

Population structure and genetic diversity of *Tamarix chinensis* as revealed with microsatellite markers in two estuarine flats

Zhao-Yu Jiang^{*}, Ao-Ao Yang^{*}, Hai-Guang Zhang, Wen-Bo Wang and Ru-Hua Zhang

Linyi University, College of Life Science, Linyi City, Shandong Province, China

^{*}These authors contributed equally to this work.

ABSTRACT

Background. *Tamarix chinensis* Lour. is a 3–6-meter-tall small tree with high salt- and alkali- tolerance and aggressive invasiveness, mainly distributed in the eastern part of China in warm-temperate and subtropical climate zones, yet there is little information available regarding genetic diversity and population structure.

Methods. A total of 204 individuals of nine *T. chinensis* populations were investigated for genetic diversity and population structure using a set of 12 highly polymorphic microsatellite markers.

Results. The total number of alleles detected was 162, the average number of effective allele was 4.607, the average polymorphism information content (PIC) value of the 12 loci was 0.685, and the mean observed heterozygosity (H_o) and the mean expected heterozygosity (H_e) was 0.653 and 0.711, respectively. Analysis of molecular variance (AMOVA) showed a 5.32% genetic variation among *T. chinensis* populations. Despite a low population differentiation, Bayesian clustering analysis, discriminant analysis of principal components (DAPC) and the unweighted pair group method with arithmetic mean (UPGMA) clearly identified three genetic clusters correlated to the populations' geographic origin: the northern populations including those from Yellow River Delta, the Fangshan (FS) population from Beijing, the Changyi (CY) population from Bohai Bay, the Huanjiabu (HHJ) population from Hangzhou Bay, and the remaining two populations from Hangzhou Bay. There was a significant relationship between the genetic distance and geographical distance of the paired populations. Gene flow (N_m) was 4.254 estimated from F_{ST} .

Conclusion. *T. chinensis* possessed high genetic diversity comparable to tree species, and although the population differentiation is shallow, our results classified the sampled populations according to sampling localities, suggesting the different origins of the study populations.

Subjects Biodiversity, Genetics, Plant Science, Population Biology

Keywords *Tamarix chinensis* Lour., Genetic diversity, Population structure, Microsatellite markers, Estuarine flats

Submitted 10 January 2023
Accepted 20 July 2023
Published 11 September 2023

Corresponding author
Ru-Hua Zhang,
ruhuazh2000@126.com

Academic editor
Sushil Kumar

Additional Information and
Declarations can be found on
page 14

DOI 10.7717/peerj.15882

© Copyright
2023 Jiang et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

INTRODUCTION

The old world temperate genus *Tamarix* L. (Tamaricaceae) comprises 68 species native to Eurasia and Africa, typically inhabiting deserts and sub-deserts in temperate and subtropical regions (Baum, 1978; Zohary, 1987). Originating at least prior to Eocene in Mediterranean basin or Irano-Turanian regions where comparatively high number of species are found, the genus spread eastward and westward from the diversity centers (Villar, Juan & Alonso, 2014; Qaiser, 1981). *Tamarix* reached the western China in Oligocene (Wang & Yang, 1980; Song, 1958) and Tarim Basin is the secondary distribution center with 16 species (five being endemic) (Zhang, Pan & Yin, 2003; Wang, 1992). *Tamarix chinensis* and its two congeners (*T. austromongolica* and *T. ramosissima*) further colonized eastward, possibly along the Yellow River corridor according to molecular evidence, and it established populations along East coast of China (Liang et al., 2018b).

T. chinensis Lour. is a salt- and alkali- tolerant deciduous shrub with 3–6-meter-height with significant ecological and economic importance (Jiang, Chen & Bao, 2012; Wang et al., 2009). Through secreting excessive salts with glands in leaves, it can tolerate high salt concentration in soil and has been widely used for establishing coastal protection forests in the East Sea coast of China (Zhao et al., 2008). Its bark and twig has high tamarixone and tamarixolare content which can cure cold (Jiang et al., 2011). It is an important firewood source where coal is unaffordable, especially in some regions of western China. Its wood is hard and flexible and wear-resistant, which is often for weaving use. Its wood fiber can be used as raw material for paper and fiberboard. It blooms three times a year from May to October and is often used as a garden plant to beautify the environment. However, this species has rapidly dwindled in the last 20 years due to overexploitation for heating, weaving and medical use. Moreover, the taxonomic position of *T. chinensis* and *T. austromongolica* remains unsolved. Although *T. austromongolica* is distributed in the upper reaches of the Yellow River and *T. chinensis* in the lower reaches (Liang et al., 2018b), their phenotypic traits are quite close and some botanists placed them as the same species.

The genetic diversity of a species plays a crucial role in providing evolution potential and affecting its adaptation to various environmental selective pressures including temperature extremes, drought and flood, and high salt and alkali concentration (Hughes et al., 2008). Therefore, evaluating the genetic diversity within and among the sampled populations of a species is elementary for further clarifying the distribution of genetic variation across its range, estimating the gene flow levels between populations, and determining the evolutionary significant units (ESU) (Jiang et al., 2019; Islam et al., 2012). Up to data, several studies have been reported on the population diversity and structure of *T. chinensis* and its congeners. With six polymorphic microsatellite markers, Liang et al. (2018a) performed the genetic analysis of the populations collected along the Yellow River and revealed moderate levels of genetic diversity (a range of the mean expected heterozygosity (H_e) from 0.366 to 0.740 for the studied population) and low population genetic differentiation ($F_{ST} = 0.053$). Random primers analysis of populations distributed in the Yellow River Delta with Random Amplified Polymorphic DNA (RAPD) (Zhao et al., 2008) and Inter-simple Sequence Repeat (ISSR) (Jiang, Chen & Bao, 2012) also showed

low population genetic differentiation levels ($G_{st} = 0.0507$ for RAPD analysis and $G_{st} = 0.0707$ for ISSR) but comparatively high genetic diversity levels (Nei index $= 0.4061$ for RAPD analysis and Nei index $= 0.276$ for ISSR analysis). [Zhu et al. \(2016\)](#) conducted a correlation analysis between population genetic diversity and soil salinity, and the results showed that overall genetic diversity within populations decreased progressively along with the increasing soil salinity. Of note, [Lee et al. \(2018a\)](#) studied the population genetic structure and genetic diversity with SNPs derived from genotyping-by-sequencing (GBS) and identified genetic discontinuities among natural populations occurring in different river basins of the northwest of USA. Of studies on Chinese *T. chinensis* populations, no one included any of the extensively distributed populations in Hangzhou Bay, an estuarine flat in the subtropical zone.

Simple sequence repeat (SSR; microsatellite) markers are widely used in population diversity and structure of various trees and shrubs for their codominant characteristics, high polymorphism and stableness ([Fernández et al., 2018](#); [Nybom, 2004](#)). In this study, we used 12 SSR primers to detect and estimate the variation of *T. chinensis* mainly from two estuarine flats of different climate zones: Yellow River Delta of the warm-temperate zone and Hangzhou Bay of the subtropical zone. Our specific aims were to (1) evaluate the genetic diversity and genetic differentiation among nine *T. chinensis* populations, (2) determine whether population structure relative to sampling locations exists, and (3) compare the population genetic indices of the two different climate regions. Our study can deepen understanding of the spatial distribution pattern of genetic variation of *T. chinensis* and lay foundation for further exploring its adaptation to heterogeneous environments.

