

Genetic diversity analysis and EST-SSR fingerprinting construction of natural *Populus euphratica* and *Populus pruinosa* populations

Xiang X. Chen

Tarim University;

Zhong S. Gai

Tarim University;

Jun T. Zhai

Tarim University;

Xue F. Guo

Tarim University;

Zhi H. Wu

Zhejiang Normal University

Zhi J. Li (✉ lizhijun0202@126.com)

Tarim University;

Research Article

Keywords: *Populus euphratica*, *Populus pruinosa*, Genetic diversity, EST-SSR markers, Fingerprinting

Posted Date: November 4th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-2226822/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Populus euphratica and *Populus pruinosa* are important tree species in northwest China. In this study, the codominant EST-SSR markers related to drought resistance genes was used to reveal the genetic diversity of *P. euphratica* and *P. pruinosa* populations. The results showed that those populations exhibited rich genetic diversity levels. Compared with other *P. euphratica* and *P. pruinosa* populations respectively, the genetic diversity of populations in southern Xinjiang was all the highest. The average fixed index of different populations indicated that the genetic relationship among *P. euphratica* populations was relatively far, while the relationship among the *P. pruinosa* populations was closer. The molecular variance analysis showed that the variation within populations was bigger than that among populations in the *P. euphratica* and *P. pruinosa* populations respectively. The cluster analysis showed that *P. euphratica* and *P. pruinosa* populations could be mainly classified into two clades represented by southern and northern Xinjiang respectively. The population structure and principal co-ordinates analysis indicated that the 47 natural *P. euphratica* populations have been divided into 4 clades, represented by the two populations in southern or southern Xinjiang respectively. The 17 natural *P. pruinosa* populations had been divided into 4 clades, consisted by the three County in southern Xinjiang, and Nilka County in northern Xinjiang. At the same time, The EST-SSR fingerprints of 47 natural *Populus euphratica* and 17 natural *Populus pruinosa* have been constructed successfully. This study laid the foundation for the later priority protection of the genetic diversity populations of excellent drought resistance.

Introduction

Populus euphratica Oliv. (*P. euphratica*) and *Populus pruinosa* Schrenk (*P. pruinosa*) belong to the Sect. *Turanga* Bge in *Populus* of Salicaceae. *P. euphratica* is commonly known as variable leaf or heterophylly poplar, and mainly distributed in Central and western Asia, followed by north Africa and southern Europe (Wang 1996; Wei 1990). About 61% of the world's *Populus euphratica* forests are distributed in China, while 91.1% of China's natural *Populus euphratica* forests are distributed in Xinjiang (Li 2020). It had an excellent resistance characteristics such as salt alkali resistance, drought resistance, heat resistance, wind and sand resistance (Li et al. 2011), and played an extremely important role in maintaining the fragile ecological environment in desert areas (Cao et al. 2012). Compared with *P. euphratica*, *P. pruinosa* populations are only found in Central Asia and West Asia along the riparian forest, and mainly distributed in the upper reaches of Tarim River, alongside the Yerqiang River and Hotan River in southern Xinjiang (Wei 1990). Plant genetic diversity is the essence of species evolution and one of the cores of conservation genetics. The research of conservation genetics mainly included morphological research, cytological research, physiological and biochemical research and molecular biology research. In recent years, SSR markers were used to analyze the population genetic diversity because of its advantages of good repeatability, high polymorphism, codominance (Luo et al. 2019). EST (expressed sequence tag) was usually less than 1000bp length of cDNA (Dillon et al. 2014) and provided a valuable source for the development of SSR markers (Bérubé et al. 2007; Poncet et al. 2006). EST-SSR (expressed sequence tag SSR) was used to detect the SSR characteristics at the transcriptional level. Compared with ordinary

gSSR (genomic SSR), EST-SSR was the generally conservative and accurate, and became an important tool to find important agronomic traits (Gao et al. 2003; Saha et al. 2004).

The natural *P. euphratica* populations used by SSR markers had high genetic diversity and low population differentiation; Geographical distance had no significant effect on the genetic differentiation of different populations, and the genetic variation among populations was greater than that within population (Wang et al. 2011a). The natural *P. pruinosa* populations also showed the similar results (Zhang et al. 2012). The different *P. euphratica* populations also existed the gene flow from the north and south of Xinjiang. At the same time, the distribution contraction caused by climate oscillation further accelerated the divergence of these regional populations (Zeng et al. 2018). Zhang et al. (Zhang et al. 2011) analyzed the 15 poplar materials by using genomic SSR and EST-SSR markers, and compared two different DNA molecular marker methods. The results showed that EST-SSR markers had better universality than genomic SSR, and the genetic diversity revealed by EST-SSR loci was higher than genomic SSR loci, The results of cluster analysis showed that EST-SSR markers could more accurately reflect the genetic differences and genetic relationships among different poplar species. Of course, molecular markers were powerful tools for the identification of plant species or individuals (Karihaloo 2015). For example, the 286 different potato varieties could be completely identified by using 4 to 5 pairs of SSR markers (Moisan-Thiery et al. 2005). The 14 SSR markers were used to discriminate the 17 most common tomato varieties distributed in Italy (Caramante et al. 2011). The 20 SSR markers were tested with 10 genotypes to construct the unique fingerprints in the different mango varieties (Kumar et al. 2013). The DNA fingerprints of 128 elite oil camellia cultivars were constructed by SSR primer pairs screened from the SSR primer database of oil camellia (Chen et al. 2016). A SSR high-throughput identification system was constructed according to six pairs high efficiency EST-SSR primers to gain the DNA fingerprint spectrum of different *Lolium multifloru* varieties (Liu et al. 2017).

