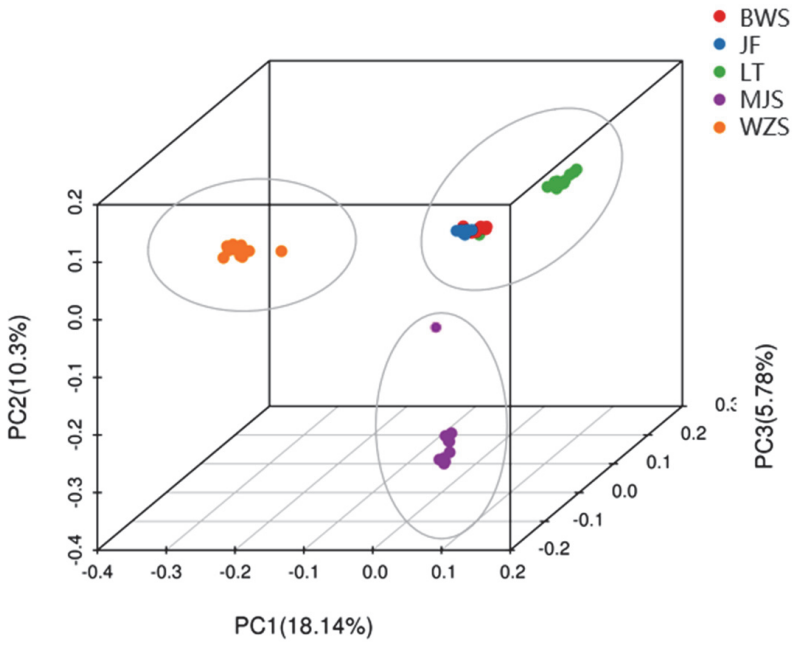

Figure 4. PCA plot of the 52 *P. bungeana* individuals based on the population SNPs.

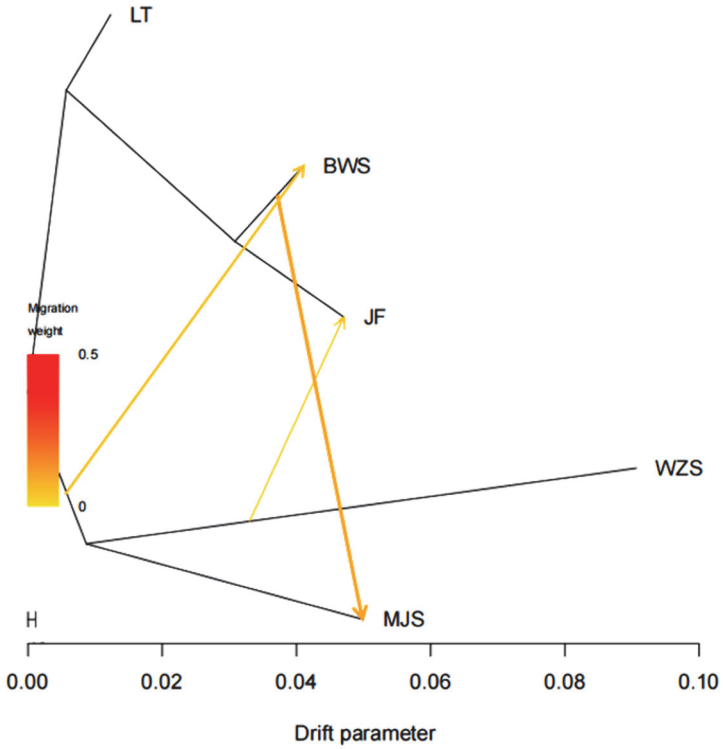

Figure 5. Maximum-likelihood tree and gene flow among *P. bungeana* populations inferred using TreeMix. Arrows indicate the direction of gene flow, while the line colors and thickness represent the migration weight based on sample number. Each branch represents a population, and the length of the horizontal branch refers to the proportion of genetic drift that occurs on each branch. The abscissa axis represents the genetic drift parameters.