MATERIALS AND METHODS

Sample collection and DNA extraction

Samples were taken from nine natural populations of *T. chinensis* distributed from Beijing, Shandong Province and Zhejiang Province. Sampling mainly focused on two estuarine flats of different climate zones: Yellow River Delta of the warm-temperate zone, and Hangzhou Bay of the subtropical zone ([Fig. 1](#), [Table 1](#)). These two estuarine flats have many *T. chinensis* populations of different sizes and ages and populations of Yellow River Delta usually have larger size than those of Hangzhou Bay. Four populations (YHK, YDG, YXX and YHD) from Yellow River Delta and three populations (HHJ, HLS and HCX) from Hangzhou Bay were sampled. One population (CY) of the Bohai Sea coast from Changyi City and one inland population (FS) from Beijing were also included. A total of 19–29 individuals per population were collected according to the population size and a total of 204 individuals from nine populations was sampled. We collected young leaves from individual trees spaced at least 50 m apart, in order to avoid sampling related individuals. Fresh leaves were desiccated in silica gel. Geographical coordinates and altitude were acquired with GPS ([Table 1](#)). A modified cetyltrimethyl ammonium bromide (CTAB) method ([Su et al., 1998](#)) was used to extract total genomic DNA. Quality and quantity of the extracted DNA were measured with 1.0% (w/v) agarose gel electrophoresis and a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), respectively. DNA was diluted to 20 ng/ μ l and stored at -20°C for later use.

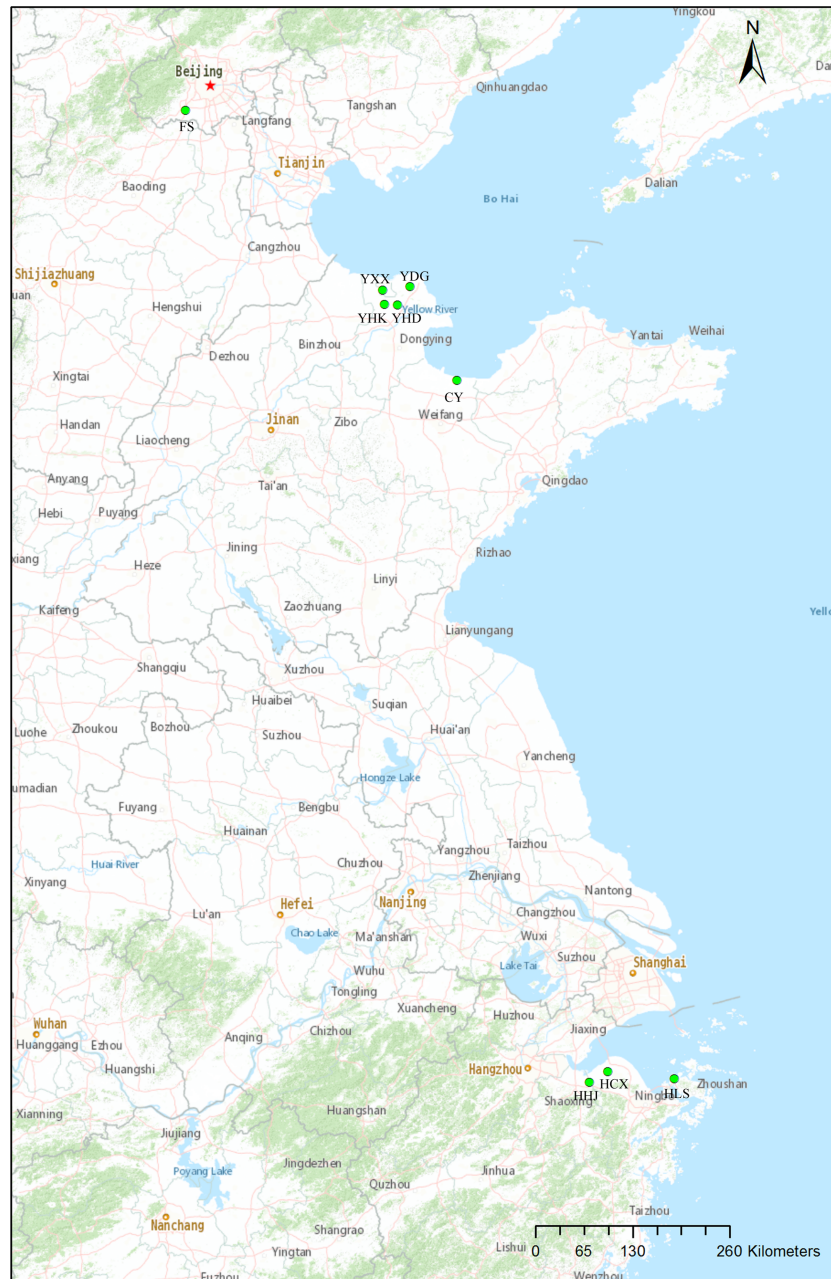


Figure 1 Locations of the nine natural *Tamarix chinensis* populations. The map was produced with the software ArcGIS 10.5.

Full-size  DOI: [10.7717/peerj.15882/fig-1](https://doi.org/10.7717/peerj.15882/fig-1)

We used a set of 12 highly polymorphic SSR markers (Table 2) for population genetic analysis, six of which were EST-SSRs (Zhang *et al.*, 2011) and the remaining six of which were genomic SSRs (Zhang, Wen & Xu, 2019). The polymerase chain reaction (PCR) was carried out according to (Zhang, Wen & Xu, 2019). Forward primers were labeled with four color fluorescent dyes (6-FAM, HEX, TAMRA and ROX) (Invitrogen, Shanghai, China). The PCR conditions of both kinds of SSR were as follows: 4 min denaturation at

Table 1 Geographical information of nine natural *T. chinensis* populations in China.

Populations	ID	Sample size	X (°N)	Y (°E)	Elevation
Helou Shandong	YHK	23	37.853	118.491	2.1–2.6
Changyi Shandong	CY	24	37.127	119.356	1.9–2.3
Fangshan Beijing	FS	20	39.674	116.095	48.3–47.7
Dongyinggang Shandong	YDG	24	38.02	118.795	1.8–2.9
Xinhu Shandong	YXX	22	37.988	118.467	2.7–3.3
Hekoudong Shandong	YHD	22	37.847	118.542	2.0–3.2
Huangjiabu Zhejiang	HHJ	22	30.222	120.952	1.1–2.2
Linshan Zhejiang	HLS	19	30.239	121.972	1.7–2.3
Cixi Zhejiang	HCX	29	30.445	121.172	1.4–2.2

Table 2 Genetic diversity among 12 simple sequence repeat (SSR) loci of nine *T. chinensis* populations.

