forests

MDPI

Article

Genetic Diversity and Population Structure of Natural *Pinus koraiensis* Populations

Yue W. Tong [1,2], Bernard J. Lewis [1], Wang M. Zhou [1], Cheng R. Mao [1,2], Yan Wang [1,2], Li Zhou [1], Da P. Yu [1], Li M. Dai [1] and Lin Qi [1,*]

1 CAS Key Laboratory of Forest Ecology and Management, Institute of Applied Ecology, Shenyang 110016, Liaoning, China; tyw8704123@163.com (Y.W.T.); lewisbern@gmail.com (B.J.L.); zhouwangming@126.com (W.M.Z.); mcr0707@163.com (C.R.M.); wyan19951008@163.com (Y.W.); zhouli930@iae.ac.cn (L.Z.); yudp2003@iae.ac.cn (D.P.Y.); lmdai@iae.ac.cn (L.M.D.)

2 College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China

* Correspondence: qilin@iae.ac.cn; Tel.: +86-024-8397-0328

Received: 19 November 2019; Accepted: 23 December 2019; Published: 26 December 2019

Abstract: Studying the genetic diversity and population structure of natural forest populations is essential for evaluating their ability to survive under future environmental changes and establishing conservation strategies. *Pinus koraiensis* is a conifer species with high ecological and economic value in Northeast China. However, its natural forests have been greatly reduced in recent years, mostly due to over exploitation and over utilization. Here, we evaluated the genetic diversity and population structure of seven populations of *P. koraiensis* located throughout its native distribution. A total of 204 samples were genotyped with nine polymorphic nuclear SSR (simple sequence repeat) markers. The results showed high genetic diversity in all populations, with an average expected heterozygosity of 0.610, and the northern-most populations (Dailin (DL) and Fenglin (FL)) showed slightly higher diversity than the other five populations. The level of genetic differentiation among populations was very low (F_{ST} = 0.020). Analysis of molecular variance (AMOVA) showed that only 2.35% of the genetic variation existed among populations. Moreover, STRUCTURE analysis clearly separated the seven populations into two clusters. Populations DL and FL from the Xiaoxinganling Mountains comprised cluster I, while cluster II included the five populations from the Changbai Mountains and adjacent highlands. Our research on the genetic diversity and population structure of *P. koraiensis* in natural forests of China can provide a basis for the implementation of programs for the conservation and utilization of *P. koraiensis* genetic resources in the future.

Keywords: *Pinus koraiensis*; genetic diversity; genetic differentiation; population structure; conservation strategies

1. Introduction

Pinus koraiensis Sieb. et Zucc., also known as Korean pine, is an evergreen tree belonging to the *Pinus* genus and Pinaceae family [1]. It is mainly distributed in the mountainous area of Northeast China, as well as in North Korea, Japan, and Far Eastern region of Russia [1–3]. *P. koraiensis* plays a key ecological role as the most important component of natural broad-leaved Korean pine forests, and it is also famous for its economic value, such as the production of good quality timber and edible seeds [4,5]. However, in recent decades, with the increased demand for timber and pine nuts from *P. koraiensis*, its natural forests have greatly declined due to excessive harvesting [6–9], the original forest has become extremely small, and the genetic resources of the species have been threatened by deforestation. Due to the status of *P. koraiensis* as a rare and nationally endangered species in China

(http://www.plant.csdb.cn/endangeredplants), protecting its resources, especially its genetic resources, has become urgent.

An important indicator of genetic resource conservation is the amount of genetic diversity, which is widely recognized as a key determinant of the long-term survival of species [10,11]. Genetic diversity in forests is determined by gene flow, genetic drift, selection, mutation, and other processes [12,13], and it provides the raw material for the adaptation, evolution, and survival of species under changing environmental conditions [14,15]. Another indicator is population structure, which is the distribution pattern of genes and genotypes in time and space, which is informative for understanding genetic diversity [16]. Studying the genetic diversity and population structure of forest trees with a long life cycle, a wide distribution, and high ecological and economic value is of great significance for genetic resource conservation and forest ecosystem management [13,17–19].

There are relatively few genetic investigations of *P. koraiensis*. Potenko and Velikov (1998) investigated the genetic diversity and differentiation of 19 natural Russian populations of *P. koraiensis* using 15 enzyme systems [2] and found relatively low levels of genetic diversity and low differentiation among populations (F_{ST} = 0.015). Kim et al. (2005) studied and compared the genetic variation in Korean pine from 12 natural populations in Korea, China, and Russia using allozymes and random amplified polymorphic DNA (RAPD) [20]. Their results showed that differentiation among the three different regions was low, and genetic variation decreased from south (Korea) to north (Russia) on latitude gradients. Furthermore, Feng et al. (2006) analyzed the genetic diversity and structure of four natural *P. koraiensis* populations in China by using inter-simple sequence repeats (ISSRs) [4] and applied sequence-related amplified polymorphism (SRAP) markers to study the genetic diversity of 24 different provenances of *P. koraiensis* in a seed orchard in China [21]. Both studies indicated that the genetic diversity levels of *P. koraiensis* were high, and no close relationship could be established between genetic diversity and geographical distances. However, these studies still had some limitations, such as small sample sizes that could not represent the populations, small populations that could not represent the whole distribution in China, research characteristics that could not represent the natural forests and outdated detection technology, and failure to investigate relationships between genetic diversity and environmental factors. Therefore, research based on the natural resource distribution of all existing populations, a more comprehensive sample size, and a suitable means of detection is necessary to study the genetic diversity and population structure of *P. koraiensis*.