4. Discussion

4.1. Genetic Diversity of P. bungeana

The extent of genetic diversity within a species' populations is the core measure in forest germplasm resources evaluation. Earlier studies evaluated *P. bungeana* genetic diversity based on morphological phenotypic differences [14,32]; however, these attributes, while important for population adaptation and survival, are affected by the environment, and thus the evaluation does not reflect the true genetic differences. Molecular markers represent a direct reflection of genetic polymorphisms at the DNA level and have been widely used in the evaluation and genetic analysis of many species. *P. bungeana* genetic variation assessment via molecular markers has been carried out before using AFLP [14], SSRs [15] and SCoT [16]. Common among these studies is the observed extremely low level of diversity within *P. bungeana* populations. While these studies enriched our understanding of the extent of genetic diversity of *P. bungeana*, they harboured an obvious limitation caused by the limited number of markers used, together with their low density. In the present study, we used SLAF-Seq to develop SNP marker sets representing the whole *P. bungeana* genome (i.e., density and number). Our results produced higher observed heterozygosity (Ho: 0.3537 ± 0.0589~0.4245 ± 0.0367) and expected heterozygosity (He: 0.3587 ± 0.0001~0.3914 ± 0.0014) than that reported using SSRs [15] and SCoT [16]; additionally, our estimate of nucleotide diversity (π: 0.2427 ± 0.2123~0.2842 ± 0.2110) was also much higher than that based on few nuclear sites [33]. It can be seen that compared with previously used DNA markers, SNP molecular markers at a large number and with a wide distribution across the genome resulted in detecting higher population structure resolution and polymorphism. Additionally, our results indicated that the genetic diversity level of *P. bungeana* populations gradually decreases from west to east across the species' distribution areas, result confirming previous observations [34], and highlighting the fact that the species western region is relatively genetically enriched. This could be attributed

to the complex terrain of this area, which provides a natural and stable protective barrier for *P. bungeana*, reduces the risk of human interference or climate change impact, resulting in the accumulation and preservation of rich genetic diversity. In fact, this area used to be a refuge for many conifers during the glacial period, such as *Pinus armandii* [35] and *Pinus tabuliformis* [36]. Then, in the process of population expansion and migration from the western region to the eastern region, potential founder effects may have resulted to the gradual reduction of genetic diversity form west to east.

4.2. Population Structure and Geographical Differentiation of P. bungeana

The population genetic structure and PCA indicated that populations (LT, BWS and JF) located in the central and eastern regions were clustered together, while those from the western regions (MJS and WZS) were separated; results were consistent with those of a previous study [10]. The phylogenetic tree analysis showed that almost all individuals in each region were clustered, indicating that the populations' evolution in different geographical regions was relatively independent and there was minimal gene admixture and introgression originating from populations of different regions. Additionally, a Mantel test was used to verify the effect of geographic distance on population differentiation. The results obtained indicated that the genetic and geographic distances among populations were significantly correlated, confirming once more results of previous studies [37]. At present, the habitats of *P. bungeana* natural populations were seriously fragmented, and most of the wild populations retreated to the peaks of mountains and cliffs [6]. The natural dispersal of pollen and seeds was hindered due to the reasons above, resulting in insufficient gene exchange between different populations. The significance of the estimated Mantel test values, is an indication that the *P. bungeana* populations are geographically isolated and this is an important factor causing the observed differentiation.

In previous studies, Zhao et al. [10] inferred that the western areas of Qinling Mountain and Daba Mountains could have served as refuges during the glacial period, and indicated that gene flow barriers among different *P. bungeana* regions mainly existed between populations in the west of Qinling Mountains and other populations, and that gene flow is rarely occurring among them. However, Yang et al. [7] inferred that the western areas of Qinling-Daba Mountains and the Lvliang Mountains could be potential refuges during the glacial period, and that a second contact or introgression might have occurred between the two refuges after the glaciation. In the present study, the maximum likelihood tree and potential migration events inferred from TreeMix analysis indicated the presence of historical genetic exchanges between MJS in the west of Qinling Mountains and BWS in the Lvliang Mountains (Figure 1). The directions of gene flow were from BWS to MJS (weight = 0.39), and from the common ancestor of MJS and WZS to BWS (weight = 0.13).

In addition, gene flow from WZS to JF was also detected (weight = 0.07); however, it should be stated that this gene flow is probably not among natural populations, but the result of artificial introduction and cultivation. The JF population may have originated from artificial migration (i.e., man-made introduction and plantings) of a small number of individuals from the western and central regions, so this may explain the low genetic diversity of the JF population, and potentially indicate the origin of this artificial population.

