

Study on Genetic Diversity of *Phellodendron amurense* based on cpDNA

Hongsheng Yang¹, Donghong Yang^{1*}, Xuewen Yang¹, Qingbo Zhou¹, Haitao Cheng¹, Lili Li¹ and Decai Liu^{1*}

¹Jiamusi University, Jiamusi City, Heilongjiang Province, 154007, China

Abstract. In this study, 3 haplotypes were found in populations of *Phellodendron amurense* based on two combined cpDNA regions (psbA-trnH and trnT-trnL). Nucleotide diversity and haplotype diversity were 0.43×10^{-3} and 0.41, respectively at the level of species. The AMOVA revealed that only 8.53% of the variation was explained by differences among geographical groups, whereas inter-population and intra-population differences explained 18.32% and 71.35% of the variation, respectively. Phylogeographical relationships showed that all haplotypes were clustered into two lineages. Haplotype H₁ and H₂ cluster together, and Haplotype H₃ composed a group. TCS network of haplotypes showed that haplotype H₁ located in the center of the lineage, and it appears to be an ancestral haplotype. So we hypothesized that Northeast China populations and North China populations had a common origin. The mismatch distribution of this species suggested that all populations and populations in North China had not undergone recent expansion, but populations in Northeast China had undergone recent expansion. The results were consistent with the results of Tajima's D and Fu's and Li's D test.

1 Introduction

Phellodendron amurense is mainly distributed in Northeast China. As an important timber tree and medicinal plant, the wild *P. amurense* has been cut down in large quantities, and its quantity is sharply reduced. In 1987, *P. amurense* was classified as a endangered species in the list of rare and endangered protected plants in China. *P. amurense* is a relic species of the ancient tropical flora from the Tertiary period, which is known as one of the top three broad-leaved hardwood trees and an important accompanying tree species. The study of its genetic diversity and spatial distribution pattern is of great scientific significance for understanding the population history of this species and species protection under the background of climate change.

The genetic diversity of *P. amurense* was also reported. Yan et al. [1] Analyzed 10 wild populations of *P. amurense* using single molecular marker AFLP technology. The results showed that the genetic diversity of *P. amurense* at the species level was higher than that at the population level, and the genetic variation might be related to the habitat. Based on the microsatellite data of *P. amurense*, Wan et al. [2] used the model to simulate the priority reserves of *P. amurense* and the distribution changes under the background of climate change. The results showed genetic diversity of *P. amurense* was positively correlated with the habitat suitability. Yang et al. [3] analyzed the genetic diversity of 17 natural populations of *P. amurense* using ISSR markers. Mantel test showed that there was a significant positive correlation between geographical distance and

genetic distance among populations of *P. amurense*. The above results made us have a certain understanding of the genetic diversity of *P. amurense*, and lay a foundation for further study on population genetics of *P. amurense*. In this paper, the genetic diversity of 9 natural populations of *P. amurense* were analyzed by using the chloroplast noncoding region sequences of maternal inheritance. The results revealed the distribution of genetic diversity of *P. amurense*, and provided reasonable strategies for population protection of *P. amurense*.

2 Materials and Methods

2.1 Experimental Materials

Nine populations, BJ, QL, HR, FS, WQ, WC, LK, YC and HL, were selected as the research objects. Eight samples were randomly selected from each population for cpDNA analysis.

2.2 primer screening

Two pairs of cpDNA fragments (psbA-trnH and trnT-trnL) were sequenced according to the reference [4].

2.3 PCR amplification system and reaction procedure

After PCR amplification, the products were detected by 1% agarose gel electrophoresis, and the samples were

* Co-Corresponding author: 769649064@qq.com;109952111@qq.com

sent to Jilin Kumi Biotechnology Co., Ltd. for sequencing.

2.4 Data Analysis

The sequences were processed and analyzed using software such as Bioedit, DNASTar, DnaSP V5, Permut 2.0, TCS 1.2, Modeltest 3.7, MrBayes3.1.2, and Arlequin V3.5.

