GENETIC DIVERSITY AND DIFFERENTIATION OF *BUPLEURUM CHINENSE* DC. IN THE QINBA MOUNTAINS REVEALED BY NUCLEAR AND CHLOROPLAST MICROSATELLITE MARKERS

PING WANG[1], BING ZHANG[1], PIN LIU[1], YOUNING HU[2] AND HUAIZHU LI[1*]

[1]Department of Life Sciences, Xianyang Normal University, Xianyang, Shaanxi, 712000, People's Republic of China
[2]School of Biological and Environmental Engineering, Xi'an University, Xi'an, Shaanxi, 710002, People's Republic of China

*Corresponding author's email: lihuaizhu121@126.com

Abstract

Bupleurum chinense DC. is an important medicinal plant widely cultivated in the Qinba Mountains of China. In this study, nuclear and chloroplast SSR markers specific to the genus *Bupleurum* were employed to genotype, 232 individuals from 19 *B. chinense* populations across the region, in order to assess genetic diversity and population structure. A low to moderate level of genetic diversity ($nH_E = 0.235$, $cpH = 0.195$) and high genetic differentiation ($nF_{ST} = 0.22$, $cpF_{ST} = 0.53$) were revealed. Principal coordinate analysis, UPGMA phenogram, and STRUCTURE analysis, which were based on the nSSR dataset, divided the 19 populations into three clusters, with no clear genetic structure observed. Genetic differentiation in the cpSSR dataset revealed weak geographical patterns. This was consistent with a non-significant isolation-by-distance pattern, suggesting frequent gene flow among populations. Seed exchange and historical genetic mixing between cultivated and wild populations likely influence the genetic patterns of *B. chinense* in the Qinba Mountains. Our findings provide crucial baseline data for conserving wild genetic resources and managing breeding programs for *B. chinense*.

Key words: Genetic diversity; Genetic differentiation; *Bupleurum chinense*; nSSR; cpSSR

Introduction

Bupleurum chinense DC., a perennial diploid herbaceous plant belonging to the Apiaceae family, has significant botanical and medicinal importance, particularly in China. The dried root, known as Bupleuri radix (Chinese name: *Chaihu*), is widely used to treat cold, fever, influenza, menstrual disorders, and hepatitis (Wang *et al.*, 2023). First recorded in the *Shen Nong's Herbal Classic*, *Chaihu* has been used by Chinese herbalists for over 2,000 years and is a key ingredient in several well-known prescriptions, including Xiao-Chaihu-Tang, Da-Chaihu-Tang, and Chai-Ge-Jieji-Tang. During the COVID-19 pandemic, *Chaihu* proved effective in relieving muscle discomfort by expelling internal heat (Su *et al.*, 2024). The demand for *Chaihu* in China has reached 20,000 tons annually, with artificial plantations becoming the primary source of medicinal material. These plantations are primarily located in Gansu, Shanxi, and Shaanxi provinces (Ma *et al.*, 2020; Zhang *et al.*, 2022). The large-scale agricultural planting of *B. chinense* began in the 1980s, and the current cultivation area is approximately 200,000 acres (Ma *et al.*, 2020). The cultivation area continues to expand, following the frequent trading of seeds. Meanwhile, wild harvesting remains an important source in some rural regions. The challenge arises from the multiple and mixed sources of genotypes of *B. chinense*, which results in the unstable quality of *Chaihu* (Wang *et al.*, 2023).

The Qinba Mountains (30-36° N, 102-114° E), which primarily include the Qinling and Daba Mountains, form a transitional zone between northern and southern China, acting as a major east-west ecological corridor. The mountain range spans across Shaanxi, Gansu, Sichuan,

Chongqing, Hubei, and Henan Provinces (Zhang *et al.*, 2021b). Owing to its complex, diverse, and unique natural environment, the area harbors over 6,000 species of plants and animals, including more than 3,000 species of medicinal herbs (Wang *et al.*, 2018; Yang *et al.*, 2021). The genus *Bupleurum* includes 6 species, 4 subspecies, and 2 forms within this region, of which *B. chinense* is found on sunny slopes and grassy areas across both the northern and southern sides of the range (Fu, 1981). Leveraging its abundant wild resources and strong herbal medicine culture, many areas in the Qinba Mountains began cultivating *Chaihu* in the 1980s and has since developed into a key industry. In Chencang District, Shaanxi Province, the cultivated area exceeds 100,000 acres, and "Xishan Chaihu" has earned national geographic indication status. To ensure the sustainable utilization and development of *B. chinense*, understanding its genetic diversity and population structure is crucial, especially in the face of human interventions like cultivation and wild harvesting (Wu *et al.*, 2023). However, few studies have been focused on the genetic diversity and population structure of *B. chinense* resources.