Although sporadic molecular markers have been reported in the *P. euphratica* population research (Dong and Bai 1998; sheng et al. 2008; Wang et al. 2011a), these markers weren't related to the extreme drought resistance genes of *P. euphratica*. For example, Genes such as *WRKY*, *GRAS*, *DREB*, *MYB* and hormone ABA signaling pathway components were widely involved in plant drought resistance (Jing et al. 2017; Ma et al. 2010; Sha et al. 2013; Yan et al. 2012). Therefore, this study identified the drought gene related EST-SSR markers of *P. euphratica* and *P. pruinosa* populations at the whole genome level, and used these markers to study the genetic diversity, constructed fingerprinting of different population samples, which could not only reveal the genetic diversity characteristics in the different geographical distribution areas and identify different populations, but also had an important guiding significance for further mining and breeding gene resources with excellent drought resistance.

Materials And Methods

Sample collection and DNA extraction

According to the longitude and latitude gradient survey sampling strategy from south to north and west to east, the natural *P. euphratica* and *P. pruinosa* forests in Xinjiang, Qinghai, Ningxia, Inner Mongolia and Gansu were investigated in Northwest China (Gai et al. 2021). Finally, 693 samples were collected from 47 natural *P. euphratica* populations and 235 samples were collected from 17 natural *P. pruinosa* populations. The spacing of sample plants in each population should be at least 100m. Pick the leaves at the third node from the base of the current year branches, dry the leaves with silica gel for preservation, send them back to the laboratory for DNA extraction and measure the concentration of the samples (Saghai-Marooft et al. 1984).

Est-ssr Primer Designing And Screening

Based on the published genome sequencing data of *P. euphratica* and *P. pruinosa* species, the SSR sequences was developed by using MISA (microsatellite identification tool). Set the parameter of SSR identification, definition (unit_size, min repeats): 1–10 2–5 3–4 4–3 5–3 6–3, which meant that the minimum repeat numbers of single nucleotide repeat, dinucleotide repeat, trinucleotide repeat, tetranucleotide repeat, pentanucleotide repeat and hexanucleotide repeat was 10, 5, 4, 3, 3 and 3 respectively; Interruptions (max_difference_for_2_ssrs): 100, the distance between the two microsatellites was less than 100bp.

The identification of drought resistance genes of *P. euphratica* included two parts: (1) Integrate the transcriptome data of drought resistance of *P. euphratica* in our research group to identify the genes with differential response in the process of stress. (2) Use the known drought resistant genes and pathways of dicotyledonous model plant Arabidopsis, homology comparison to the drought resistant genes of *P. euphratica* and *P. pruinosa*, and combine the genes from the two sources to construct the drought resistant gene collection of *P. euphratica* and *P. pruinosa*. 10kb upstream and downstream of these drought genes (including the gene itself) were extracted by self-writing script. 500 pairs of EST-SSR primers were developed in batch according to the extracted gene sequence. 133 pairs of suitable primers were screened from the 500 pairs of primers according to the conditions of conventional primer screening and sent to Sangon Biotech company (Shanghai) for primer synthesis.

Genetic Diversity Analysis And Fingerprinting Construction

12 trees were randomly selected from 12 different *P. euphratica* populations for PCR amplification. The specific PCR amplification system was as follows (Tab. S1). The reaction procedure of PCR amplification was as follows: initial denaturation at 95 °C for 3 min; primer denaturation at 95 °C for 15s, primer annealing at 58 °C ~ 60 °C for 15s, primer extension at 72 °C for 1min, 35 cycles; final extension at 72 °C for 10min. Finally, 10 pairs of core primers with clear bands and good specificity were selected from 133 pairs of suitable primers (Tab. S2). A total of 929 samples of *P. euphratica* and *P. pruinosa* were amplified by PCR. The PCR products were detected by 1% agarose gel electrophoresis. Using a marker of 1500bp as a scale, the bright and clear band at the same site were marked as 1, and the absent band

were marked as 0, and 9 indicated no amplification (Xia and Lu 2009). ignore weak or heterozygous bands and build the 0/1 matrix of the original data.