3 results and analysis

3.1 sequence variation and haplotype distribution

After sequencing, the length of psbA-trnH sequence was 455 bp, and one polymorphic site (1 inversion) was detected. The length of trnT-trnL sequence was 926 bp, and one polymorphic site was detected, which was a large insertion deletion with 22 bases. After splicing and combining the two chloroplast gene fragments, the combined length of the chloroplast fragments was 1381 bp, and 3 haplotypes were found.

The H₁ haplotype was widely distributed in North China population and Northeast China population, while the H₂ haplotype was unique to northeast China population. And the H₃ haplotype was common in North China population and Changbaishan population. Among 9 populations of *P. amurense*, it was found that there were 3 haplotypes in 3 populations, 2 haplotypes in 5 populations, and 1 haplotype in only 1 population (YC).

3.2 genetic diversity

DnaSP analysis showed that in nine populations of *P. amurense* the nucleotide polymorphism (π) ranged from 0 to 0.92×10^{-3} , with an average of 0.43×10^{-3} (Table 1). WQ population had the highest nucleotide polymorphism (0.92×10^{-3}). The other two populations in Changbai Mountain including FS population and HR population, also had high nucleotide polymorphism, while YC population had the lowest nucleotide polymorphism (0). The haplotype polymorphism (H) ranged from 0 to 0.75 with an average of 0.41. WQ, FS and HR populations all had 3 haplotypes with high haplotype polymorphism, with values of 0.92, 0.79 and 0.59, respectively. YC population had no haplotype polymorphism, showing a single haplotype.

Table 1 Genetic diversity of cpDNA haplotype

Population	Sample number	H	$\pi \times 10^{-3}$	Frequency of haplotype
YC	8	0	0	H ₁ (100)
DFH	8	0.21	0.19	H ₁ (91),H ₂ (9)
LK	8	0.49	0.43	H ₁ (85),H ₂ (15)
WC	8	0.52	0.58	H ₁ (80),H ₂ (20)
WQ	8	0.75	0.92	H ₁ (55),H ₂ (25),H ₃ (20)
FS	8	0.71	0.79	H ₁ (50),H ₂ (10),H ₃ (40)
HR	8	0.65	0.59	H ₁ (12),H ₂ (18),H ₃ (70)
QL	8	0.23	0.21	H ₁ (28),H ₃ (72)
BJ	8	0.11	0.16	H ₁ (11),H ₃ (89)

Permut software analysis showed that the total genetic diversity Ht (Se) of *P. amurense* based on cpDNA sequence was 0.471(0.0632). The average genetic diversity within population Hs(Se) was 0.323(0.043). The population genetic differentiation coefficient(G_{st}) was 0.314, and N_{st} was 0.351. U statistical test showed that there was no significant difference between G_{st} and N_{st} ($P > 0.05$), which indicated that the geographical distribution and phylogeny of the haplotype had not significant correlation. The gene flow(N_m) among populations calculated by the differentiation coefficient was 0.546.

AMOVA results (Table 2) showed that the genetic variation within the population accounted for 73.15% of the total variation, and the genetic variation among populations accounted for 18.32% of the total variation. And the variation between North China and Northeast China was 8.53%. The F_{st} value ($F_{st}=0.2685$) indicated that the main genetic differentiation occurred within the population, while the differentiation between regions and populations was small.

Table 2 Analysis of molecular variance for populations of *P. amurense* based on cpDNA

Source of variation	D.f.	Sum of squares	Variance components	Percentage of variation
Between groups	1	18.717	0.20188	8.53
Between populations	7	60.679	0.43358	18.32
Within group				
Within population	135	233.688	1.73102	73.15

3.3 Phylogeographic analysis based on haplotype

The best model selected by modeltest 3.7 was TVM+I+G. Taking *Phellodendron chinense* as the outgroup, the phylogenetic tree of three cpDNA haplotypes about *P. amurense* was constructed. The haplotypes were divided into two branches. The haplotype H1 and H2 were divided into one group, and the haplotype H3 was divided into one group. This result was basically consistent with the reticular tree constructed by TCS (Fig.1). The haplotype H1 was in the center, and the haplotypes H2 and H3 were located on both sides of H1.