Molecular markers, particularly at the DNA level, offer a reliable method for assessing genetic diversity in plant germplasm collections. Among these, simple sequence repeats (SSRs) are commonly used due to their high reproducibility, polymorphism, and codominant inheritance (Powell *et al.*, 1996; Hameed *et al.*, 2023). SSRs have proven effective in exploring genetic diversity within the Apiaceae family, including species like *Helosciadium repens* (Herden *et al.*, 2020), *Angelica dahurica* (Huang *et al.*, 2022), and *Glehnia littoralis* (Li *et*

al., 2024). However, studies on the genus *Bupleurum* remain limited. Nuclear SSR markers (nSSRs) are biparentally inherited, highly reproducible, and useful for studying population differentiation. In contrast, chloroplast SSR markers (cpSSRs) are maternally inherited, highly conserved, and evolve slowly, making them valuable tools for studying genetic relationships and historical dynamics (Djedid *et al.*, 2021; Peng *et al.*, 2024). The combined use of nSSR and cpSSR markers offers a comprehensive approach for analyzing population genetic diversity and provides insights into species history and dynamics.

In this study, 6 nSSR and 6 cpSSR combinations were utilized to evaluate the genetic diversity and differentiation of 19 *B. chinense* populations in the Qinba Mountains. The research aims to (1) estimate the genetic diversity of *B. chinense* populations across the region, (2) assess the genetic structure of the populations, and (3) provide foundational information for the conservation and breeding management of *B. chinense*.

Experimental methods

Species sampling: A total of 232 individuals representing 19 *B. chinense* populations were collected across the Qinba Mountains in China (Fig. 1; Table S1). To represent the broad distribution of *B. chinense* across the Qinba Mountains, we stratified our sampling into two major geographical units which differ in climate and vegetation (Liu *et al.*, 2024): the northern part of the Qinling Range (NQL), the southern part of the Qinling Range (SQL,

including the southern Qinling Mountains and Bashan Mountains). Each population consisted of 4-20 individuals, with a minimum distance of 10 meters between sampled plants. Fresh leaf samples were collected and dried with silica gel. Voucher specimens were deposited at the Herbarium of Xianyang Normal University.