The genetic diversity parameters of *P. euphratica* and *P. pruinosa* populations were evaluated in terms of the number of alleles (N_a), effective number of alleles (N_e), Shannon's information index (H), observed heterozygosity (H_o), expected heterozygosity (H_e), unbiased expected heterozygosity (UHe), fixed index (F), AMOVA (Analysis of Molecular Variance) (Peakall and Smouse 2012), Nei's (1983) population genetic distance (D) and PCoA (principal coordinates analysis) were calculated by using GenALEX 6.5. Use NTSYS-pc 2.10s software to generate UPGMA clustering map (Rohlf 1998) and use STRUCTURE v2.3.4 software to analyse population structure (Earl and Vonholdt 2012). At the same time, the 0/1 matrix dates of *P. euphratica* and *P. pruinosa* was used to construct the fingerprinting of different populations samples.

Results

Genetic diversity analysis

In the natural *P. euphratica* populations (Tab. S3), the results showed the average level of the observed heterozygosity (H_o) was greater than the expected heterozygosity (H_e), indicating that the proportion of heterozygotes in the overall level population was higher than that required to achieve balance. there was excess heterozygotes and the genetic relationship of different was far. From the fixed index (F), the MYX population ($F = 0.002$) was relatively close to 0, indicating that there was little difference between the expected heterozygosity and the observed heterozygosity of the population, which was close to the ideal population. However, In the fixed index (F) (Tab. S4), there were seven populations whose F value was less than 0, and other ten populations were more than 0, indicating that at the population level, the expected heterozygosity was relatively bigger than the observed heterozygosity, there was lack of heterozygotes, excess homozygotes and more close relatives among populations. In general, the average level of genetic diversity of natural *P. euphratica* and *P. pruinosa* populations in southern Xinjiang was higher than that other areas.

Population Genetic Variation

In the analysis of molecular variance (AMOVA) of natural *P. euphratica* population (Table 1), the variation among populations accounted for 23% of the total variation of *P. euphratica*, and the variation within populations accounted for 77% of the total variation. The variation within populations included 11% of the variation among individuals and 66% of the variation within individual. The results showed that the variation within population was the main source of the total variation of natural *P. euphratica* populations. However, the variation among populations accounted for 24% of the total variation of *P. pruinosa*, and the variation within populations accounted for 76% of the total variation (Tab. S5). The variation within populations included 22% of the variation among individuals and 54% of the variation

within individual. The results showed that the variation within population was the main source of the total variation of natural *P. pruinosa* populations.

Table 1
Molecular variance analysis of *Populus euphratica* populations

Source	Degree of freedom (df)	Sum of square (SS)	Mean square (MS)	Variance component	Proportion (%)
Among populations	46	612.319	13.311	0.400	23%
Among individuals	645	984.521	1.526	0.195	11%
Within individual	692	786.000	1.136	1.136	66%
Total	1383	2382.840	-	1.731	100%

In the cluster analysis based the genetic distance of 47 natural *P. euphratica* populations (Fig. S1), when the genetic distance among populations was approximately equal to 0.1, *P. euphratica* populations were mainly divided into two clades, of which *P. euphratica* populations represented by Bazhou, Kashgar and Hotan in southern Xinjiang were clustered into the one, and *P. euphratica* populations represented by Changji, Tacheng and Altay in Northern Xinjiang were clustered into the other. In the cluster analysis based the genetic distance of 17 natural *P. pruinosa* populations (Fig. S2), when the genetic distance among populations was approximately equal to 0.105, *P. pruinosa* populations were mainly divided into two clades, of which *P. pruinosa* populations represented by Bazhou, Kashgar and Hotan in southern Xinjiang were clustered into the one, and *P. pruinosa* populations represented by Ili in Northern Xinjiang were clustered into the other. In general, there were different in the genetic distance among different populations, indicating that natural *P. euphratica* and *P. pruinosa* populations ha certain genetic differentiation respectively.

Population Structure Analysis

The bayesian cluster analysis was carried out on the 693 samples of 47 natural *P. euphratica* populations by using structure software (Fig. 1). The results showed that, when k value was equal to four, ΔK had the maximum peak, (Fig. S4) which meant the natural *P. euphratica* population was optimally divided into four clades. The analysis of principal coordinate axis also supported four clades of different *P. euphratica* populations (Fig. S6), the *P. euphratica* populations were mainly clustered into one group represented Kashgar Prefecture, Bayingol Mongolian Autonomous Prefecture populations in southern Xinjiang, and the other group represented the Tarbagatay Prefecture, Changji Hui Autonomous Prefecture populations in northern Xinjiang respectively. There was certain gene flow among different *P. euphratica* populations. The bayesian cluster analysis was carried out on the 235 samples of 17 natural *P. pruinosa* populations by using structure software (Fig. S3). The results showed that, when k value was equal to four, ΔK has the

maximum peak (Fig. S5) which meant the natural *P. pruinosa* population was optimally divided into four clades. The analysis of principal coordinate axis also supported four clades of different *P. pruinosa* population (Fig. S7). the *P. pruinosa* populations were mainly clustered into one group represented the Yuli County, Qiemo county, Karakax County in southern Xinjiang, and the other group represented the Nilka County in northern Xinjiang respectively. There existed a certain gene flow among different *P. pruinosa* populations.