In this context, we investigated the genetic diversity and population structure of seven *P. koraiensis* populations distributed across the species' natural range in Northeast China by means of simple sequence repeat markers (SSRs). The aim of this study was to (a) identify regions of high genetic diversity of *P. koraiensis*, (b) determine the number of genetic clusters for population structure, (c) understand whether any populations or regions are genetically distinct, and (d) test if genetic diversity is related to environmental or climatic gradients.. The findings will be useful for the genetic conservation, exploration, and development of breeding programs of this species.

2. Materials and Methods

2.1. Plant Material and Sampling

A total of 204 georeferenced tender leaf samples of *P. koraiensis* were collected from 28 to 30 randomly selected trees in seven different populations, representing most of the distribution area of this species in Northeast China (Figure 1; Table 1). The detailed population information and climate parameters of each population are listed in Table 1. To avoid sampling from closely related individuals, the distance between sampled trees was at least 200 m. All collected samples were enclosed in plastic bags, brought back to the laboratory and stored at −80 °C before DNA extraction. This work was guided by "Observation Methodology for Long-term Forest Ecosystem Research" of National Standards of the People's Republic of China (GB/T 33027-2016).

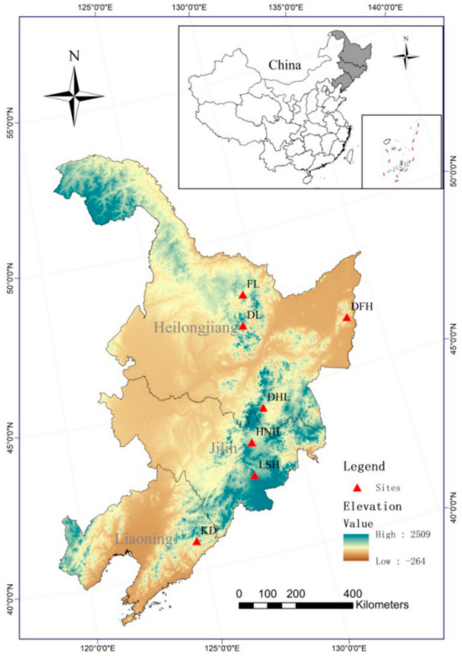

Figure 1. Map illustrating the location of the seven populations of *P. koraiensis* in this study.

Table 1. Population information and geographical characteristics of seven populations of *P. koraiensis*.

Population	ID	N	Location (Province)	Latitude (°)	Longitude (°)	Annual Mean Temperature (°C)	Annual Mean Precipitation (mm)
Kuandian	KD	30	Liaoning	40.91	124.78	6.29	973.04
Lushuihe	LSH	28	Jilin	42.53	127.80	4.12	783.68
Huangnihe	HNH	30	Jilin	43.55	128.01	3.76	635.42
Dahailin	DHL	28	Heilongjiang	44.52	128.86	1.14	619.72
Dongfanghong	DFH	29	Heilongjiang	46.58	133.58	3.17	648.99
Dailin	DL	29	Heilongjiang	47.18	128.85	0.30	574.59
Fenglin	FL	30	Heilongjiang	48.13	129.19	0.04	596.15
All populations		204					

2.2. DNA Extraction and Microsatellite Genotyping

Total genomic DNA was extracted from the samples using a DNAsecure Plant Kit (DP320, Tiangen, Beijing, China) according to the manufacturer's instructions. A total of nine nuclear microsatellite primers developed for *P. koraiensis* [22,23] were selected, as shown in Table 2, and the forward primers were labelled with a fluorescent dye (FAM or HEX).

Polymerase chain reaction (PCR) was performed in a 20 µL reaction volume consisting of 40 ng of genomic DNA template, 0.3 µM concentrations of each primer, 2 µL of 10× buffer, 0.1 mM dNTPs (TransGen Biotech, Beijing, China), and 1.0 U of Taq DNA polymerase (TransGen Biotech, Beijing, China). Amplification was conducted by the following cycling parameters: Initial denaturation at 94 °C for 5 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing for 35 s (for the annealing temperatures see Table 2), and extension at 72 °C for 40 s and a final extension at 72 °C for 3 min. Then, the PCR products were run on an ABI 3730XL DNA Analyzer (Applied Biosystems, Foster City, CA, USA), using GS-500 LIZ as an internal size standard. Allele binning and genotyping were performed with GeneMarker version 2.20 (Soft Genetics, State College, PA, USA).

Table 2. Primer characteristics of nine microsatellite markers in *Pinus koraiensis*. The forward primers were labelled with a fluorescent dye (FAM or HEX) [22,23].

Locus	Dye	Primer Sequences (5′–3′)	Repeat Motif	Range (bp)	T_a (°C)
P5	FAM	F: ATTCCTACTTTTCCCGTTT R: ACAGAGACCCCGTTTACAT	$(CA)_{11}$	108–118	55
P6	FAM	F: TCAAATTACCAGACAATAA R: GAATTCGCCAATGAAATCA	$(TA)_3(GT)_{15}$	107–128	55
P29	HEX	F: TGTCAACTTTGAACCCTGAA R: AGGCCAATCCTCATACATTT	$(CA)_9$	134–148	55
P45	HEX	F: CTTACATTTTGCTGCTTTTC R: TTGTCAGTTTTAGGTTGGAT	$(TG)_{16}(AG)_{17}$	165–203	55
P51	HEX	F: CCTAAGAGCAATGTAAAATG R: AGCTTGACAACGACTAACT	$(AG)_{15}$	188–225	55
P52	FAM	F: CCATCCTTCAAATTTTCCT R: GCCATTCTTTCTACCACTT	$(AG)_{26}$	113–145	56
P62	FAM	F: CAAGGAGGAAAACAATAAGG R: CTACAACAGAAACTAGCCAGA	$(CT)_{10}$	127–133	56
P63	HEX	F: CTCCTTCTTCATCCATCCATT R: TGAGGTGAGCCTGCATATAGT	$(CT)_{19}$	218–252	55
P79	HEX	F: CCACCGCCAAGTCCATTA R: GCTTTGTTAGCCGTCCAG	$(CAA)_7$	183–201	55

Note: T_a = annealing temperature.