As the existing natural populations of *P. bungeana* are endangered and the terrains of distribution areas are relatively dangerous, sampling is dangerous and highly difficult. As mentioned above, the sample size per population was relatively small. However, we compensated for this shortcoming by selecting representative populations from the core of the species distribution and using extensive genomic SNPs data for population genetic analysis. The fact that the results of the analyses carried out were consistent and complementary to each other (i.e., F_{is} and H_o and H_e; admixture and phylogenetic tree, and Mantel and gene flow) supports previous theoretical findings related to sample size vs. marker number and density debate [38–40]. Compared with previous molecular marker methods, extensive genomic SNPs data can sensitively evaluate genetic diversity and accurately infer genetic structure and population differentiation history, representing an

effective tool for genetic analysis for some rare and endangered species that cannot offer ample individual numbers for classical population genetic analyses.

4.3. Suggestions on Conservation of P. bungeana Germplasm Resources

According to the present status of *P. bungeana* disrupted and fragmented distribution, the observed level of populations genetic diversity, genetic structure and differentiation, and other biological characteristics (e.g., fecundity, recruitment etc.), devising proper conservation strategies requires a dual prong approach, namely, in situ and ex situ conservation modes. Current recommendations are focusing on the implementation of in situ conservation for the natural populations in the western region, such as the MJS and WZS populations, by preventing harvesting, designating protected areas, and maintaining the original habitats together with the high genetic diversity of the species harbored within, as much as possible. Meanwhile, for those populations with low genetic diversity or areas with seriously damaged habitats, it is recommended to assist and facilitate their natural regeneration by applying the appropriate forest management practices, or if needed to apply artificial regeneration with local forest reproductive material, in order to expand the existing populations and enhance their reproductive capacity and fitness. For ex situ conservation, germplasm collection specially from highly differentiated populations, that will be large enough to include rare or unique alleles, and from as many populations located in different areas as possible should take the highest priority; moreover, sampling of genetic material should also be conducted in areas with high genetic diversity to secure their genetic resources in ex situ collections.

5. Conclusions

In summary, we used high density and strong specificity SNP loci to carry out genetic analysis on 52 *P. bungeana* individuals from 5 populations representing the main regions of the species distribution in China. Genetic diversity assessment indicated a trend of genetic diversity gradual decrease from west to east across the species distribution areas. Population genetic structure, PCA and phylogenetic clustering consistently indicated that populations in the central and eastern regions were clustered together, while those from the western regions were separated. Mantel test values indicated the presence of geographic isolation among populations, an important factor contributing to the observed genetic differentiation. The maximum likelihood tree and potential migration events inferred from TreeMix analysis indicated the presence of historical genetic exchanges between the west of Qinling Mountains and the Lvliang Mountains populations. Based on the generated genetic information, in situ and ex situ conservation strategies for *P. bungeana* germplasm resources are proposed, these strategies could be valuable for the conservation, protection and genetic improvement of this endangered species.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/f13020326/s1, Table S1. Sampling points information.

Author Contributions: Conceptualization, Q.T., Y.A.E.-K. and W.L.; data curation, Q.T.; formal analysis, Q.T.; funding acquisition, W.L.; methodology, Q.T.; resources, W.L.; supervision, W.L.; writing—original draft, Q.T.; writing—review and editing, Y.A.E.-K. and W.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by grant from the National Natural Science Foundation of China (31770713, 31860221).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available from corresponding authors upon reasonable request.