Fingerprinting Construction

In this study, 693 samples from 47 *P. euphratica* populations and 235 samples from 17 *P. pruinosa* populations were amplified by using 10 EST-SSR primers associated with drought resistance, resulting in a total of 23 bright and clear polymorphic bands. The band numbers of each primer pair ranged from 2 to 3. According to the computer language, EST-SSR fingerprinting of each samples was converted into digital matrix (1, presence; 0, absence; 9 no amplification). According to the formula $P = (1/2)^n$, the 10 EST-SSR primer pairs could detect 23 polymorphic alleles, and the probability of the same fingerprints appearing in the tested population was as follow: $P = (1/2)^{23} = 1.192 \times 10^{-7}$. Therefore, the 10 pairs of EST-SSR primers could fully distinguish the 47 natural *P. euphratica* populations and 17 *P. pruinosa* populations (Tab. S6, S7).

Discussion

In this study, 10 pairs of codominant EST-SSR primers related to drought stress of *P. euphratica* and *P. pruinosa* were used to analyze the genetic diversity of 693 samples of 47 natural *P. euphratica* populations and 235 samples of 17 natural *P. pruinosa* populations distributed in Xinjiang, Gansu and Inner Mongolia. The average level of the observed allele number (N_a) was an important index to evaluate the polymorphism of SSR locis and the degree of population variation. The average level of the observed allele number (N_a) (Tab. S3) was 1.743 ($N_a = 1.743$) and 1.712 ($N_a = 1.712$) in the *P. euphratica* and *P. pruinosa* populations respectively, which was lower than that the 20 natural *Pinus yunnanensis* populations of using SSR markers ($N_a = 3.7$) (Xu et al. 2015), and the 27 *Rosa odorata* var. *gigantea* (Rosaceae) populations used the 7 pairs of SSR markers ($N_a = 3.9$) (Meng et al. 2016). The average level of the effective allele number (N_e), Shannon's information index (I), observed heterozygosity (H_o) and expected heterozygosity (H_e) were important parameters to evaluate genetic diversity of populations. Through the research results of natural *P. euphratica* ($N_e = 1.407$, $I = 0.34$, $H_o = 0.238$, $H_e = 0.215$) (Tab. S3) and *P. pruinosa* ($N_e = 1.382$, $I = 0.322$, $H_o = 0.204$, $H_e = 0.205$) (Tab. S4) populations, its Shannon's information index was all lower than that of the wild apricot populations ($I = 0.46$) (He et al. 2007) and *Malus sieversii* populations ($I = 0.41$) (Zhang et al. 2007), its expected heterozygosity (H_e) was lower than that of north American *Pinus strobus* ($H_e = 0.531$) (Mandák et al. 2013), *Populus alba* ($H_e = 0.368$), *Populus tremula* ($H_e = 0.490$) (Lexer et al. 2005), *Populus euphratica* ($H_e = 0.787$) (Wang et al. 2011a). This might be due to the fact that the sequences were derived from the drought stress-related genome sequences of *P. euphratica* and *P. pruinosa*, so the polymorphism bands acquired the codominant EST-

SSR markers was lower than that the genomic SSR (Powell et al. 1996). Previous studies had also shown that the polymorphism of genomic SSR was higher than that of EST-SSR marker. For example, Eujayl et al. (Eujayl et al. 2001) studied the wheat genetic diversity by using genomic SSR and EST-SSR molecular markers, and the results showed that the polymorphism of genomic SSR was higher. Li et al. (Li et al. 2007) also found the fact that the genomic SSR polymorphism was higher, but EST-SSR markers could reflect accurately the genetic differences and genetic relationships among different wheat genotypes.

The molecular variance analysis (AMOVA) of *P. euphratica* populations showed that the variation among and within populations of total variation was 23% and 77% respectively (Table 1). In natural *P. pruinosa* populations (Tab. S5), the variation among populations accounted for the 24% of the total variation, while the variation within populations was 76%. the variation of *P. euphratica* and *P. pruinosa* populations showed that the variation within population was all the main source of the total population variation. the variation within population (77%, 76%) was less than that of previous researches respectively. For example, Wei (Wei 2010) analyzed 16 *Populus simonii* populations with 20 pairs of SSR primers, and the results showed that the variation within population was 85.97% of the total variation. Dong and bai (Dong and Bai 1998) analyzed 8 natural *P. euphratica* populations with four isozymes, and the results showed that 93.7% of the variation derived from intrapopulation. Wang et al. (Wang et al. 2011a) analyzed the *P. euphratica* populations using 8 pairs of SSR markers in Northwest China, and its results showed that 94.79% of the variation existed the intrapopulations. However, the variation within population *P. euphratica* and *P. pruinosa* was more than the previous research that Sheng et al (Sheng et al. 2008) found the variation existed the intrapopulation is 60.51% using 10 pairs of RAPD primers in the five natural *P. euphratica* populations. These results showed that, using different molecular markers to analyze the genetic variation of the populations, there were differences in the variation percentage of within population. but the overall conclusion was the same that the variation within population was bigger than that among populations at the molecular level.