2.3. Data Analysis

2.3.1. Genetic Diversity Indices

Genetic diversity per locus and population was evaluated by using GenAlEx 6.5.1 [24]. The number of different alleles (N_a), the number of effective alleles (N_e), the observed (H_o) and expected (H_e) heterozygosities, and the Shannon diversity index (I) were all calculated. In addition, we calculated allelic richness (A_r), the coefficient of differentiation between pairs of populations (F_{ST}), and the inbreeding coefficient (F_{IS}) [25] in FSTAT 2.9.3 [26].

Correlations between genetic diversity parameters and geoclimatic factors were determined using a Spearman nonparametric correlation coefficient matrix constructed in R version 3.53 [27]. The climatic data (i.e., annual mean temperature (T_{mean}), maximum temperature of the warmest month (T_{max}), minimum temperature of the coldest month (T_{min}), and annual precipitation ($Prec$)), were extracted from https://climexp.knmi.nl [28] by using each population's longitude and latitude.

2.3.2. Population Structure Analysis

First, a non-hierarchical analysis of molecular variance (AMOVA) (1000 permutations) based on the degree of genetic divergence among populations was performed using GenAlEx 6.5.1 [24]. Next, population structure was analyzed based on Bayesian clustering using STRUCTURE 2.3.4 [29]. The cluster number was set from 1 to 8 (number of populations plus 1), and the populations set as location priors (LOCPRIOR) [29] under the admixture model were used to run the Markov chain Monte Carlo (MCMC) simulation algorithm. The length of the burn-in period was set to 10,000 iterations. The number of MCMC iterations after the burn-in period was set to 100,000, and for each K value the calculation was repeated 10 times. The optimal K value was obtained by the method of Evanno [30]. The 10 runs for the optimal K were averaged by using the programs STRUCTURE HARVESTER and CLUMPP 1.1.2 [31]. After that, a hierarchical AMOVA which was calculated considering the main groups obtained from the STRUCTURE analysis was implemented by the software GenAlEx 6.5.1 [24]. The statistical significance was also tested using a nonparametric approach described in Excoffier et al. (1992) with 1000 permutations [32].

3. Results

3.1. Genetic Diversity

In total, 72 alleles were amplified from nine SSR primers across the 204 *P. koraiensis* samples, with an average of 6.7 alleles per locus (Table 3). Among loci, observed heterozygosity (H_o) and expected heterozygosity (H_e) ranged from 0.416 to 0.922 and from 0.351 to 0.846, with means of 0.741 and 0.610, respectively. H_o was higher than H_e at each microsatellite locus. A_r ranged from 2.5 (P5) to 11.9 (P45), and the average A_r was 6.8 (Table 3). Meanwhile, the F_{IS} showed negative and significant values at all loci except P62 and P79, and the overall F_{IS} across the nine loci was −0.231 (Table 3, $p < 0.05$), indicating heterozygote excess in the *P. koraiensis* populations.

Table 3. Genetic diversity parameters of number of alleles (N_a), number of effective alleles (N_e), Shannon's index (I), observed heterozygosity (H_o), expected heterozygosity (H_e), unbiased expected heterozygosity (uH_e), allelic richness (A_r), and the inbreeding coefficient (F_{IS}) for each locus. Significant values are indicated with * ($0.01 < p < 0.05$).

Locus	N_a	N_e	I	H_o	H_e	uH_e	A_r	F_{IS}
P5	2.3	1.6	0.565	0.425	0.358	0.364	2.5	−0.186 *
P6	7.1	2.3	1.214	0.688	0.561	0.571	7.4	−0.227 *
P29	3.1	1.6	0.649	0.416	0.351	0.357	3.3	−0.186 *
P45	11.6	6.1	2.043	0.922	0.836	0.851	11.9	−0.102 *
P51	11.7	6.6	2.122	0.922	0.846	0.861	11.7	−0.089 *
P52	8.3	3.7	1.560	0.868	0.714	0.726	8.3	−0.216 *
P62	3.9	2.4	0.987	0.833	0.574	0.584	4.5	−0.451
P63	9.1	4.8	1.779	0.921	0.787	0.800	9.5	−0.171 *
P79	2.7	1.9	0.699	0.676	0.466	0.474	2.6	−0.450
Mean	6.7	3.4	1.291	0.741	0.610	0.621	6.8	−0.231 *

Genetic diversity parameters at the population level are given in Table 4. Among populations, N_a and N_e ranged from 5.9 to 7.1 and 3.1 to 3.8, with means of 6.7 and 3.4, respectively. H_o ranged from 0.700 to 0.808, and H_e ranged from 0.581 to 0.636, with means of 0.741 and 0.610, respectively. H_o was higher than H_e in each population. For the H_o and H_e descriptors, the two northern-most populations (Dailin (DL) and Fenglin (FL)) had higher values than the five southern populations, but the difference was not significant. The overall average A_r was 6.6, the largest value was observed in the Lushuihe (LSH) population (7.1), and the smallest value was observed in the Kuandian (KD) population (5.8). Meanwhile, the F_{IS} showed negative and significant values in all seven populations, indicating a slight departure from Hardy–Weinberg equilibrium with heterozygote excess.

Table 4. Genetic diversity parameters for the seven *P. koraiensis* populations analyzed with nine microsatellite loci: Mean number of different alleles (N_a), mean number of effective alleles (N_e), Shannon's index (I), observed heterozygosity (H_o), expected heterozygosity (H_e), unbiased expected heterozygosity (uH_e), allelic richness (A_r) and the inbreeding coefficient (F_{IS}). Significant values are indicated with * ($0.01 < p < 0.05$) and ** ($p < 0.01$).