DNA extraction, PCR amplification, and sequencing: Genomic DNA extraction was carried out following the modified CTAB method (Wang *et al.*, 2015). A set of seven chloroplast SSR primers (cpSSR) and twenty nuclear SSR primers (nSSR) were initially selected based on expected polymorphisms within the genus *Bupleurum* (Wang *et al.*, 2023). After evaluation, six cpSSRs and six nSSRs with good polymorphism and high reproducibility were chosen for further analysis (Table S2). Forward primers were labeled with FAM, HEX or TAM. Multiplex PCRs were performed by combining three or four primer pairs per multiplex. The PCR reaction was conducted in a 20-µL volume containing 1 µL of template DNA (about 50 ng), 0.5 µL of each primer, 10 µL of PCR Premix (Premix Taq Hot Start Version, TakaRa), and 8 µL of ddH$_2$O. A touchdown PCR program was applied, where the annealing temperature decreased from 65°C to 55°C in 1°C steps over the first 10 cycles, followed by a constant annealing temperature of 55-60°C for the remaining 28 cycles. PCR products were analyzed via capillary electrophoresis using an ABI 3730XL automatic sequencer (Applied Biosystems). Chromatograms were analyzed with GeneMarker (Hulce *et al.*, 2011) and checked manually.
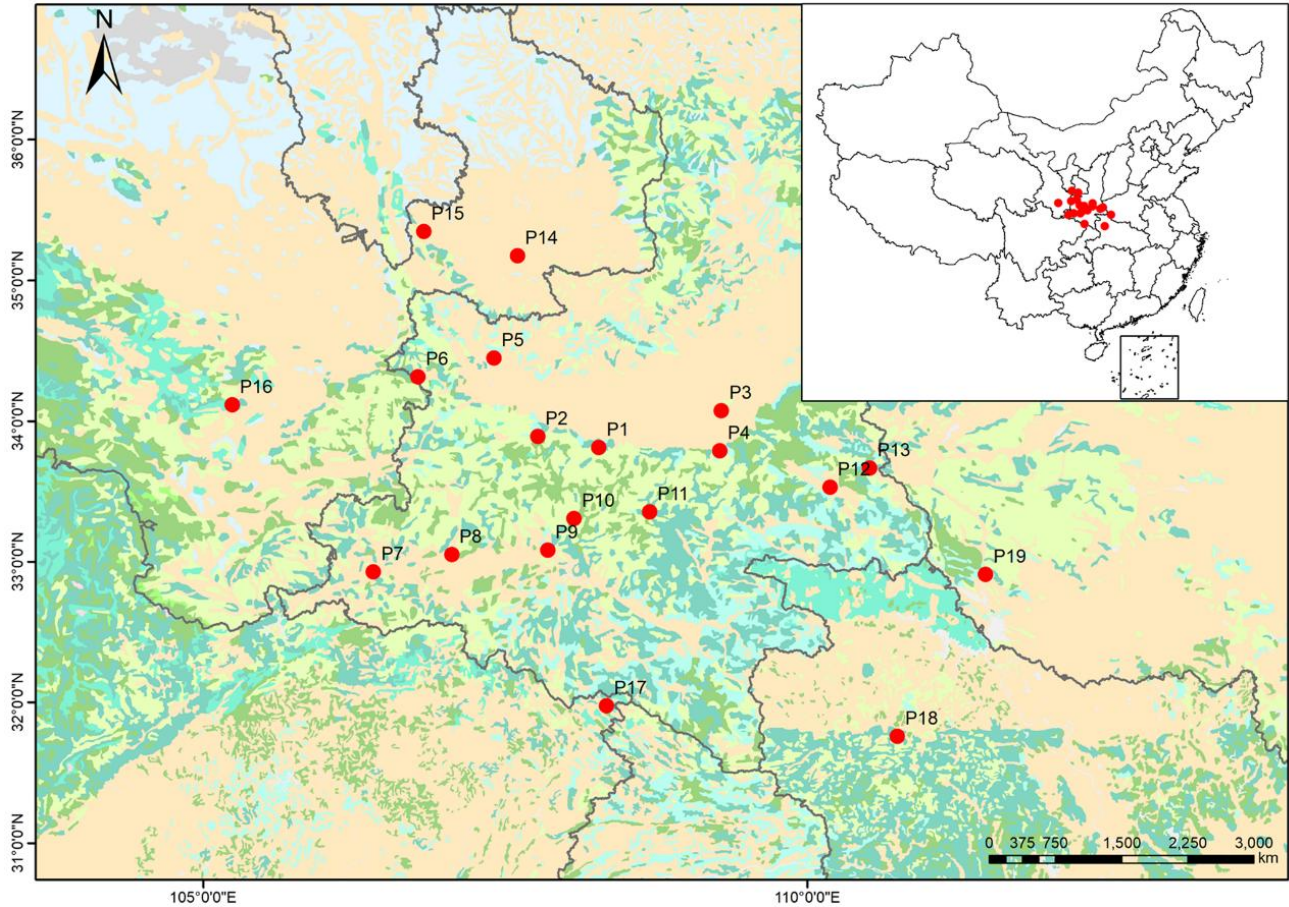

Fig. 1. Geographical distribution of 19 *Bupleurum chinense* populations across the Qinba Mountains.

Table S1. Sampling information of the 19 *B. chinense* populations in the study.