Through UPGMA cluster analysis of 47 natural *P. euphratica* populations and 17 natural *P. pruinosa* (Fig. S1, S2), the results showed that, when the genetic distance among populations was about 0.1, *P. euphratica* and *P. pruinosa* populations were all divided into subgroups represented by the populations in southern and northern Xinjiang respectively, which was consistent with the previous cluster results of *P. euphratica* populations (Chen et al. 2021). It showed that the differentiation degree of the *P. euphratica* or *P. pruinosa* populations distributed in northern and southern Xinjiang were relatively largest. Zeng et al. (Zeng et al. 2018) studied the population structure and phylogeography of *P. euphratica* populations in Northwest China by using chloroplast DNA (cpDNA) sequences and nuclear microsatellites respectively. The results showed that the uplift of Tianshan Mountains directly hindered the gene flow between *P. euphratica* from northern and southern Xinjiang, and the distribution contraction caused by climate oscillation further accelerated the divergence of in these areas. Therefore, there was an obvious differentiation among *P. euphratica* populations from southern and northern Xinjiang. At the same time, Jia et al. (Jia et al. 2020) analyzed the genetic diversity of 252 individuals of 27 natural populations with the whole genome resequencing. The results showed that there was a high degree of genetic differentiation among populations in southern and northern Xinjiang.

Bayesian cluster analysis was carried out on the 692 samples from 47 *P. euphratica* populations and 237 samples from 17 *P. pruinosa* populations with STRUCTURE software. The results showed that, when the K value was about 4, ΔK had the maximum peak value (Evanno et al. 2005) (Fig. S4, S5). At this time, the same variant of natural *P. euphratica* and *P. pruinosa* populations clustered into one group respectively (Fig. 1, S3). At the same time, 47 *P. euphratica* populations and 17 *P. pruinosa* populations were analyzed by the principal coordinate axis (Fig. S6, S7). The results showed that *P. euphratica* and *P. pruinosa* populations were also clustered into four clades, of which the populations with high genetic similarity were clustered into one clade respectively. Specifically, the difference among populations from northern and southern Xinjiang was relatively obvious, while the difference within populations in northern and southern Xinjiang was relatively low, and the genetic similarity of its populations was high. The possible reason was that the similarity among populations was often related to the ecological environment, species reproductive ability and dispersal media of geographical distribution areas. For example, species with biological characteristics such as wide distribution, long seeds longevity in vitro and wind pollination were relatively easy to overcome the geographical isolation among populations, which was conducive to the hybridization of individuals to maintain high similarity of different populations (Hamrick et al. 1979), *P. euphratica* and *P. pruinosa* had biological characteristics such as long life longevity, root tiller propagation, wind pollination and seed dispersal, which might be the main reason for the high genetic similarity among populations in the same region (Wang et al. 2011b). At the same time, the seed maturity of *P. euphratica* and *P. pruinosa* populations in southern Xinjiang often occurred in flood period (Eusemann et al. 2013), these unique environmental conditions were conducive to the population propagation of *P. euphratica* and *P. pruinosa* and build a large-scale effective population, making populations in Southern Xinjiang become the most densely distributed area in the world (Wang 1996).

The construction fingerprinting of species in different regions helped to identify different varieties or strains quickly and accurately, which also provided great convenience for the protection and breeding management of varieties in different regions in the later period (Feng et al. 2022). In this study, 10 pairs of EST-SSR primers with good polymorphism were selected from 133 pairs of primers to construct the fingerprintings of different samples of *P. euphratica* and *P. pruinosa* populations (Tab. S6, S7). Of course, the *P. euphratica* and *P. pruinosa* populations distributed along the riparian forests had a relatively narrow genetic backgrounds and relatively close genetic relationships, which caused to a certain difficulties to identify the different populations samples. To a certain extent, The EST-SSR markers developed based on the transcription level in the study had a high accuracy, were suitable for the construction fingerprintings of *P. euphratica* and *P. pruinosa* populations and the identification of different populations samples.

Conclusion

In this study, the 10 pairs of EST-SSR markers were used to analyze the genetic diversity characteristics and successfully construct the fingerprintings of different *P. euphratica* and *P. pruinosa* populations. The results showed that the genetic diversity levels of *P. euphratica* and *P. pruinosa* populations in southern Xinjiang were higher than other distribution areas respectively. There was a high degree of genetic similarity between different populations and the fingerprintings of each of them has been successfully

constructed. These primers laid the foundation for the later protection of populations with a high level of genetic diversity and the preservation of excellent and rare germplasm resources.

Declarations

Acknowledgements We thanked Shanhe Zhang, Jianhao Sun, Xiaoli Han, Xiu Li, Jianing Wang, Xiaoshan Dong, and Zijian Wang who participated in this sample collection.

Funding This work was supported by the earmarked fund for the Corps Key Laboratory of Protection and Utilization of Biological Resources in Tarim Basin (grant number BRZD2003), the Bintuan Science and Technology Program (grant number 2021BB010) and the Construction of *Populus euphratica* Germplasm Resource Garden (grant number TDZKKY202202).