Population	N_a	N_e	I	H_o	H_e	uH_e	A_r	F_{IS}
KD	5.9	3.3	1.204	0.700	0.581	0.591	5.8	−0.188 **
LSH	7.1	3.4	1.303	0.718	0.605	0.616	7.1	−0.170 **
HNH	7.1	3.4	1.332	0.730	0.618	0.629	7.0	−0.164 **
DHL	6.7	3.8	1.339	0.718	0.616	0.627	7.0	−0.148 **
DFH	6.7	3.1	1.214	0.743	0.586	0.596	6.6	−0.252 **
DL	6.6	3.5	1.324	0.808	0.636	0.647	6.5	−0.254 **
FL	6.6	3.6	1.321	0.770	0.629	0.640	6.5	−0.208 **
Mean	6.7	3.4	1.291	0.741	0.610	0.621	6.6	−0.198 *

Thus, the overall genetic diversity in *P. koraiensis* populations was relatively high, and the northern-most populations DL and FL showed the highest diversity. By comparison, the southern-most population KD showed the lowest level of genetic diversity.

3.2. Genetic Differentiation

We estimated the overall population differentiation degree among seven populations of *P. koraiensis*, and the F_{ST} was 0.020. This indicated that 2.0% of the genetic variation existed among the seven *P. koraiensis* populations, whereas 98.0% of the genetic variation existed within populations. In other words, the genetic variation within populations of *P. koraiensis* was the main source of variation. Pairwise F_{ST} values were very low (0.007 to 0.021), and 10 of 21 pairwise F_{ST} values were significant, which were mainly found in the comparisons of the two northern populations (DL and FL) (Table 5). This result illustrated that the level of genetic differentiation among populations was very low.

Table 5. Pairwise F_{ST} values for all populations of *P. koraiensis*. Significant values are indicated with * ($p < 0.05$).

KD	LSH	HNH	DHL	DFH	DL	FL	
0.000							KD
0.009	0.000						LSH
0.007	0.008	0.000					HNH
0.010	0.007	0.008	0.000				DHL
0.008	0.008	0.007	0.009	0.000			DFH
0.016 *	0.016 *	0.017 *	0.021 *	0.019 *	0.000		DL
0.015 *	0.014 *	0.014 *	0.017 *	0.014 *	0.007 *	0.000	FL

We performed non-hierarchical AMOVA among and within *P. koraiensis* populations (Table 6), and the results showed that the genetic variation among populations was only 2.35% ($p < 0.001$). Thus, most variation occurred within populations (97.65%, $p < 0.001$). These results also indicated that the genetic variation in *P. koraiensis* mainly occurred within populations, and the genetic differentiation level among populations was very low.

Table 6. Non-hierarchical and hierarchical AMOVAs of *P. koraiensis* populations. Significant values are indicated with * ($p < 0.05$).

Analysis	Source of Variation	df	Sum of Squares	Variance Component	Percentage of Variation
Non-hierarchical AMOVA	Among populations	6	45.619	0.108	2.35% *
	Within populations	197	879.895	4.466	97.65% *
	Total	203	925.515	4.574	100.00%
Hierarchical AMOVA	Between groups	1	19.414	0.169	3.62% *
	Among populations	5	26.206	0.027	0.57%
	Within populations	197	879.895	4.466	95.81% *
	Total	203	925.515	4.662	100.00%

3.3. Population Structure

The population structure analysis provided additional information on the level of genomic admixture among populations. The results from STRUCTURE showed that at K = 2, ΔK was optimal (Figure 2), indicating that the most likely division of *P. koraiensis* populations included two clusters (Figure 3). The KD, LSH, HNH, DHL, and DFH populations were assigned to cluster I (blue, F_{ST} = 0.0031); cluster II (red, F_{ST} = 0.0223) included the two northern-most populations (DL and FL).

Furthermore, a hierarchical AMOVA was carried out according to the two main groups (clusters) obtained by STRUCTURE analysis. The results showed that the molecular variance between groups

was 3.62% ($p < 0.001$), and the majority of the genetic variation in *P. koraiensis* was located within populations (95.81%, $p < 0.001$) (Table 6).

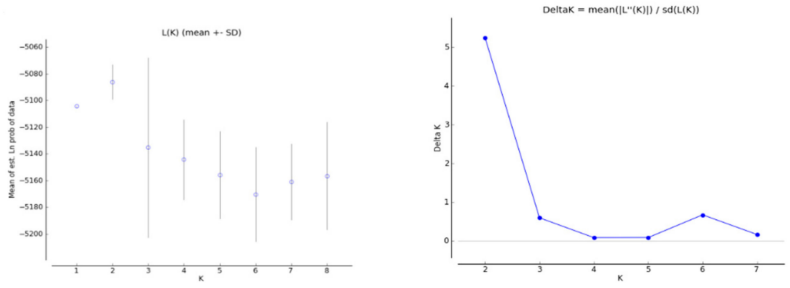

Figure 2. Diagnostic plots of L(K) and Delta-K from the STRUCTURE analysis of *P. koraiensis* populations.

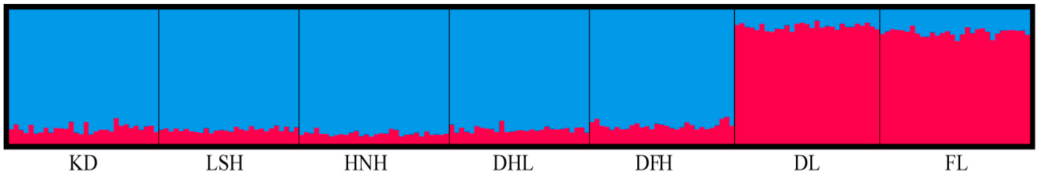

Figure 3. Results of the STRUCTURE analysis of *P. koraiensis* populations at K = 2. Each individual is represented by a single vertical bar, which is partitioned among gene pools. Colors represent genetic clusters, and the colored segments show the individual's estimated ancestry proportion.