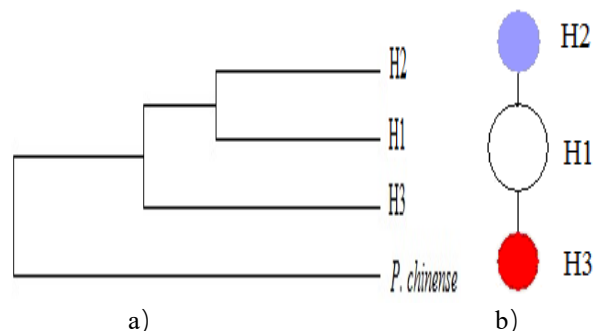
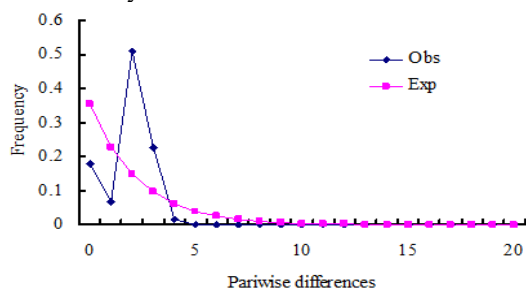


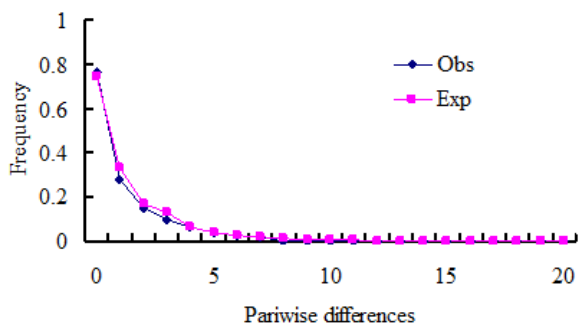
Fig.1 Phylogenetic tree of the 3 haplotypes constructed by Bayesian analysis and TCS networks

3.4 mismatch distribution analysis and neutral test.

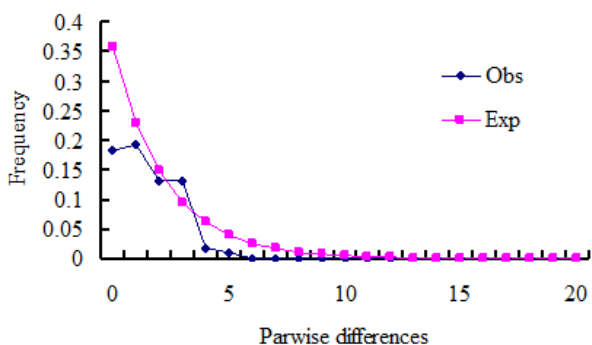
In order to test whether the *P. amurensis* population had experienced the expansion event, the DnaSP v5 software was used to analyze the mismatch distribution of the psbA-trnH and trnT-trnL combined fragments (Fig.2). At the species level of *P. amurensis*, the expected curve did not coincide with the actual observed curve, indicating that the *P. amurensis* population had not experienced a rapid expansion event at the species level. The results of mismatch distribution analysis of two population groups in Northeast China and North China showed that the expectation curve and observation curve of North China were roughly in line with the species level, which did not conform to the rapid expansion event. However, the actual observed expectation curve and the expectation curve in the Northeast population group had a good coincidence, which accorded with the expansion model. The results of neutral test were consistent with those of mismatch analysis.



a) Nine populations of *P. amurensis*



b) Northeast population



c) North China population

Fig. 2 Mismatch Distribution of the number of pairwise nucleotide differences for chloroplast DNA sequence data

4 Discussion

In this study, a total of 3 haplotypes were detected using two pairs of cpDNA fragments, and the total genetic diversity was low ($H_t = 0.471$). The population genetic differentiation coefficient (G_{st}) was 0.314, and the genetic variation mainly existed within the population. AMOVA analysis also revealed similar results ($F_{st} = 0.2685$). The difference between N_{st} and G_{st} of *P. amurensis* was not significant, and no obvious phylogenetic structure was detected. This may be related to the unique haplotype. Coincidentally, the phylogenetic structure of *Acer mono* in the same region based on cpDNA was not detected [5]. Two haplotypes of *Juglans mandshurica* and *Acer mono* with similar distribution areas were detected in both Yanshan and Northeast populations [5-6]. *A. mono* had its own haplotype in Yanshan and Northeast population, while the southern population in Northeast China shared the same haplotype with Yanshan population [5]. For *J. mandshurica*, the Yanshan population and the Northeast population also had their own haplotype, and two mixed haplotypes appeared in the Northeast population [6]. In this study, 3 haplotypes were detected, and the number of haplotype was relatively small. The haplotype diversity of tree species in the north was less than that in the south, which was a common phenomenon.