Conflict of interest Authors declare no conflict of interest.

References

1. Bérubé Y, Zhuang J, Rungis D, Ralph S, Bohlmann J, Ritland K (2007) Characterization of EST SSRs in loblolly pine and spruce. *Tree Genet Genomes* 3(3):251-259. <https://doi.org/10.1007/s11295-006-0061-1>
2. Cao D, Li J, Huang Z, Baskin CC, Baskin JM, Hao P, Zhou W, Li J, Reinhart KO (2012) Reproductive Characteristics of a *Populus euphratica* Population and Prospects for Its Restoration in China. *Plos One* 7(7):e39121. <https://doi.org/10.1371/journal.pone.0039121>
3. Caramante M, Corrado G, Monti LM, Rao R (2011) Simple sequence repeats are able to trace tomato cultivars in tomato food chains. *Food Control* 22(3):549-554. <https://doi.org/10.1016/j.foodcont.2010.10.002>
4. Chen XX, Gai ZS, Zhai JT, Xu JD, Jiao PP, Wu ZH, Li ZJ (2021) Genetic diversity and construction of core conservation units of the natural populations of *Populus euphratica* in Northwest China. *Biodiversity Science* 29 (12):1638–1649. <https://doi.org/10.17520/biods.2021249>
5. Chen YN, Dai XG, Hou J, Guan HW, Wang YX, Li Y, Yin TM (2016) DNA fingerprinting of oil camellia cultivars with SSR markers. *Tree Genetics & Genomes* 12(1):7. <https://doi.org/10.1007/s11295-015-0966-7>
6. Dillon NL, Innes DJ, Bally ISE, Wright CL, Devitt LC, Dietzgen RG (2014) Expressed Sequence Tag-Simple Sequence Repeat (EST-SSR) Marker Resources for Diversity Analysis of Mango (*Mangifera indica* L.). *Diversity* 6(1):72-87. <https://doi.org/10.3390/d6010072>
7. Dong YZ, Bai GB, (1998) Study on the genetic structure of natural population of *Populus euphratica* Oliv by isozyme technique (in Chinese). *Journal of Northeast Forestry University*, 26(5):16-20.
8. Earl DA, Vonholdt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4(2):359-361. <https://doi.org/10.1007/s12686-011-9548-7>

9. Eujayl I, Sorrells M, Baum M, Wolters P, Powell W (2001) Assessment of genotypic variation among cultivated durum wheat based on EST-SSRS and genomic SSRS. *Euphytica* 119(1), 39-43.
<https://doi.org/0.1023/A:1017537720475>
10. Eusemann P, Petzold A, Thevs N, Schnittler M (2013) Growth patterns and genetic structure of *Populus euphratica* Oliv. (Salicaceae) forests in NW China implications for conservation and management. *Forest Ecol Manag* 297:27-36. <https://doi.org/0.1016/j.foreco.2013.02.009>
11. Evanno GS, Regnaut SJ, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol Ecol* 14(8):2611-2620.
<https://doi.org/10.1111/j.1365-294X.2005.02553.x>
12. Feng PL, Han R, Wang YY, Shao DK, Li QH, Zhong QW (2022) Construction of DNA fingerprints of pepper (*Capsicum annuum* L.) varieties based on SSR markers (in Chinese). *Acta Agriculturae Boreali-occidentalis Sinica*, 31(3): 320-327.
13. Gai ZS, Zhai JT, Chen XX, Jiao PP, Zhang SH, Sun JH, Qin R, Liu H, Wu ZH, Li ZJ (2021) Phylogeography reveals geographic and environmental factors driving genetic differentiation of *Populus sect. Turanga* in northwest China. *Front Plant Sci* 12:705083.
<https://doi.org/10.3389/fpls.2021.705083>
14. Gao L, Tang J, Li H, Jia J (2003) Analysis of microsatellites in major crops assessed by computational and experimental approaches. *Mol Breeding* 12(3):245-261.
<https://doi.org/10.1023/A:1026346121217>
15. Hamrick JL, Linhart YB, Mitton JB (1979) Relationships between life history characteristics and electrophoretically detectable genetic variation in plants. *Annu Rev Ecol Syst* 10(3):173-200.
<https://doi.org/10.1146/annurev.es.10.110179.001133>
16. Xia HB, Lu BR (2009) Software for transforming co-dominant genotype scores into binary data matrix: its application to genetic diversity research (In Chinese). *Journal of Plant Genetic Resources* 10(1):97-102.
17. He TM, Chen XS, Xu Z, Gao JS, Lin PJ (2007) Using SSR markers to determine the population genetic structure of wild apricot (*Prunus armeniaca* L.) in the Ily Valley of West China. *Genet Resour Crop Ev* 54(3):563-572. <https://doi.org/10.1007/s10722-006-0013-5>
18. Jia HX, Liu GJ, li J, Zhang J, Sun P, Zhao ST, Zhou X, Lu MZ, Hu JJ (2020) Genome resequencing reveals demographic history and genetic architecture of seed salinity tolerance in *Populus euphratica*. *J Exp Bot* 71(14):4308–4320. <https://doi.org/10.1093/jxb/eraa172>
19. Jing W, Chen J, Wang L, Wang S (2017) Genome-Wide investigation of WRKY transcription factors involved in terminal drought stress response in common bean. *Front Plant Sci* 8:380.
<https://doi.org/10.3389/fpls.2017.00380>
20. Karihaloo JL (2015) DNA fingerprinting techniques for plant identification. In: *Plant biology and biotechnology* springer, Springer India, pp 205–221. https://doi.org/10.1007/978-81-322-2283-5_9
21. Kumar M, PonnuswamiV, Nagarajan P, eyakumar PJ, Senthil N (2013) Molecular characterization of ten mango cultivars using simple sequences repeat (SSR) markers. *African Journal of Biotechnology*