3.4. Correlations Between Genetic Diversity Parameters and Geoclimatic Variables

We performed Spearman's correlation analyses between the calculated genetic diversity parameters and geoclimatic variables (Table 7). The results revealed no significant correlations between any of the genetic diversity descriptors (N_a, N_e, I, H_o, H_e, uH_e, A_r, and F_{IS}) and latitude or longitude ($p > 0.05$), except for H_o, which was significantly correlated with latitude. However, some genetic parameters, such as H_o, H_e, and uH_e exhibited significant and negative correlations with the annual mean temperature and annual precipitation variables ($p < 0.05$), which implied that there were higher genetic diversities in populations that grew in regions with lower temperatures and precipitation. Furthermore, because the temperature and precipitation were highly correlated ($R^2 = 0.75$, $p = 0.012$), we performed principle component analysis (PCA, Figure S1; Table S1) of our climate data to reduce its dimensionality, and the results indicated that the component 1 (PCA1, Figure S1; Table S1) could explain 88.68% of the variance of the climate predictors. We then calculated relationships between genetic variables and PCA1, which also showed that the H_o, H_e, and uH_e parameters were significant and negatively correlated with PCA1 ($p < 0.05$).

Table 7. Spearman's correlations between genetic diversity parameters and geoclimatic variables, which were latitude, longitude, annual mean temperature (T_{mean}), mean temperature of the warmest month (T_{max}), mean temperature of the coldest month (T_{min}), annual precipitation (Prec), and the component 1 of principle component analysis of the climate data (PCA1). Significant values are indicated with * ($0.01 < p < 0.05$) and ** ($p < 0.01$).

Parameter	Latitude	Longitude	T_{mean}	T_{max}	T_{min}	*Prec*	PCA1
N_a	−0.202	0.018	0.202	−0.092	0.202	0.128	0.202
N_e	0.464	0.214	−0.679	−0.929 **	−0.643	−0.750	−0.643
I	0.321	0.214	−0.500	−0.893 **	−0.393	−0.679	−0.536
H_o	0.901 **	0.631	−0.811 *	−0.432	−0.757 *	−0.811 *	−0.847 *
H_e	0.714	0.286	−0.786 *	−0.786 *	−0.714	−0.929 *	−0.821 *
uH_e	0.714	0.286	−0.786 *	−0.786 *	−0.714	−0.929 *	−0.821 *
A_r	−0.250	−0.071	0.214	−0.179	0.179	0.107	0.179
F_{IS}	−0.536	−0.321	0.357	−0.179	0.393	0.286	0.429

4. Discussion

4.1. Genetic Diversity

To the best of our knowledge, this study was the first to use SSR molecular markers to evaluate the genetic diversity and structure of natural *P. koraiensis* populations in China. Thus, our results provide essential insight into the potential of this species to adapt to environmental changes and could play an important role in future forest management and conservation of its genetic resources.

Our results clearly showed that the level of genetic diversity in the natural populations of *P. koraiensis* was very high, and the overall H_e (0.610) was similar to that in most *Pinus* species with high levels of genetic diversity when evaluated by SSRs, including *Pinus thunbergii* (H_e = 0.782) [33], *Pinus sylvestris* (H_e = 0.586) [34], and *Pinus strobus* (H_e = 0.531) [35]. Although different SSR markers could affect the comparisons, most studies still use H_e to compare the level of genetic diversity [33–36]. The main reason for the high genetic diversity observed in *P. koraiensis* may be associated with its specific characteristics such as a long life cycle, an outcrossing mating system, wind pollination, and high fecundity [37–39]. Another factor that may have contributed to the high levels of genetic diversity was the large geographic range, which had large differences in climatic and habitat conditions. Although *P. koraiensis* are wind-pollinated and have an outcrossing mating system, all seven populations showed highly negative F_{IS} values. We presumed that this result could be due to different causes, such as the fusion of formerly isolated populations, adaptive advantage of heterozygote individuals, or their critical habitat destruction, therefore resulting in non-random mating between individuals within populations.

In addition, for *P. koraiensis*, the level of genetic diversity in this study was higher than that in previous reports employing other means of detection, such as allozymes (H_e = 0.183) [2], RAPD molecular markers (H_e = 0.169) [20], and ISSR molecular markers (H_e = 0.601) [3]. These differences may be due to the different numbers and types of genetic detection technologies used in the studies or to the different populations and biased sample sizes. Meanwhile, we found that the northern-most populations DL and FL had slightly higher genetic diversity than the other populations, which was also different from the findings of Feng [4] and Kim [20]. However, they seem to have low A_r values, which suggests these two populations may have experienced a recent bottleneck, as a faster reduction in the number of alleles and a concomitantly slight reduction of gene diversity have been generally observed in recently bottlenecked populations [40,41].

4.2. Genetic Differentiation and Population Structure

According to Wright [42], the level of genetic differentiation among populations is low when the coefficient of genetic differentiation (F_{ST}) is less than 0.25. Our results showed that the genetic differentiation level of *P. koraiensis* was very low (F_{ST} ranging from 0.007 to 0.021), and the AMOVA results showed that only 2.35% of the total genetic variation occurred among populations, which was consistent with conifers often showing low levels of genetic differentiation among populations [35,36,43,44]. Due to their wind pollination and high out-crossing rates, *Pinus* species, such as *P. koraiensis* often exhibit high gene flow among populations [45–47] and, consequently, a low level of genetic differentiation.