The topological structure of TCS based on cpDNA haplotypes showed that the North China population and the Northeast population were significantly differentiated. However, there was no significant difference between G_{st} and N_{st} ($P > 0.05$), indicating that the phylogenetic structure was not obvious. From the phylogenetic tree of haplotypes, it showed a relatively consistent topological structure. The haplotype H2 that was unique to the Northeast population formed a group with the widely distributed haplotype H1, and H3 formed a group alone. The H1 haplotype was shared by the Northeast population and the North China population. The TCS topology shows that the H1 haplotype was in the center, and the H2 and H3 haplotypes were on both sides. Generally, the innermost haplotype in the network diagram was the ancestral haplotype [7], so H1 was the ancestor haplotype, and the H1 haplotype was differentiated into H2 and H3. The H3 haplotype was shared by the North China population and the Changbai Mountain population, which indicated that the North China population and the Changbai Mountain population had a common origin.

5 Conclusion

In this study, the psbA-trnH and trnT-trnL regions of cpDNA in *P. amurensis* were amplified and sequenced for 72 individuals in 9 populations. 3 haplotypes were found in populations of *P. amurensis* based on two combined cpDNA regions (psbA-trnH and trnT-trnL) that had a length of 1381 bp. Haplotype H1 widely distributed was found in all populations. Haplotype H2 was native in Northeast China. Haplotype H3 was interspersed in North China and the Changbai Mountain. Nucleotide diversity and haplotype diversity were

0.43×10⁻³ and 0.41, respectively at the level of species. The AMOVA revealed that only 8.53% of the variation was explained by differences among geographical groups, whereas inter-population and intra-population differences explained 18.32% and 71.35% of the variation, respectively. Phylogeographical relationships showed that all haplotypes were clustered into two lineages. Haplotype H1 and H2 cluster together, and Haplotype H3 composed a group. TCS network of haplotypes showed that haplotype H1 located in the center of the lineage, and it appears to be an ancestral haplotype. So We hypothesized that Northeast China populations and North China populations had a common origin. The mismatch distribution of this species suggested that all populations and populations in North China had not undergone recent expansion, but populations in Northeast China had undergone recent expansion. The results were consistent with the results of Tajima's D and Fu's and Li's D test.

Acknowledgement

Basic Scientific Research Project of Heilongjiang Provincial Universities (2019-KYYWF-1401); Youth Innovative Talents Training Project of Jiamusi University (JMSUQP2020006) .

Reference

1. Yan Z F, Zhang B G, Zhang Z, et al. (2006) Genetic diversity in wild populations of *Phellodendron amurense*, a rare endangered medicinal plant, detected by AFLP. Biodiversity science, 14(6):488-497.
2. Wan J Z, Wang C J, Yu J H, et al. (2014) Model-based conservation planning of the genetic diversity of *Phellodendron amurense* due to climate change. Ecology and Evolution, 4(14):2884-2900.
3. Yang H S, Li F H, Wang C B, et al. (2016) ISSR analysis on genetic diversity in natural populations of rare and endangered species *Phellodendron amurense*. Journal of Northeast Agricultural University, 47(6): 26- 32.
4. Wang L J, Bao L, Wang H F, et al. (2014) Study on the genetic diversity of *Phellodendron amurense*. China Science Paper, 9(12):1397-1401.
5. Liu C P. (2013) Study on genetic structure and phylogeography of *Acer mono* natural populations. Doctoral Dissertation of Northeast Forestry University.
6. Bai W N, Liao W J, Zhang D Y. (2010) Nuclear and chloroplast DNA phylogeography reveal two refuge areas with asymmetrical gene flow in a temperate walnut tree from East Asia. New Phytologist, (188): 892-901.
7. Golding G B. (1987) The detection of deleterious selection using ancestors inferred from a phylogenetic history. Genetics Research, (49): 71-82.