- 12(47):6568–6573. <https://doi.org/10.5897/AJB2013.12797>
22. Lexer C, Fay MF, Joseph JA, NICA MS, Heinze B (2005) Barrier to gene flow between two ecologically divergent *Populus* species, *P. alba* (white poplar) and *P. tremula* (European aspen): the role of ecology and life history in gene introgression. *Mol Ecol* 14(4):1045-1057. <https://doi.org/10.1111/j.1365-294X.2005.02469.x>
23. Li BS, Qin YR, Duan H, Yin WL, Xia XL (2011) Genome-wide characterization of new and drought stress responsive microRNAs in *Populus euphratica*. *J Exp Bot* 62(11):3765-3779. <https://doi.org/10.1093/jxb/err051>
24. Li JM, Li HQ, Li XQ, Li LH, Chai SC (2007) Studies on genetic diversity of genomic-SSR and EST-SSR molecular marker in common wheat (in Chinese). *Journal of Anhui Agricultural Sciences* 35(26):8173-8175.
25. Li ZJ (2020) *Populus euphratica* forest In Xinjiang (in Chinese). China Forestry Publishing House, Beijing, pp13-16.
26. Liu H, Zhang XQ, Ma X, Zhang RZ, He GW, Pan L, Jin MY (2017) Construction of EST-SSR fingerprinting based on fluorescence detection technology for Italian ryegrass. *Scientia Agricultura Sinica* (in Chinese) 50(3):437-450.
27. Luo Z, Brock J, Dyer JM, Kutchan T, Daniel S, Megan A, Ge YF, Noah F, Abdel-Haleem H (2019) Genetic diversity and population structure of a *Camelina sativa* Spring Panel. *Front Plant Sci* 10:184. <https://doi.org/10.3389/fpls.2019.00184>
28. Ma HS, Dan L, Peng S, Xia XL, Yin WL (2010) The salt- and drought-inducible poplar GRAS protein SCL7 confers salt and drought tolerance in *Arabidopsis thaliana*. *J Exp Bot* 61(14):4011-4019. <https://doi.org/10.1093/jxb/erq217>
29. Mandák B, Hadincová V, Mahelka V, Wildová R (2013) 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* 8(7):e68514. <https://doi.org/10.1371/journal.pone.0068514>
30. Meng J, He SL, Li DZ, Yi TS (2016) Nuclear genetic variation of *Rosa odorata* var. *gigantea* (Rosaceae): population structure and conservation implications. *Tree Genet Genomes* 12(4):65. <https://doi.org/10.1007/s11295-016-1024-9>
31. Moisan-Thiery M, Marhadour S, Kerlan MC, Dessenne N, Perramant M, Gokelaere T, Hingrat YL (2005) Potato cultivar identification using simple sequence repeat markers (SSR). *Potato Research* 48(3), 191-200. <https://doi.org/10.1007/BF02742376>
32. Peakall R, Smouse PE (2012) GenAEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28(19):2537-2539. <https://doi.org/10.1093/bioinformatics/bts460>
33. Poncet V, Rondeau M, Tranchant C, Cayrel A, Hamon S, Kochko Ad, Hamon P (2006) SSR mining in coffee tree EST databases: potential use of EST–SSRs as markers for the *Coffea* genus. *Mol Genet Genomics* 276(5):436-449. <https://doi.org/10.1093/bioinformatics/bts460>