Locus	Na	Ne	Ho	He	I	PIC	F _{ST}	Nm	Source
CF199044	10	5.948	0.8162	0.8341	1.9136	0.810	0.0437	5.471	Zhang et al. (2011)
EH051607	13	4.014	0.623	0.7529	1.6869	0.716	0.0953	2.373	Zhang et al. (2011)
EH053495	18	3.041	0.755	0.6728	1.4644	0.628	0.0421	5.688	Zhang et al. (2011)
CV790971	9	2.848	0.5373	0.6505	1.3133	0.592	0.0539	4.388	Zhang et al. (2011)
EG967460	23	7.428	0.7407	0.8677	2.3415	0.854	0.0417	5.745	Zhang et al. (2011)
AB298390	33	10.637	0.8469	0.9083	2.7643	0.899	0.0684	3.405	Zhang et al. (2011)
MG856344	15	4.618	0.7487	0.7855	1.8329	0.758	0.0491	4.842	Zhang, Wen & Xu (2019)
MG856346	13	5.545	0.6546	0.8218	1.9932	0.787	0.0582	4.046	Zhang, Wen & Xu (2019)
MG856349	15	4.089	0.7473	0.7731	1.4882	0.745	0.0728	3.814	Zhang, Wen & Xu (2019)
Gssr3	4	2.478	0.501	0.5361	0.7687	0.477	0.110	2.023	unpublished
Gssr4	3	2.312	0.3939	0.4073	0.5401	0.406	0.075	3.083	unpublished
Gssr5	6	2.322	0.4752	0.5230	0.7210	0.553	0.036	6.694	unpublished
mean	13.50	4.607	0.6533	0.7111	1.569	0.685	0.0622	4.245	

Notes.

Na, number of alleles; Ne, effect number of alleles; Ho, observed heterozygosity; He, expected heterozygosity; I, Shannon's information index; PIC, polymorphism information content; F_{st}, genetic differentiation coefficient; Nm, gene flow.

94 °C; followed by 30 cycles of 94 °C denaturation for 30s, varied annealing temperatures ranging from 49.6–58.3 °C for 30s (Table S1), and an elongation at 72 °C for 30s; and a final extension step at 72 °C for 5 min. PCR products were run on an ABI3730 XL Analyzer (Applied Biosystems, Foster City, CA, USA) and allele sizes were determined with GeneMapper 4.0 software (Applied Biosystems).

Data analysis

The occurrence of null alleles has a biased effect on the estimation of population genetic parameters, so we used MICROCHECKER v2.2.3 (Van Oosterhout et al., 2004) to check for the presence of null alleles for all loci. Hardy-Weinberg equilibrium exact tests and multilocus linkage disequilibrium tests were performed with GENEPOP (version 1.2) (Raymond & Rousset, 1995), and sequential Bonferroni correction tests were applied to adjust the significance levels of genotypic linkage disequilibrium. GenALEX 6.502

(Peakall & Smouse, 2005) was used to calculate genetic diversity indices including observed heterozygosity (H_o), expected heterozygosity (H_e), number of allele (N_a), number of effective alleles (N_e), inbreeding coefficient (F_{IS}), and Shannon diversity index (I). The genetic differentiation coefficient (F_{ST}) across all populations at each locus and over all loci, gene flow ($N_m = (1 - F_{ST})/4 F_{ST}$), and allelic richness (Ar) were calculated using FSTAT version 2.9.3 software (Goudet, 2001). We used CERVUS version 3.0.3 (Kalinowski, Taper & Marshall, 2007) to calculate the polymorphism information content (PIC) value of each locus according to: $PIC = 1 - \sum_{i=1}^n f_i^2$ where f_i is the frequency of the i th allele and n is the allele number.

Analysis of molecular variance (AMOVA) was performed using ARLEQUIN 3.5 (Excoffier & Lischer, 2010) to partition the genetic variability into within- and among-population components. The significance was tested based on 1,000 permutations for the AMOVA results (Excoffier, Laval & Schneider, 2005). Bottleneck events of the studied populations were tested using BOTTLENECK 1.2.02 (Piry, Luikart & Cornuet, 1999) with stepwise mutation model (SMM) and two-phased mutation model (TPM). Recently bottlenecked populations have higher observed heterozygosity (unbiased gene diversity) than expected from the observed number of alleles across microsatellite loci. Variance for TPM was set to 30 and the proportion of SMM in TPM was set to 80%. The “mode-shift” of allele frequency distribution was also used to identify bottlenecked populations (Luikart et al., 1998).

Bayesian clustering analysis was performed using STRUCTURE v2.3.4 software (Pritchard, Stephens & Donnelly, 2000) to determine the number of genetically homogeneous groups of the collected individuals and to evaluate the amount of admixture between individuals with the admixture model and allele frequencies correlated. We run the program with 100,000 burn-in steps followed by 100,000 Markov chain Monte Carlo (MCMC) steps, and 10 independent runs. The optimal K was determined through the online program STRUCTURE HARVESTER (Earl & Von Holdt, 2012) for each K from 1 to 10. Graphics were displayed with the DISTRUCT program (Rosenberg, 2004). We also implemented the Discriminant Analysis of Principal Components (DAPC) (Jombart, Devillard & Balloux, 2010) using the “adegenet” package (Jombart, 2008) for software R version 3.4.2 (R Core Team, 2017) to evaluate population structure. DAPC is a multivariate approach that is free of assumptions of Hardy–Weinberg equilibrium (HWE) and maximizes the genetic differentiation between groups with unknown prior clusters, thus improving the discrimination of populations. A dendrogram was also constructed using the unweighted pair-group method with arithmetic means (UPGMA) algorithm based on Nei’s genetic distance among populations with the NTSYS version 2.1 software (Rohlf, 2000).

We used the mantel function in the “vegan” R package (Dixon, 2003) to test the correlation between genetic distance (pairwise F_{ST}) and Euclidean geographical or environmental distance. The climatic layers of 19 bioclimatic variables under current periods were downloaded from the WorldClim database (<http://www.worldclim.org/>) with a resolution of 2.5 arc-minutes to evaluate the effects of environment on genetic variation. The Vifstep function of the R package usdm (Naimi et al., 2014) was used to select the

geographical and environmental variables with variance inflation factor (VIF) below 10 for Euclidean distance calculation. VIF measures the collinearity of the variables and the larger the VIF, the stronger the linear relationship between two or more predictor variables. After forward selection, five bioclimate variables (Bio2, Bio5, Bio8, Bio15 and Bio18) were chosen to constitute the environmental variable matrix. Euclidean distances of the five chosen environmental variables from different populations constituted the environmental variable matrix for the Mantel test. The significance of the correlation tests was evaluated with 9,999 permutations. To determine whether geographical or environmental variables may have affected genetic distance, we also conducted a partial Mantel test using the `mantel.partial` function in the R package `vegan` with 9,999 permutations.