The STRUCTURE analysis clustered the seven natural populations into two main groups, with the two northern populations forming one group and the five southern populations forming another group. The main reason for this clustering might be that the two northern populations (DL and FL), which are located in the Xiaoxinganling Mountains, are the closest geographically and belong to the same range continuum (with the absence of a physical barrier between them), leading to their tendency to cluster together. Meanwhile, the other five populations, which are located in the Changbai Mountains and adjacent highlands, also had barrier-free gene flow and formed another genetic cluster. Furthermore, in this study, the northern populations were collected from regions that had similar climates, and the genetic parameters were significantly and negatively correlated with climatic factors (i.e., T_{mean} and *Prec*) (Table 7). This implied that geographical and climatic factors could result in

strong and discrete genetic differentiation, which has been found in other studies [36,48]. In addition, human activities might also have an impact on the populations [49,50]. The northern populations were located in nature reserves, and their habitat was protected and relatively intact; however, the southern region had long been deforested by the activities of local villagers. Nevertheless, based on our data, we conclude that *P. koraiensis* was differentiated into two groups on the basis of its current population genetic structure.

4.3. Implications for Conservation

The main genetic concern in the conservation of an endangered species is to find suitable strategies for maintaining its current genetic diversity and ensuring its long-term evolutionary potential [18,51–53]. Therefore, based on the genetic diversity and population structure of *P. koraiensis* in this study, several necessary approaches should be taken for conservation by the Chinese government. First of all, considering the high genetic diversity of *P. koraiensis* populations in our study, an in situ conservation strategy should be carried out to protect their habitats, prohibit timber harvesting, and implement sustainable management regimes in order to maintain the genetic diversity of this species. Our results also showed that there are two genetic clusters, so managers should not move seeds between the northern and southern areas if they want to preserve the genetic distinctiveness of the different clusters. Furthermore, some breeding programs such as seed orchards should be established to produce seeds with high genetic diversity, which is another way to maintain the genetic diversity of this species.

5. Conclusions

Pinus koraiensis is a conifer species of ecological and economic importance in Northeast China which has been excessively exploited in recent years. In order to protect and manage the genetic resources of this important species efficiently, greater knowledge of its genetic diversity and population structure is needed. In the present study, we assessed the genetic diversity and population structure of seven natural populations representing most of the *P. koraiensis* range by using nine nuclear simple sequence repeats. First, we found that the level of genetic diversity in the natural populations was still high, and the level of genetic differentiation among populations was very low. Second, we detected a genetic population structure and a separation between northern and southern populations. Also, genetic variation within populations followed a geographic pattern. Taken together, our results will inform efforts for the conservation and management of *P. koraiensis* and provide guidance for future studies of population genetics and breeding programs.

Supplementary Materials: The following are available online at http://www.mdpi.com/1999-4907/11/1/39/s1, Figure S1: The loadings of each climate variable of PCA, Table S1: Principle component analysis of the climate data (PCA).

Author Contributions: Experimental design, Y.W.T. and L.Q.; investigation, Y.W.T., C.R.M., and Y.W.; formal analysis, Y.W.T. and W.M.Z.; funding acquisition, L.M.D. and L.Q.; project administration, L.Z., D.P.Y., and L.Q.; supervision, L.M.D.; writing—original draft, Y.W.T.; writing—review & editing, Y.W.T., B.J.L., L.M.D., and L.Q. All the co-authors have read and approved the submitted version of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by two projects of National Natural Science Foundation of China, grant number "41877549" and "41701052". This work is also supported by CFERN & BEIJING TECHNO SOLUTIONS Award Funds on excellent academic achievements.

Acknowledgments: We would like to acknowledge Qi Zhu, Jian Wu, Yang Tang, Hong Chen, and all the people from the Forestry Bureau of seven counties that helped us with the sampling.