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

References

1. Fu, L.G.; Li, N.; Mill, R.R. *Flora of China*; Wu, C.Y., Raven, P.H., Eds.; Science Press and Missouri Botanical Garden Press: Beijing, China, 1999; Volume 4.
2. Zhao, Y.; Zhang, X.Z.; Wang, X.A. A study on geographical distribution law of *pinus bungeana* natural forests in China. *Acta Bot. Boreali Occident. Sin.* **1995**, *15*, 161–166.
3. Qin, H.N.; Yang, Y.; Dong, S.Y.; He, Q. Threatened Species List of China's Higher Plants. *Biodiv. Sci.* **2017**, *25*, 696–744. [CrossRef]
4. Li, B.; Gu, W.C. Conservation genetics of *Pinus bungeana*—evaluation and conservation of natural populations' genetic diversity. *Sci. Silvae Sin.* **2005**, *1*, 57–64.
5. Li, B.; Gu, W.C. Distribution characteristics and research progress of *Pinus bungeana*. *For. Res.* **2003**, *16*, 225–232.
6. Zhao, H. *Evaluation and Conservation Strategy of Genetic Resources of Pinus bungeana Zucc*; Chinese Academy of Forestry: Beijing, China, 2012.
7. Yang, Y.X.; Wang, M.L.; Liu, Z.L.; Zhu, J.; Yan, M.Y. Nucleotide polymorphism and phylogeographic history of an endangered conifer species *Pinus bungeana*. *Biochem. Syst. Ecol.* **2016**, *64*, 89–96. [CrossRef]
8. Frankham, R.; Ballou, J.D.; Ralls, K.; Eldridge, M.; Dudash, M.R.; Fenster, C.B.; Lacy, R.C.; Sunnucks, P. *Genetic Management of Fragmented Animal and Plant Populations*; Oxford University Press: Oxford, UK, 2017.
9. Li, B.; Gu, W.C.; Zhou, S.L. Conservation genetics of *Pinus bungeana* I. Gene protection analysis. *Biodiversity* **2003**, *11*, 28–36.
10. Zhao, H.; Zheng, Y.Q.; Li, B.; Lin, F.R.; Zhang, C.H.; Cheng, P.P.; Huang, P. Genetic structure analysis of natural populations of *Pinus bungeana* in different geographical regions. *J. Plant Genet. Resour.* **2013**, *14*, 395–401.
11. Liddell, E.; Cook, C.N.; Sunnucks, P. Evaluating the use of risk assessment frameworks in the identification of population units for biodiversity conservation. *Wildl. Res.* **2020**, *47*, 208–216. [CrossRef]
12. Falk, D.A.; Holsinger, K.E. *Genetics and Conservation of Rare Plants*; Oxford University Press: New York, NY, USA, 1991.
13. Hogbin, P.M.; Peakall, R. Evaluation of the contribution of genetic research to the management of the endangered plant *Zieria prostrata*. *Conserv. Biol.* **1999**, *13*, 514–522. [CrossRef]
14. Li, B. *Study on Genetic Diversity and the Sampling Strategy for Conserving Core Collection in Pinus bungeana*; Beijing Forestry University: Beijing, China, 2002.
15. Li, X.M.; Jin, Z.Y.; Su, A.R.; Yang, L.X.; Su, S.; Wang, L.; Zhang, J. Distribution characteristics and primer development of EST-SSR sequence of *Pinus bungeana*. *For. Ecol. Sci.* **2019**, *34*, 266–272.
16. Zhang, X.X.; Liu, B.G.; Li, Y.; Liu, Y.; He, Y.X.; Qian, Z.H.; Li, J.X. Landscape genetics reveals that adaptive genetic divergence in *Pinus bungeana* (Pinaceae) is driven by environmental variables relating to ecological habitats. *BMC Evol. Biol.* **2019**, *19*, 1–13. [CrossRef] [PubMed]
17. Mammadov, J.; Aggarwal, R.; Buyyarapu, R.; Kumpatla, S. SNP markers and their impact on plant breeding. *Int. J. Plant Genom.* **2012**, *2012*, 728398. [CrossRef] [PubMed]
18. Zavinon, F.; Adoukonou, S.H.; Keilwagen, J.; Lehnert, H.; Ordon, F.; Perovic, D. Genetic diversity and population structure in Beninese pigeon pea [*Cajanus cajan* (L.) Huth] landraces collection revealed by SSR and genome wide SNP markers. *Genet. Resour. Crop. Evol.* **2020**, *67*, 1–18. [CrossRef]
19. Bradbury, I.R.; Hamilton, L.C.; Dempson, B.; Robertson, M.J.; Bourret, V.; Bernatchez, L.; Verspoor, E. Transatlantic secondary contact in Atlantic Salmon, comparing microsatellites, a single nucleotide polymorphism array and restriction-site associated DNA sequencing for the resolution of complex spatial structure. *Mol. Ecol.* **2015**, *24*, 5130–5144. [CrossRef] [PubMed]
20. Jeffries, D.L.; Copp, G.H.; Handley, L.L.; Olsén, K.H.; Sayer, C.D.; Hänfling, B. Comparing RADseq and microsatellites to infer complex phylogeographic patterns, an empirical perspective in the Crucian carp, *Carassius carassius*, L. *Mol. Ecol.* **2016**, *25*, 2997–3018. [CrossRef] [PubMed]
21. Bohling, J.; Small, M.; Bargen, J.V.; Louden, A.; DeHaan, P. Comparing inferences derived from microsatellite and RADseq datasets: A case study involving threatened bull trout. *Conserv. Genet.* **2019**, *20*, 329–342. [CrossRef]
22. Du, X.; Zhang, X.; Bu, H.D.; Zhang, T.C.; Lao, Y.C.; Dong, W.X. Molecular analysis of evolution and origins of cultivated hawthorn (*Crataegus* spp.) and related species in China. *Front. Plant Sci.* **2019**, *10*, 443. [CrossRef]
23. Rodger, Y.S.; Pavlova, A.; Sinclair, S.; Pickup, M.; Sunnucks, P. Evolutionary history and genetic connectivity across highly fragmented populations of an endangered daisy. *Heredity* **2021**, *126*, 846–858. [CrossRef]
24. Murray, M.G.; Thompson, C.L.; Wendel, J.F. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* **1980**, *8*, 4321–4326. [CrossRef]
25. McKenna, A.; Hanna, M.; Banks, E.; Sivachenko, A.; Cibulskis, K.; Kernytsky, A.; Garimella, K.; Altshuler, D.; Gabriel, S.; Daly, M.; et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **2010**, *20*, 1297–1303. [CrossRef]
26. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The sequence alignment/map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [CrossRef] [PubMed]
27. Danecek, P.; Auton, A.; Abecasis, G.; Cornelis, A.A.; Banks, E.; Mark, A.D.; Robert, E.H.; Lunter, G.; Gabor, T.M.; Stephen, T.S.; et al. The variant call format and VCFtools. *Bioinformatics* **2011**, *27*, 2156–2158. [CrossRef] [PubMed]
28. Alexander, D.H.; Novembre, J.; Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **2009**, *19*, 1655–1664. [CrossRef] [PubMed]