RESULTS

Genetic diversity

The micro-checker analysis revealed one locus in population FS and population HHJ, three loci in population HCX, four loci in population YHD, five loci in population YHK with null alleles, while no null alleles were detected in four populations (CY, YDG, YXX, and HLS) (Table S2). Specifically, loci 2, 4, 5, 9, 11 revealed null alleles in two populations; loci 3, 8 and 10 in one population, while loci 1, 6, 7, 12 showed no null alleles. Hardy-Weinberg exact tests showed two loci in one populations (YXX), three loci in four populations (CY, FS, YHD, and HCX), four loci in two populations (YDG and HHJ), five loci in one population (YHK), and seven loci in one population (HLS) that deviated from equilibrium at $p < 0.05$ or at $p < 0.01$ (Table S3A). At population level, only one population (CY) was in HWE, and other populations deviated from equilibrium at $p < 0.05$ or $p < 0.01$ (Table S3B). Except for populations YDG, YXX and YHD, genotypic linkage disequilibrium (LD) was detected with a range of one locus pair (for population HYD) to 16 locus pairs (for population HHJ) after Bonferroni correction (Table S4A). Across all populations, 12 out of 132 locus pairs showed linkage disequilibrium, with loci 2 and 4 having the lowest number of pairs (1 pair) in LD (Table S4B).

We used 12 pairs of SSR primers to generate 162 alleles from 204 individuals of nine *T. chinensis* populations (Table 2). The results showed an average value of 13.50 for all alleles ranging from three (Gssr4) to 33 (AB298390) per locus. The effective number of alleles (N_e) ranged from 2.312 (Gssr4) to 10.636 (AB298390), with an average of 4.607. The observed heterozygosity (H_o) for each locus varied from 0.3939 (Gssr4) to 0.8469, with an average of 0.6533 and the expected heterozygosity (H_e) varied from 0.4073 (Gssr4) to 0.9083 (AB298390), with an average of 0.7111. Shannon's information index (I) ranged from 0.5401 (Gssr4) to 2.7643 (AB298390), with an average of 1.569. The genetic differentiation coefficient (F_{st}) (0.036 (Gssr5)–0.110 (Gssr3)) and the polymorphic information content (PIC) values (0.406 (Gssr4) –0.899 (AB298390)) were detected at all loci. Gene flow (N_m) ranged from 2.023 (Gssr3) to 6.694 (Gssr5), with an average of 4.245.

The genetic diversity analysis of nine *T. chinensis* population of was shown in Table 3. Allele richness ranged from 8.56 (HHJ) to 10.38 (YXX) with an average of 9.37. The number of alleles (N_a) varied from 7.25 for the HHJ population to 9.5 for the YXX population

Table 3 Genetic diversity parameters at the population level based on 12 SSR markers.

Population	Na	Ne	Ar	Np	I	Ho	He	Fis
YHK	8.000	4.984	9.25	2	1.651	0.570	0.690	0.177
CY	7.375	4.516	9.53	2	1.585	0.668	0.678	0.022
FS	7.375	4.022	8.85	3	1.529	0.667	0.655	-0.021
YDG	8.375	5.325	9.84	3	1.727	0.746	0.712	-0.072
YXX	9.500	5.914	10.38	7	1.854	0.743	0.739	-0.011
YHD	8.375	4.953	9.85	7	1.693	0.687	0.700	0.018
HHJ	7.250	3.978	8.56	5	1.547	0.749	0.678	-0.099
HLS	7.750	4.560	8.76	6	1.648	0.586	0.702	0.148
HCX	8.375	5.047	9.27	16	1.698	0.662	0.706	0.061
mean	8.042	4.811	9.37	5.667	1.659	0.674	0.696	0.025

Notes.

Abbreviations: Na, the number of different alleles; Ne, the number of effective alleles; Ar, the allelic richness; Np, the number of private alleles; I, Shannon's information index; Ho, the observed heterozygosity; He, the expected heterozygosity; Fis, inbreeding coefficient.

with an average of 8.042 and the effective number of alleles (N_e) varied from 3.978 for the HHJ population to 5.915 for the YXX population with an average of 4.811. Shannon's information index (I) varied from 1.547 (HHJ) to 1.854 (YXX), with a mean of 1.659. Mean value of expected heterozygosity ($H_e = 0.696$) was slightly higher than that of observed heterozygosity ($H_o = 0.674$). The fixation index (F_{is}) ranged from -0.099 for the HHJ population to 0.177 for the HHK population, with an average of 0.025. Out of the nine studied populations, population YXX from Yellow River Delta showed the highest values of genetic diversity parameters except for the number of private allele ($N_p = 7$), which was only lower than that of population HCX ($N_p = 16$). Altogether, 54 private alleles were identified and populations YHK and CY had the lowest number ($N_p = 2$) while population HCX displayed most private alleles ($N_p = 16$).

Averagely, higher values of genetic diversity indices ($N_a = 8.5625$, $N_e = 5.294$, $I = 1.73125$, $H_o = 0.680$, $H_e = 0.710$) for populations from Yellow River Delta were obtained than those of populations from Hang Zhou Bay ($N_a = 7.7916$, $N_e = 4.528$, $I = 1.631$, $H_o = 0.666$, $H_e = 0.695$) except for the number of private allele ($N_p = 4.75$ for populations from Yellow River Delta and $N_p = 9$ for populations from Hangzhou Bay) (Table S8). Inbreeding coefficients of populations from Yellow River Delta (-0.072 – 0.017 , average = 0.028) are lower than those of populations from Hangzhou Bay (-0.099 – 0.148 , average = 0.0367).

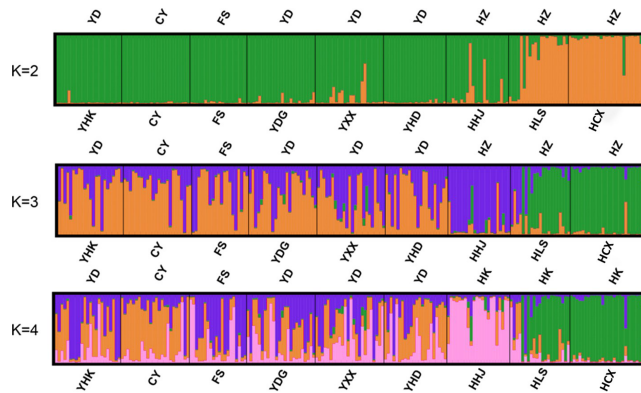
Under SMM and TPM, our Wilcoxon tests detected no significant heterozygote excess for the nine populations (Table S7). Likewise, mode-shift tests revealed an L-shaped distribution of alleles, suggesting the absence of a recent bottleneck.

Population differentiation and genetic structure

Low population genetic differentiation ($F_{ST} = 0.0532$, $p < 0.001$) (Table 4) was detected across *T. chinensis* populations. F_{ST} values of the 12 loci ranged from 0.036 (for Gssr5) to 0.110 (for Gssr3) (Table 2). AMOVA test revealed significant differentiation detected among regions ($F_{ct} = 0.0252$, $p < 0.05$), among populations within regions ($F_{sc} = 0.028$,

Table 4 The partition of SSR variation of *T. chinensis* by analysis of molecular variance (AMOVA).