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

References

1. Ma, J.L.; Zhang, L.W.; Cheng, D.; Li, J.W. Geographic distribution of *Pinus koraiensis* in the world. *J. Northeast. For. Univ.* **1992**, *20*, 40–47. (In Chinese)
2. Potenko, V.V.; Velikov, A.V. Genetic diversity and differentiation of natural populations of *Pinus koraiensis* (Sieb. et Zucc.) in Russia. *Silvae Genet.* **1998**, *47*, 202–208.
3. Feng, F.J.; Sui, X.; Chen, M.M.; Zhao, D.; Han, S.J.; Li, M.H. Mode of pollen spread in clonal seed orchard of *Pinus koraiensis*. *J. Biophys. Chem.* **2010**, *1*, 33–39. [CrossRef]
4. Feng, F.J.; Han, S.J.; Wang, H.M. Genetic diversity and genetic differentiation of natural *Pinus koraiensis* population. *J. For. Res.* **2006**, *17*, 21–24. [CrossRef]
5. Aizawa, M.; Kim, Z.S.; Yoshimaru, H. Phylogeography of the Korean pine (*Pinus koraiensis*) in northeast Asia: Inferences from organelle gene sequences. *J. Plant Res.* **2012**, *125*, 713–723. [CrossRef] [PubMed]
6. Yu, D.P.; Zhou, L.; Zhou, W.M.; Ding, H.; Wang, Q.W.; Wang, Y.; Wu, X.Q.; Dai, L.M. Forest management in Northeast China: History, problems, and challenges. *Environ. Manag.* **2011**, *48*, 1122–1135. [CrossRef] [PubMed]
7. Qi, L.; Yang, J.; Yu, D.; Dai, L.; Contrereas, M. Responses of regeneration and species coexistence to single-tree selective logging for a temperate mixed forest in eastern Eurasia. *Ann. For. Sci.* **2016**, *73*, 449–460.
8. Sun, Y.; Zhu, J.; Sun, O.J.; Yan, Q. Photosynthetic and growth responses of *Pinus koraiensis* seedlings to canopy openness: Implications for the restoration of mixed-broadleaved Korean pine forests. *Environ. Exp. Bot.* **2016**, *129*, 118–126. [CrossRef]
9. Chen, X.W.; Zhou, G.S.; Zhang, X.S. Spatial characteristics and change for tree species along the North East China Transect (NECT). *Plant Ecol.* **2003**, *164*, 65–74. [CrossRef]
10. Hughes, A.R.; Inouye, B.D.; Johnson, M.T.J.; Underwood, N.; Vellend, M. Ecological consequences of genetic diversity. *Ecol. Lett.* **2008**, *11*, 609–623. [CrossRef]
11. Gapare, W.J. Merging applied gene conservation activities with advanced generation breeding initiatives: A case study of *Pinus radiata* D. Don. *New For.* **2014**, *45*, 311–331. [CrossRef]
12. Hamrick, J.L.; Godt, M.J.W.; Sherman-Broyles, S.L. Factors influencing levels of genetic diversity in woody plant species. *New For.* **1992**, *6*, 95–124. [CrossRef]
13. Shi, X.M.; 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]
14. Potter, K.M.; Jetton, R.M.; Bower, A.; Jacobs, D.F.; Man, G.; Hipkins, V.D.; Westwood, M. Banking on the future: Progress, challenges and opportunities for the genetic conservation of forest trees. *New For.* **2017**, *48*, 153–180. [CrossRef]
15. Reed, D.H.; Frankham, R. Correlation between fitness and genetic diversity. *Conserv. Biol.* **2003**, *17*, 230–237. [CrossRef]
16. Eltaher, S.; Sallam, A.; Belamkar, V.; Emara, H.A.; Nower, A.A.; Salem, K.F.M. Genetic diversity and population structure of $F_{3:6}$ Nebraska winter wheat genotypes using genotyping-by-sequencing. *Front. Genet.* **2018**, *9*, 76. [CrossRef]
17. Sutherland, W.J.; Albon, S.D.; Allison, H.; Armstrong-Brown, S.; Bailey, M.J.; Brereton, T.; Boyd, I.L.; Carey, P.; Edwards, J.; Gill, M.; et al. The identification of priority policy options for UK nature conservation. *J. Appl. Ecol.* **2010**, *47*, 955–965. [CrossRef]
18. Guo, H.Y.; Wang, Z.L.; Huang, Z.; Chen, Z.; Yang, H.B.; Kang, X.Y. Genetic Diversity and Population Structure of *Alnus cremastogyne* as Revealed by Microsatellite Markers. *Forests* **2019**, *10*, 278. [CrossRef]
19. Thakur, S.; Choudhary, S.; Singh, A.; Ahmad, K.; Sharma, G.; Majeed, A.; Bhardwaj, P. Genetic diversity and population structure of *Melia azedarach* in North-Western Plains of India. *Trees* **2016**, *30*, 1483–1494. [CrossRef]
20. Kim, Z.S.; Hwang, J.W.; Lee, S.W.; Yang, C.; Gorovoy, P.G. Genetic variation of Korean pine (*Pinus koraiensis* Sieb. et Zucc.) at allozyme and RAPD markers in Korea, China and Russia. *Silvae Genet.* **2005**, *54*, 235–246. [CrossRef]
21. Feng, F.J.; Chen, M.M.; Zhang, D.D.; Sui, X.; Han, S.J. Application of SRAP in the genetic diversity of *Pinus koraiensis* of different provenances. *Afr. J. Biotechnol.* **2009**, *8*, 1000–1008.