29. Sudhir, K.; Glen, S.; Li, M.; Christina, K.; Koichiro, T. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549.
30. Price, A.L.; Patterson, N.J.; Plenge, R.M.; Weinblatt, M.E.; Shadick, N.A.; Reich, D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **2006**, *38*, 904–909. [CrossRef] [PubMed]
31. Pickrell, J.K.; Pritchard, J.K. Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet.* **2012**, *8*, e1002967. [CrossRef]
32. Wang, X.P.; Liu, J.L.; Wang, J.L.; Liu, C.J. Geographic variation of morphological characteristics of *Pinus bungeana* seeds and cones. *J. Beijing For. Univ.* **1998**, *20*, 28–34.
33. Yang, Y.X.; Zhi, L.Q.; Jia, Y.; Zhong, Q.Y.; Liu, Z.L.; Yue, M.; Li, Z.H. Nucleotide diversity and demographic history of *Pinus bungeana*, an endangered conifer species endemic in China. *J. Syst. Evol.* **2020**, *58*, 282–294. [CrossRef]
34. Zhao, H.; Zheng, Y.Q.; Li, B.; Zhang, C.H.; Lin, F.R.; Yu, X.D.; Cheng, B.B.; Huang, P. Genetic diversity analysis of *Pinus bungeana* natural populations with EST-SSR markers. *For. Res.* **2014**, *27*, 474–480.
35. Wang, R.N.; Bai, J.Q.; Wang, M.L.; Yan, X.H.; Dong, P.B.; Wang, N.; Li, Z.H. Evolutionary History of an Endemic Conifer Tree *Pinus armandii* in China. *Acta Bot. Boreali Occident. Sin.* **2019**, *39*, 42–51.
36. Xia, H.H.; Wang, B.; Pan, J.; Zhao, W.; Li, Y.; Mao, J.F.; Wang, X.R. Combiningmitochondrial and nuclear genome analyses to dissect the effects of colonization, environment, and geography on population structure in *Pinus tabuliformis*. *Evol. Appl.* **2018**, *11*, 1931–1945. [CrossRef] [PubMed]
37. Zhou, H.J. *Genetic Diversity and Population Structure of Natural Endangered Forest Tree Pinus bungeana in China*; Northwestern University: Xi'an, China, 2013.
38. Li, S.; Jakobsson, M. Estimating demographic parameters from large-scale population genomic data using Approximate Bayesian Computation. *BMC Genet.* **2012**, *13*, 22. [CrossRef] [PubMed]
39. Felsenstein, J. Accuracy of coalescent likelihood estimates: Do we need more sites, more sequences, or more loci? *Mol. Biol. Evol.* **2006**, *23*, 691–700. [CrossRef] [PubMed]
40. Robinson, J.; Bunnefeld, L.; Hearn, J.; Stone, G.; Hickerson, M.J. ABC inference of multi-population divergence with admixture from unphased population genomic data. *Mol. Ecol.* **2014**, *18*, 4458–4471. [CrossRef]