Source of variation	d.f.	Sum of squares	Variation components	Percentage of variation
Among Regions	3	43.127	0.08068 Va	(Fsc) 2.8 [*]
Among Pops within regions	5	35.648	0.08787 Vb	(Fct) 2.52 ^{**}
Within populations	403	1,242.834	3.08396 Vc	94.68 ^{**}
Total	411	3.252	100%	

Notes.* $p < .05$.** $p < .01$.**Figure 2** Individual and population memberships to genetic clusters for $k = 2, 3$ and 4 using STRUCTURE.Full-size DOI: [10.7717/peerj.15882/fig-2](https://doi.org/10.7717/peerj.15882/fig-2)

$p < 0.001$), and 94.68% of variation was within-populations. The pairwise F_{ST} values between populations are shown in Table S5. The genetic differentiation between the CY and HCK populations was highest at 0.110.

Bayesian-based structure analysis supported the populations studied were classified into two clusters almost without admixture with a maximum delta value of 72.40 at $K = 2$ (Table S6). Populations from Yellow River Delta, CY population, FS population from Beijing, and one population from Hangzhou Bay (HHJ) formed the first cluster, and the two other populations, Hangzhou Bay HCK and HLS, formed the second cluster (Fig. 2A). At $K = 3$, HHJ population of the first cluster was assigned to the third cluster only with 2 admixed individuals (Fig. 2B), which is consistent with the DAPC results. At $K = 4$, HHJ population and two other Hangzhou Bay populations (HLS and HCK) formed two distinct clusters almost without admixture, while the individuals from the remaining populations were mostly admixed with two other clusters with approximate proportion (Fig. 2C).

The nine *T. chinensis* populations were grouped into two groups based on UPGMA (Fig. 3), two populations from Zhejiang Province, and the remaining populations. HHJ population was grouped into the second group with populations from Shandong Province and Beijing, but with the shortest branch length to populations of the first group. DAPC identified three distinct groups (Fig. 4), which is consistent with structure result at $K = 3$.

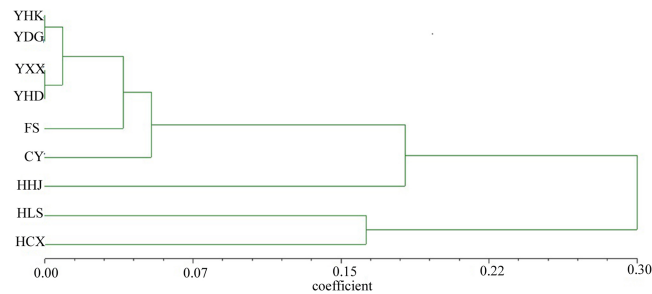


Figure 3 UPGMA tree of the nine *T. chinensis* populations using Nei's genetic distance.

Full-size [DOI: 10.7717/peerj.15882/fig-3](https://doi.org/10.7717/peerj.15882/fig-3)

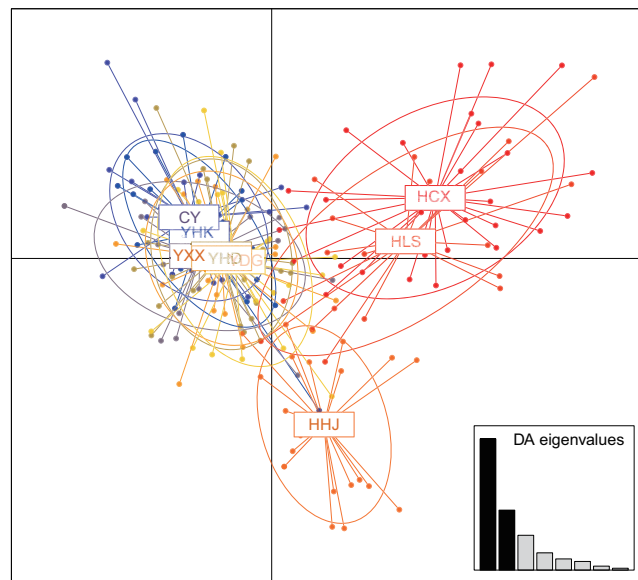


Figure 4 Clustering results of *T. chinensis* populations obtained by discriminant analysis of principal components (DAPC, PCs = 40).

Full-size [DOI: 10.7717/peerj.15882/fig-4](https://doi.org/10.7717/peerj.15882/fig-4)

Mantel test

Mantel tests showed the studied populations conformed to the patterns of isolation by distance (IBD) ($R^2 = 0.5784$, $p = 0.012$) (Fig. 5A) and isolation by environment (IBE) ($R^2 = 0.5923$, $p = 0.015$) (Fig. 5B) at significant levels. Given the autocorrelation of geographical distance and environmental distance ($R^2 = 0.9357$, $p = 0.002$) (Fig. 5C), partial mantel tests were performed, which detected highly significant IBD ($r = 0.7324$, $p = 0.002$), but no IBE patterns in the populations studied.

DISCUSSION

In this study, we evaluated the genetic diversity level and population genetic structure of *T. chinensis* mainly distributed in two coastal flats of different climate zones. Our study revealed a high genetic diversity and low differentiation level and strong population genetic

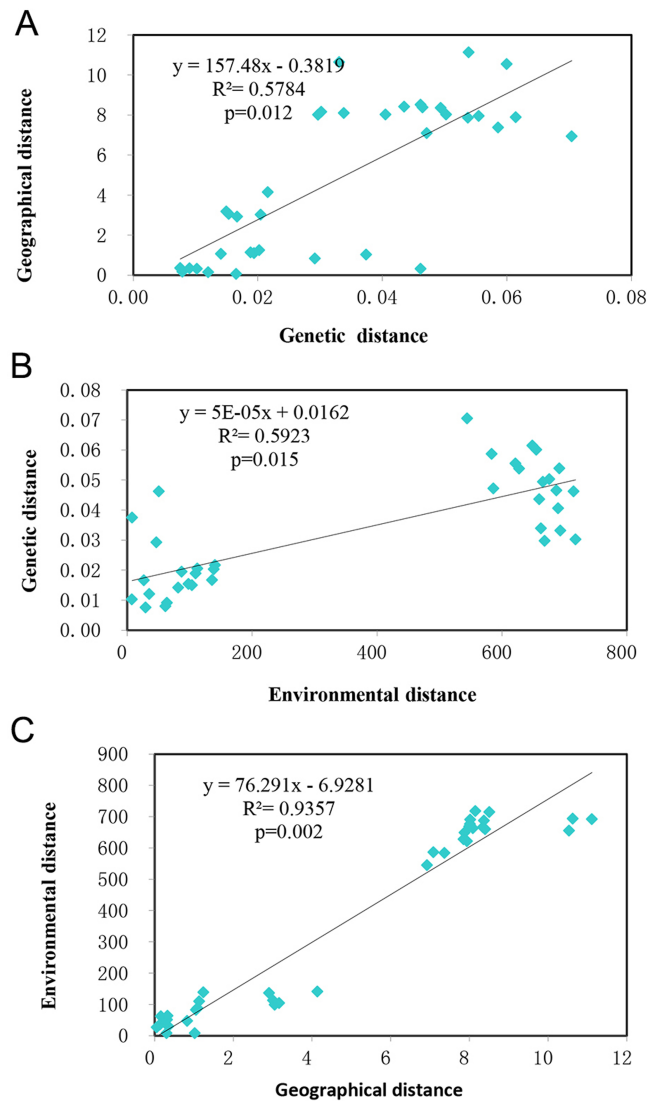


Figure 5 The relationship of genetic, geographical, and environmental distances. (A) The relationship between genetic distance and environmental distance of *T. chinensis*. (B) The relationship between genetic distance and environmental distance of *T. chinensis*. (C) The relationship between geographical distance and environmental distance of *T. chinensis*.