22. Yu, J.H.; Chen, C.M.; Tang, Z.H.; Yuan, S.S.; Wang, C.J.; Zu, Y.G. Isolation and characterization of 13 novel polymorphic microsatellite markers for *Pinus koraiensis* (Pinaceae). *Am. J. Bot.* **2012**, *99*, 421–424. [CrossRef] [PubMed]
23. Jia, D.; Zhen, Z.; Zhang, H.G.; Tang, J.H. EST-SSR marker development and transcriptome sequencing analysis of different tissues of Korean pine (*Pinus koraiensis* Sieb. et Zucc.). *Biotechnol. Biotechnol. Equip.* **2017**, *31*, 679–689.
24. 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] [PubMed]
25. Weir, B.S.; Cockerham, C.C. Estimating F-statistics for the analysis of population structure. *Evolution* **1984**, *38*, 1358–1370.
26. Goudet, J. FSTAT (version 1.2): A computer program to calculate F-statistics. *J. Hered.* **1995**, *86*, 485–486. [CrossRef]
27. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019.
28. Trouet, V.; Van Oldenborgh, G.J. KNMI climate explorer: A web-based research tool for high-resolution paleoclimatology. *Tree Ring Res.* **2013**, *69*, 3–13. [CrossRef]
29. Hubisz, M.J.; Falush, D.; Stephens, M.; Pritchard, J.K. Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Resour.* **2009**, *9*, 1322–1332. [CrossRef]
30. Evanno, G.; Regnaut, S.; Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* **2010**, *14*, 2611–2620. [CrossRef]
31. Jakobsson, M.; Rosenberg, N.A. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **2007**, *23*, 1801. [CrossRef]
32. Excoffier, L.; Smouse, P.E.; Quattro, J.M. Analysis of molecular variance inferred from metric distances among DNA haplotypes-application to human mitochondrial-DNA restriction data. *Genetics* **1992**, *131*, 479–491. [PubMed]
33. Iwaizumi, M.G.; Miyata, S.; Hirao, T.; Tamura, M.; Watanabe, A. Historical seed use and transfer affects geographic specificity in genetic diversity and structure of old planted *Pinus thunbergii* populations. *For. Ecol. Manag.* **2018**, *408*, 211–219. [CrossRef]
34. Toth, E.G.; Vendramin, G.G.; Bagnoli, F.; Cseke, K.; Hoehn, M. High genetic diversity and distinct origin of recently fragmented Scots pine (*Pinus sylvestris* L.) populations along the Carpathians and the Pannonian Basin. *Tree Genet. Genomes* **2017**, *13*, 47. [CrossRef]
35. Mandak, B.; Hadincova, V.; Mahelka, V.; Wildova, R. European Invasion of North American *Pinus strobus* at Large and Fine Scales: High Genetic Diversity and Fine-Scale Genetic Clustering over Time in the Adventive Range. *PLoS ONE* **2013**, *8*, e68514. [CrossRef]
36. Xu, Y.L.; Cai, N.H.; Woeste, K.; Kang, X.Y.; He, C.Z.; Li, G.Q.; Chen, S.; Duan, A.A. Genetic diversity and population structure of *Pinus yunnanensis* by simple sequence repeat markers. *For. Sci.* **2016**, *62*, 38–47. [CrossRef]
37. Hamrick, J.L.; Linhart, Y.B.; Mitton, J.B. Relationships between life-history characteristics and electrophoretically detectable genetic-variation in plants. *Annu. Rev. Ecol. Syst.* **1979**, *10*, 173–200. [CrossRef]
38. Hellmann, J.J.; Pineda-Krch, M. Constraints and reinforcement on adaptation under climate change: Selection of genetically correlated traits. *Biol. Conserv.* **2007**, *137*, 599–609. [CrossRef]
39. Gaudeul, M.; Taberlet, P.; Till-Bottraud, I. Genetic diversity in an endangered alpine plant, *Eryngium alpinum* L. (Apiaceae), inferred from amplified fragment length polymorphism markers. *Mol. Ecol.* **2000**, *9*, 1625–1637. [CrossRef]
40. Karmin, M.; Saag, L.; Vicente, M.; Sayres, M.A.W.; Jarve, M.; Talas, U.G.; Rootsi, S.; Ilumaee, A.M.; Maegi, R.; Mitt, M.; et al. A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome Res.* **2015**, *25*, 459–466. [CrossRef]
41. Peery, M.Z.; Kirby, R.; Reid, B.N.; Stoelting, R.; Doucet-Beer, E.; Robinson, S.; Vasquez-Carrillo, C.; Pauli, J.N.; Palsboll, P.J. Reliability of genetic bottleneck tests for detecting recent population declines. *Mol. Ecol.* **2012**, *21*, 3403–3418. [CrossRef]
42. Wright, S. The Interpretation of Population Structure by F-Statistics with Special Regard to Systems of Mating. *Evolution* **1965**, *19*, 395–420. [CrossRef]

43. Belletti, P.; Ferrazzini, D.; Piotti, A.; Monteleone, I.; Ducci, F. Genetic variation and divergence in Scots pine (*Pinus sylvestris* L.) within its natural range in Italy. *Eur. J. For. Res.* **2012**, *131*, 1127–1138. [CrossRef]
44. Iwaizumi, M.G.; Tsuda, Y.; Ohtani, M.; Tsumura, Y.; Takahashi, M. Recent distribution changes affect geographic clines in genetic diversity and structure of *Pinus densiflora* natural populations in Japan. *For. Ecol. Manag.* **2013**, *304*, 407–416. [CrossRef]
45. Nybom, H. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Mol. Ecol.* **2004**, *13*, 1143–1155. [CrossRef]
46. 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]
47. Rubio-Moraga, A.; Candel-Perez, D.; Lucas-Borja, M.E.; Tiscar, P.A.; Vinegla, B.; Linares, J.C.; Gómez-Gómez, L.; Ahrazem, O. Genetic diversity of *Pinus nigra* Arn. populations in southern Spain and northern Morocco revealed by inter-simple sequence repeat profiles. *Int. J. Mol. Sci.* **2012**, *13*, 5645–5658. [CrossRef]
48. Wang, B.; Mao, J.F.; Zhao, W.; Wang, X.R. Impact of geography and climate on the genetic differentiation of the subtropical pine *Pinus yunnanensis*. *PLoS ONE* **2013**, *8*, e67345. [CrossRef]
49. Young, A.; Merriam, H. Effects of forest fragmentation on the spatial genetic structure of *Acer saccharum* Marsh.(sugar maple) populations. *Heredity* **1994**, *72*, 201. [CrossRef]
50. Aldrich, P.R.; Hamrick, J.L.; Chavarriaga, P.; Kochert, G. Microsatellite analysis of demographic genetic structure in fragmented populations of the tropical tree *Symphonia globulifera*. *Mol. Ecol.* **1998**, *7*, 933–944. [CrossRef]
51. Forest, F.; Grenyer, R.; Rouget, M.; Davies, T.J.; Cowling, R.M.; Faith, D.P.; Balmford, A.; Manning, J.C.; Proches, S.; van der Bank, M.; et al. Preserving the evolutionary potential of floras in biodiversity hotspots. *Nature* **2007**, *445*, 757. [CrossRef]
52. Stojnić, S.; Avramidou, E.V.; Fussi, B.; Westergren, M.; Orlović, S.; Matović, B.; Trudić, B.; Kraigher, H.; Aravanopoulos, F.A.; Konnert, M. Assessment of Genetic Diversity and Population Genetic Structure of Norway Spruce (*Picea abies* (L.) Karsten) at Its Southern Lineage in Europe. Implications for Conservation of Forest Genetic Resources. *Forests* **2019**, *10*, 258. [CrossRef]
53. Carabeo, M.; Simeone, M.C.; Cherubini, M.; Mattia, C.; Chiocchini, F.; Bertini, L.; Caruso, C.; La Mantia, T.; Villani, F.; Mattioni, C. Estimating the genetic diversity and structure of *Quercus trojana* Webb populations in Italy by SSRs: Implications for management and conservation. *Can. J. For. Res.* **2017**, *47*, 331–339. [CrossRef]