34. Powell W, Machray GC, Provan J (1996) Polymorphism revealed by simple sequence repeats. *Trends Plant Sci* 1(7):215-222. [https://doi.org/10.1016/1360-1385\(96\)86898-1](https://doi.org/10.1016/1360-1385(96)86898-1)
35. Rohlf FJ (1998) NTSYSpc Version 2.0: User Guide. Applied Biostatistics Inc, New York, pp 6-43.
36. Saghai-Maroo MA, Soliman KM, Jorgensen RA, Affiliations RWAI (1984) Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. *PNAS* 81(24):8014–8018. <https://doi.org/10.1073/pnas.81.24.8014>
37. Saha M, Mian MAR, Eujayl I, Zwonitzer JC, Wang L, May GD (2004) Tall fescue EST-SSR markers with transferability across several grass species. *Theor Appl Genet* 109(4):783-791. <https://doi.org/10.1007/s00122-004-1681-1>
38. Sha T, Liang H, Yan D, Zhao Y, Xiao H (2013) *Populus euphratica*: the transcriptomic response to drought stress. *Plant Mol Biol* 83(6):539-557. <https://doi.org/10.1007/s11103-013-0107-3>
39. Sheng G, Abdureyim A, MA XR, Rahman E (2008) Study on geographical isolation and genetic diversity of *Populus euphratica* in three basins of Xinjiang by RAPD Analysis. *Biochemical and Biophysical Research Communications*. *Biotechnology* 18(5):18-21.
40. Wang J, Li ZJ, Guo QH, Ren GP, Wu YX (2011a) Genetic variation within and between populations of a desert poplar (*Populus euphratica*) revealed by SSR markers. *Ann Forest Sci* 68(6):1143-1149. <https://doi.org/10.1007/s13595-011-0119-6>
41. Wang J, Wu YX, Ren GP, Guo QH, Liu JN, Lascoux M (2011b) Genetic differentiation and delimitation between ecologically diverged *Populus euphratica* and *P. pruinosa*. *Plos One* 6(10):e26530. <https://doi.org/10.1371/journal.pone.0026530>
42. Wang SJ (1996) The Status, Conservation and recovery of global resources of *Populus euphratica*. world forestry research (in Chinese). *World Forestry Research* (6):38-45.
43. Wei QJ (1990) *Populus euphratica*. China Forestry Publishing House, Beijing.
44. Wei ZZ (2010) Evaluation of genetic resources and identification of SSRs associated with important traits in *Populus simonii*. Degree, Beijing Forestry University.
45. Xu YL, Cai NH, Xu Y, He CZ, Wang DW, Chen S, Duan AA (2015) Genetic variation and conservation units analysis in *pinus yunnanensis* natural populations (in Chinese). *Forest Research* 28(6):883-891.
46. Yan DH, Fenning T, Tang S, Xia XL, Yin WL (2012). Genome-wide transcriptional response of *Populus euphratica* to long-term drought stress. *Plant Sci* 195:24-35. <https://doi.org/10.1016/j.plantsci.2012.06.005>
47. Zeng YF, Zhang JG, Abuduhamiti B, Wang WT, Jia ZQ (2018) Phylogeographic patterns of the desert poplar in Northwest China shaped by both geology and climatic oscillations. *BMC Ecology and Evolution* 18(1):75. <https://doi.org/10.1186/s12862-018-1194-1>
48. Zhang CY, Chen XS, He TM, Liu XL, Feng T, Yuan, ZH (2007) Genetic Structure of *Malus sieversii* Population from Xinjiang, China, Revealed by SSR Markers. *J Genet Genomics* 34(10), 947-955. [https://doi.org/10.1016/S1673-8527\(07\)60106-4](https://doi.org/10.1016/S1673-8527(07)60106-4)

49. Zhang L, Jiao PP, Li ZJ (2012) Genetic diversity of *Populus pruinosa* populations in Xinjiang of China based on SSR analysis (in Chinese). Chinese Journal of Ecology 31(11):2755-2761.

50. Zhang YD, Peng C, Li ZF, Yang YL, Hu XY (2011) Genetic diversity of Genomic-SSR and EST-SSR markers in interspecies of poplar (in Chinese). Journal of Northeast Forestry University 39(12):8-11.

Figures

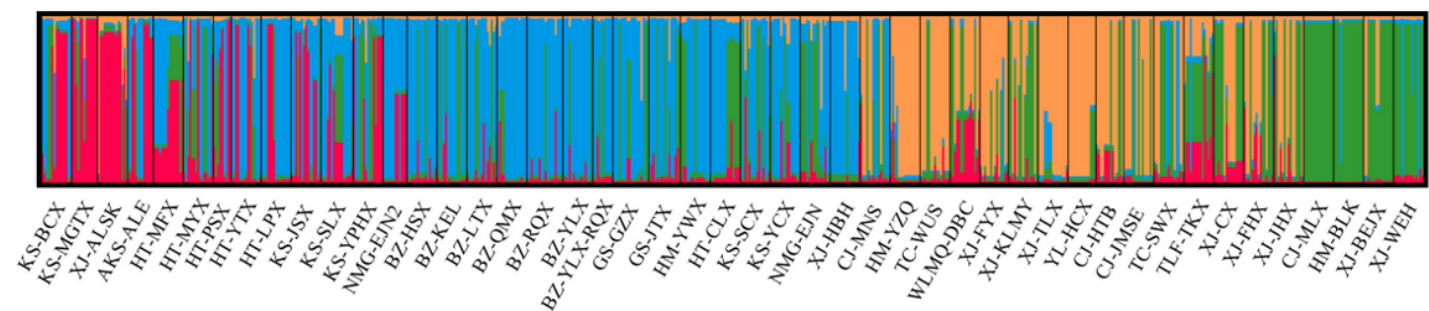


Figure 1

Analysis of population structure of 47 natural *Populus euphratica* populations

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplementarymaterial.docx](#)