Full-size DOI: [10.7717/peerj.15882/fig-5](https://doi.org/10.7717/peerj.15882/fig-5)

structure correlated to populations' location that previous studies did not detect. The genetic diversity level of populations from Yellow River Delta is higher than that of those from Hangzhou Bay on average

High genetic diversity of *T. chinensis*

With highly polymorphic microsatellite markers, we detected high genetic diversity of *T. chinensis* (12 nSSRs, $H_e = 0.7111$) (Table 2) comparable to the widely distributed long-lived tree species including *Quercus acutissima* (12 nSSRs, $H_e = 0.760$) (Zhang *et al.*, 2013), *Q. liaotungensis* (19 nSSRs, $H_e = 0.801$) (Guo *et al.*, 2021), *Q. rubra* (15 SSRs, $H_e = 0.78$) (Lind

& Gailing, 2014), and *Q. ellipsoidalis* (15 SSRs, $H_e = 0.76$) (Lind & Gailing, 2014). Similarly, population genetic analysis through dominant markers including RAPD and ISSR also detected high genetic diversity in this species (Nei index = 0.4061 for RAPD analysis and Nei index = 0.276 for ISSR analysis) (Jiang, Chen & Bao, 2012; Zhao et al., 2008). The high level of genetic diversity of *T. chinensis* may be attributable to its high outcrossing rate, wide distribution, and long life span. Generally, the widely distributed long-lived plant species often maintain relatively high genetic diversity compared with narrow geographical ranges or endangered species (Wang, Bernhardsson & Ingvarsson, 2020; Hamrick, Godt & Shermann-Broyles, 1992). *T. chinensis* is an insect-pollinated plant with high outcrossing rate, although sometimes its pollen is dispersed *via* wind to the stigma of another plant. Our results showed there was no clonal reproduction in this species (data not shown) despite that it can be reproduced by cuttings and root suckers. It has a short generation time of 1-2y and its raceme flowers produce a large number tiny hairy seeds which can easily be transported by wind and/or water. Its seed can tolerate high salt concentration (0.15 g/l) (Zhang & Zhang, 2019) for a long time.

Our measurement of population genetic diversity showed significantly higher level than that of Liang et al. (2018a) where the genetic diversity ranged from 0.366 to 0.740 for populations along the Yellow River at six SSR markers and that of Zhu et al. (2016) where an average population heterozygosity of 0.493 was detected with 7 SSR markers in 26 populations from Yellow River Delta, and also that of its congener *T. taklamakanensis* (7SSRs, $H_e = 0.432$) (Su et al., 2017). The relatively low genetic diversity detected by those *T. chinensis* studies may be due to their usage of the microsatellite markers sets which developed through biotinylated-oligonucleotide capture method from *T. ramosissima* Ledeb. (Gaskin, Pepper & Manhart, 2006). These 10 microsatellite markers actually showed relatively low polymorphism, altogether two of which amplified length-variant alleles when PCR products were separated on PAGE gels (Zhang, 2011).

Population genetic differentiation and structure

T. chinensis populations showed a low level of population differentiation ($F_{ST} = 0.0532$) (Table 2), in comparison with the result from its congener *T. taklamakanensis* ($F_{ST} = 0.21$) (Su et al., 2017). The level of population differentiation was consistent with findings from two population genetic studies of *T. chinensis* with SSR markers: (Zhu et al., 2016) ($F_{ST} = 0.0533$) and Liang et al. (2018a) ($F_{ST} = 0.058$). The low level of population differentiation has also been revealed by using RAPD markers (Zhao et al., 2008) and ISSR markers (Jiang, Chen & Bao, 2012). Overall, high outcrossing rate and gene flow facilitate the formation of continuously distributed homogenous populations across its natural range of warm-temperate and subtropical zones, thus resulting in comparatively low population genetic differentiation. With highly polymorphic SSR markers, we robustly identified population genetic structure corresponding to populations' geographical locations. Our results detected three genetic clusters: the northern cluster including the four populations of Yellow River Delta, the fifth Bohai coastal population CY, and Beijing population, the second genetic cluster including two populations of Huang Zhou Bay, and the third genetic cluster containing the remaining Hang Zhou Bay population. The northern (warm-temperate

zone) populations were clustered together as one genetic group and populations of Hang Zhou Bay fell into two other genetic clusters, which may imply the different origins of the studied populations of the two different climate zones. Previous studies revealed more nucleus ribosome DNA (nrDNA) haplotypes in the populations of the lower reaches of the Yellow River, implying this species may have dispersed eastward along the direction of the flowing water (Sun *et al.*, 2019). The continuously distributed populations in the Bohai Sea coastal flats are possibly the results of the northward and southward dispersal of seeds from Yellow River Delta by sea currents, which has yet to be tested by extensive sampling along coastal flats of the Bohai Sea and new polymorphic markers (*e.g.*, chloroplast microsatellite markers) for population genetic analysis. The inland population FS from Beijing was clustered together with populations of Bohai Bay, which is consistent with a study where a Beijing population is genetically closer to a population of the Bohai Sea coastal flat from Korea (Lee, Gaskin & Young, 2018b). The subtropical (the East Sea) populations were genetically different from populations from the Bohai Sea coastal flats, which may be the result of seed dispersal *via* the Yellow River at different times when 150 years ago the Yellow River gorged into the Huanghai Sea instead of the Bohai Sea. Yet we did not include populations from the old estuary flat of the Yellow River in the Huanghai Sea coast in this study.

Isolation pattern of *T. chinensis*

Using Mantel tests, we detected a highly significant correlations between genetic distance and geographical and environmental variables. Because the autocorrelation between geographical and environmental variables is also high ($R^2 = 0.9357$), we performed partial Mantel tests to disentangle the role of geographic distance and environmental variables in shaping the genetic variation detected by microsatellite markers. Our high correlation ($R^2 = 0.5224$ for IBD and $R^2 = 0.5225$ for IBE) between genetic distance and geographical and environmental variables contrasts with no or low significant correlation ($R^2 = 0.1773$, $p < 0.01$, (Liang *et al.*, 2018a); $p = 0.39$, (Zhu *et al.*, 2016)) detected before, which may partly be due to their relatively small sampling areas (Yellow River Delta for Zhu *et al.*, 2016, and the Yellow River reaches for Liang *et al.*, 2018a) and the usage of less variable microsatellite loci.

CONCLUSIONS

In this study, we evaluated the genetic diversity and population structure of *T. chinensis* using 12 SSR primer pairs. *T. chinensis* was found to maintain a high level of genetic diversity. On average, populations of Yellow River Delta showed a higher level of genetic diversity than those from Hang Zhou Bay. The studied populations showed strong IBD and IBE patterns and significant population genetic structure corresponding to sampling sites was detected despite a quite low level of population genetic differentiation. Our results revealed population genetic structure related to environmental variables of two climate zones in a fast spread shrub species. Further extensive sampling and more informative variable loci analysis will enable us to determine the detailed dispersal routine and better understand genetic variation conferring local environmental adaptations of *T. chinensis*.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by the Foundation of Linyi Univeristy (No. A5130673). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Linyi Univeristy: A5130673.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Zhaoyu Jiang performed the experiments, prepared figures and/or tables, and approved the final draft.
- Aoao Yang performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Haiguang Zhang analyzed the data, prepared figures and/or tables, and approved the final draft.
- Wenbo Wang analyzed the data, prepared figures and/or tables, and approved the final draft.
- Ruhua Zhang conceived and designed the experiments, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:
Raw data are available in the [Supplemental Files](#).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.15882#supplemental-information>.

REFERENCES

- Baum BR. 1978.** *The genus Tamarix*. Jerusalem: Israel Academy of Sciences and Humanities, 209.
- Dixon P. 2003.** Vegan, a package of R functions for community ecology. *Journal of Vegetation Science* **14**:927–930 DOI [10.1111/j.1654-1103.2003.tb02228.x](https://doi.org/10.1111/j.1654-1103.2003.tb02228.x).
- Earl DA, Von Holdt BM. 2012.** Structure harvester: a website and program for visualizing structure output and implementing the evanno method. *Conservation Genetics Resources* **4**:359–361 DOI [10.1007/s12686-011-9548-7](https://doi.org/10.1007/s12686-011-9548-7).

- Excoffier L, Laval G, Schneider S. 2005.** Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* **1**:47–50 DOI [10.1177/117693430500100003](https://doi.org/10.1177/117693430500100003).
- Excoffier L, Lischer HEL. 2010.** Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**:564–567 DOI [10.1111/j.1755-0998.2010.02847.x](https://doi.org/10.1111/j.1755-0998.2010.02847.x).
- Fernández MA, Romero RC, Navarro C R, Abril N, Jorrín NJ, Dodd R. 2018.** Population genetic diversity of *Quercus ilex* subsp. ballota (Desf.) Samp. Reveals divergence in recent and evolutionary migration rates in the Spanish Dehesas. *Forests* **9**(6):337–349 DOI [10.3390/f9060337](https://doi.org/10.3390/f9060337).
- Gaskin JF, Pepper AE, Manhart JM. 2006.** Isolation and characterization of ten polymorphic microsatellites in saltcedars (*Tamarix chinensis* and *Tamarix ramosissima*). *Molecular Ecology Notes* **6**(4):1147–1149 DOI [10.1111/j.1471-8286.2006.01466.x](https://doi.org/10.1111/j.1471-8286.2006.01466.x).
- Goudet J. 2001.** FSTAT, a program to estimate and test gene diversities and fixation indices. version 2.9.3 Université de Lausanne..
- Guo B, Hao XC, Han LJ, Zhai Y, Zhou S, Chen S. 2021.** Unraveling the genetic diversity and structure of *Quercus liaotungensis* population through analysis of microsatellite markers. *PeerJ* **9**:e10922 DOI [10.7717/peerj.10922](https://doi.org/10.7717/peerj.10922).
- Hamrick JL, Godt MJW, Shermann-Broyles SL. 1992.** Factors influencing levels of genetic diversity in woody plant species. *New Forests* **6**:95–124 DOI [10.1007/BF00120641](https://doi.org/10.1007/BF00120641).
- Hughes AR, Inouye BD, Johnson MT, Underwood N, Vellend M. 2008.** Ecological consequences of genetic diversity. *Ecology Letters* **11**:609–623 DOI [10.1111/j.1461-0248.2008.01179.x](https://doi.org/10.1111/j.1461-0248.2008.01179.x).
- Islam MS, Lian C, Kameyama N, Hogetsu T. 2012.** Analyses of genetic population structure of two ecologically important mangrove tree species, *Bruguiera gymnorhiza* and *Kandelia obovat* a from different river basins of Iriomote Island of the Ryukyu Archipelago, Japan. *Tree Genetics & Genomes* **8**(6):1247–1260 DOI [10.1007/s11295-012-0512-9](https://doi.org/10.1007/s11295-012-0512-9).
- Jiang B, Gao CQ, Wang YC, Yu LL, Yang CP. 2011.** Cloning and expression analysis of a glycine-rich rna-binding protein gene from *tamarix hispida*. *Forest Research*.
- Jiang ZM, Chen YX, Bao Y. 2012.** Population genetic structure of *Tamarix chinensis* in the Yellow River Delta, China. *Plant Systematics and Evolution* **298**(1):147–153 DOI [10.1007/s00606-011-0532-1](https://doi.org/10.1007/s00606-011-0532-1).
- Jiang XL, Gardner EM, Meng HH, Deng M, Xu GB. 2019.** Land bridges in the Pleistocene contributed to flora assembly on the continental islands of South China: Insights from the evolutionary history of *Quercus championii*. *Molecular Phylogenetics and Evolution* **132**:36–45.
- Jombart T. 2008.** Adegnet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**:1403–1405 DOI [10.1093/bioinformatics/btn129](https://doi.org/10.1093/bioinformatics/btn129).
- Jombart T, Devillard S, Balloux F. 2010.** Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* **11**:245–258 DOI [10.1186/1471-2156-11-94](https://doi.org/10.1186/1471-2156-11-94).

- Kalinowski ST, Taper ML, Marshall TC. 2007.** Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* **16**:1099–1006 DOI [10.1111/j.1365-294x.2007.03089.x](https://doi.org/10.1111/j.1365-294x.2007.03089.x).
- Lee SR, Gaskin JF, Young DK. 2018b.** Molecular diagnosis for a *Tamarix* species from two reclaimed lands along the yellow sea in Korea inferred from genome wide SNP markers. *Journal of Systematics & Evolution* **57**:247–255 DOI [10.1111/jse.12432](https://doi.org/10.1111/jse.12432).
- Lee SR, Jo YS, Park CH, Friedman JM, Olson MS. 2018a.** Population genomic analysis suggests strong influence of river network on spatial distribution of genetic variation in invasive saltcedar across the southwestern US. *Molecular Ecology* **27**(3):636–646 DOI [10.1111/mec.14468](https://doi.org/10.1111/mec.14468).
- Liang HY, Feng ZP, Pei B, Li Y, Yang XT. 2018b.** Demographic expansion of two *tamarix* species along the Yellow River caused by geological events and climate change in the Pleistocene. *Scientific Reports* **147**(1):13–22 DOI [10.1007/s10709-018-0047-6](https://doi.org/10.1007/s10709-018-0047-6).
- Liang HY, Liu CR, Li Y, Wang YC, Kong YH, Quan JN. 2018a.** Low population genetic differentiation in two tamarix species (*Tamarix austromongolica* and *Tamarix chinensis*) along the Yellow River. *Genetica* **147**:13–22 DOI [10.1007/s10709-018-0047-6](https://doi.org/10.1007/s10709-018-0047-6).
- Lind JF, Gailing O. 2014.** Genetic structure of *Quercus rubra* L. and *Quercus ellipsoidalis* E. J. Hill populations at gene-based EST-SSR and nuclear SSR markers. *Tree Genetics & Genomes* **9**:707–722 DOI [10.1007/s11295-012-0586-4](https://doi.org/10.1007/s11295-012-0586-4).
- Luikart G, Allendorf FW, Cornuet JM, Sherwin WB. 1998.** Distortion of allele frequency distributions provides a test for recent population bottlenecks. *Journal of Heredity* **89**(3):238–247 DOI [10.1093/jhered/89.3.238](https://doi.org/10.1093/jhered/89.3.238).
- Naimi B, Hamm NAS, Groen TA, Skidmore AK, Toxopeus AG. 2014.** Where is positional uncertainty a problem for species distribution modelling? *Ecography* **37**:191–203 DOI [10.1111/j.1600-0587.2013.00205.x](https://doi.org/10.1111/j.1600-0587.2013.00205.x).
- Nybom H. 2004.** Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* **13**:1143–1155 DOI [10.1111/j.1365-294X.2004.02141.x](https://doi.org/10.1111/j.1365-294X.2004.02141.x).
- Peakall R, Smouse PE. 2005.** GenAEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**:288–295.
- Piry S, Luikart G, Cornuet JM. 1999.** BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity* **90**:502–503 DOI [10.1093/jhered/90.4.502](https://doi.org/10.1093/jhered/90.4.502).
- Pritchard JK, Stephens M, Donnelly P. 2000.** Inference of population structure using multilocus genotype data. *Genetics* **155**:945–959 DOI [10.1093/genetics/155.2.945](https://doi.org/10.1093/genetics/155.2.945).
- Qaiser M. 1981.** The genus *Tamarix* Linn, (Tamaricaceae) in Pakistan. *Pakistan Journal of Botany* **13**:107–158.
- R Core Team. 2017.** R: A language and environment for statistical computing. Version 3.4.2. Vienna: R Foundation for Statistical Computing. Available at <https://www.r-project.org>.
- Raymond M, Rousset F. 1995.** GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**:248–249 Rice, W. R. 1989 DOI [10.1093/oxfordjournals.jhered.a111573](https://doi.org/10.1093/oxfordjournals.jhered.a111573).

- Rohlf FJ. 2000. NTSYS-pc: numerical taxonomy and multivariate analysis system-version 2.1.
- Rosenberg NA. 2004. Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes* 4:137–138.
- Song ZC. 1958. Tertiary spore and pollen complexes from the red beds of chiuchuan. Kansu and their geological and botanical significance. *Acta Palaeontology Sinica* 6(2):159–167.
- Su YJ, Wang T, Yang WD, Huang C, Fan GK. 1998. DNA extraction and RAPD analysis of *Podocarpus*. *Acta Scientiarum Naturalium Universitatis Sunyatseni* 37:11–18.
- Su Z, Richardson BA, Zhuo L, Jiang XL, Li WJ, Kang XS. 2017. Genetic diversity and structure of an endangered desert shrub and the implications for conservation. *Aob Plants* 9(3):1–10.
- Sun LK, Liu GX, Zhang BG, Zhang GS. 2019. Effects of environmental factors on population genetic diversity of *Tamarix chinensis*. *Acta Prataculturae Sinica* 28(10):178–186 DOI 10.11686/cyxb2018691.
- Van Oosterhout C, Hutchinson WF, Wills DP, Shipley P. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4(3):535–538 DOI 10.1111/j.1471-8286.2004.00684.x.
- Villar JL, Juan A, Alonso MA. 2014. *Tamarix hohenackeri* Bunge, a new record for the flora of Mexico. *Acta Botanica Mexicana* 106:117–128.
- Wang C, Yang C, Gao C, Wang YC. 2009. Cloning and expression analysis of 14 lipid transfer protein genes from *Tamarix hispida* responding to different abiotic stress. *Tree Physiology* 29(12):1607–1619 DOI 10.1093/treephys/tpp082.
- Wang DX, Yang HQ. 1980. Petroleum spore-pollen assemblages and oil source rock of Yumenoil-bearing region in Gansu. *Acta Botanica Sinica* 22(3):280–285.
- Wang HS. 1992. *Floristic geography*. Beijing: Science Press, 84.
- Wang X, Bernhardsson C, Ingvarsson PK. 2020. Demography and natural selection have shaped genetic variation in the widely distributed conifer Norway spruce (*Picea abies*). *Genome Biology and Evolution* 12:3803–3817 DOI 10.1093/gbe/evaa005.
- Zhang DY, Pan BR, Yin LK. 2003. The photogeographical studies of *Tamarix* (Tamaricaceae). *Acta Botanica Yunnanica* 25(4):415–427.
- Zhang R. 2011. Study on the genetic variation of *Tamarix chinensis* populations. Thesis, Nanjing Forestry University.
- Zhang R, Li R, Zhao J, Xu L. 2011. Development and population detection of microsatellite markers from expressed sequence tags for *Tamarix* spp. *Molecular Plant Breeding* 9:1291–1296 DOI 10.5376/mpb.cn.2011.09.0038.
- Zhang R, Wen Q, Xu L. 2019. Development and characterization of genomic SSR markers for *Tamarix chinensis* (Tamaricaceae). *Applications in Plant Sciences* 7(2):e1219 DOI 10.1002/aps3.1219.
- Zhang R, Zhang L. 2019. Effects of different NaCl concentration on the germination of *Tamarix chinensis* seeds. *Journal of Shanxi Agricultural Sciences* 47(7):1160–1162 DOI 10.3969/j.issn.1002-2481.2019.07.10.

- Zhang Y-Y, Fang Y-M, Mukui K, Li X-X, Xia T. 2013.** Molecular characterization and genetic structure of *Quercus acutissima* germplasm in China using microsatellites. *Molecular Biology Reports* **40**:4083–4090 DOI [10.1007/s11033-013-2486-6](https://doi.org/10.1007/s11033-013-2486-6).
- Zhao JK, Li-An Xu, Xie HF, Zhao DY, Huang MR. 2008.** Rapd analysis of population genetic diversity of *Tamarix chinensis* in yellow river delta. *Journal of Nanjing Forestry University* **32**(5):56–60.
- Zhu Z, Zhang LY, Gao LX, Tang SQ, Zhao Y, Yang J. 2016.** Local habitat condition rather than geographic distance determines the genetic structure of *Tamarix chinensis* populations in Yellow River Delta, China. *Tree Genetics & Genomes* **12**(1):14 DOI [10.1007/s11295-016-0971-5](https://doi.org/10.1007/s11295-016-0971-5).
- Zohary M. 1987.** Tamarix L. In: Zohary M, ed. *Flora palaestina 2*. Jerusalem: Israel Academy of Sciences and Humanities, 350–362.