Population code	Locations	Region	No. of sample	Latitude, Lonitude	Type	No. of voucher
P1	Zhouzhi, Shaanxi Province	NQL	12	34.019N,108.175E	wild	ZZYS001-003
P2	Yingge, Shaanxi Province	NQL	10	34.084N,107.649E	Cultivated	YGZP001-003
P3	Lintong, Shaanxi Province	NQL	14	34.302N,109.221E	wild	LTYS001-003
P4	Lantian, Shaanxi Province	NQL	12	34.013N,109.214E	wild	LTYS001-003
P5	Qianyang, Shaanxi Province	NQL	15	34.635N,107.247E	Cultivated	QYZP001-003
P6	Tuoshi, Shaanxi Province	NQL	10	34.478N,106.599E	Cultivated	TSZP001-003
P7	Ningqiang, Shaanxi Province	SQL	15	33.068N,106.293E	wild	NQYS001-003
P8	Mianxian, Shaanxi Province	SQL	14	33.216N,106.951E	wild	MXYS001-003
P9	Yangxian, Shaanxi Province	SQL	15	33.273N,107.765E	wild	YXYS001-003
P10	Foping, Shaanxi Province	SQL	15	33.504N,107.982E	wild	FPYS001-003
P11	Ankang, Shaanxi Province	SQL	4	33.565N,108.626E	wild	AKYS001-003
P12	Shangluo, Shaanxi Province	SQL	9	33.763N,110.163E	Cultivated	SLZP001-003
P13	Shangluo, Shaanxi Province	SQL	15	33.900N,110.502E	wild	SLYS001-003
P14	Pingliang, Gansu Province	NQL	9	35.371N,107.420E	wild	PL01YS001-003
P15	Pingliang, Gansu Province	NQL	15	35.518N,106.594E	wild	PL02YS001-003
P16	Longnan, Gansu Province	NQL	14	34.211N,105.018E	wild	LNYS001-003
P17	Wanyuan, Sichuan Province	SQL	4	32.170N,108.297E	wild	WYYS001-003
P18	Fangxian, Hubei Province	SQL	15	31.978N,110.744E	wild	FXYS001-003
P19	Xichuan, Henan Province	SQL	15	33.138N,111.491E	wild	XCYS001-003

Table S2. Characterization of the 6 nSSRs and 6 cpSSRs.

Locus	Forward primer sequence (5'-3')	Reverse primer sequence (5'-3')	Type	Size (bp)
nSSR				
nr01	GAGAGAAATGAAGCAGTGTTG	GACATCAGATAGCCAGTAAAG	AAAG	194-202
nr02	CTATGCATTTTCTAGCTCCAC	TCAACTCCTCCTCTAAGTCAT	AAAT	204-224
nr03	CACAGTGTCTTGAACTAAAC	GAGCCAAAGTACACCATTAT	ATG	156-159
nr06	GAAGTCCATGAGTAATGATAG	CATGATTTTCACCTCCAATGC	CCT	216-222
nr08	GTTATGATGATGAGGAGGCG	GATGGATGAAGAAGGCGAAGA	GCT	210-222
nr11	GTGTCAGAACCAGAGCAATAA	GCCCTTACTTACAACATCACA	AAG	193-211
cpSSR				
cp01	GCCTTTCTTTGATTACTTTG	ACGCCTCCATTATACACG	AT	209-221
cp02	AGTTCAAATCTGGTTCCTGGC	AAAGACCGAGATGGAGAAAAG	AT	154-172
cp03	TTTGCCTCCTACTAATGAGTG	GAGGACATGCAAAAAGCGATA	AT	112-122
cp04	TTTTTTTTGGAATCCCCTCG	ACTAAACTATACCCGCCACA	ATT	160-166
cp05	TGGCATCTTGAATCATATCC	GGAATTGTTGATCTTGTAGC	ATTTC	197-202
cp07	CTTGTTTTCATCTTACCTAGGA	CCCAACACCAATCTATTCTAAA	TA	308-314

Genetic diversity and structure: For the nuclear SSR dataset, genetic diversity parameters including observed alleles (Na), Shannon's information index (I), observed heterozygosity (nHo), expected heterozygosity (nHe), and fixation index (F) were calculated using GenAlEx v.6.5 (Peakall & Smouse, 2012). The FSTAT v.2.9.4 software (Goudet, 2003) was used to assess interpopulation differentiation (F_{ST}) and inbreeding (F_{IS}). For the chloroplast SSR dataset, genetic diversity was analyzed using Haplotype Analysis v.1.05 (Eliades & Eliades, 2009), with parameters such as the number of haplotypes (cpA), number of private haplotypes (cpP), effective number of haplotypes ($cpNe$), haplotypic richness (Rh), haplotype diversity (cpH), and mean within-population genetic distance between haplotypes (D^2sh). F_{ST} for cpSSRs was also calculated using FSTAT.

To explore genetic clustering among the 19 *B. chinense* populations, principal coordinate analysis (PCoA) was performed in GenAlex. Bayesian cluster analysis was carried out in STRUCTURE v2.3.4 (Pritchard *et al.*, 2000) with a burn-in of 10,000 steps, followed by 100,000 iterations. The number of subpopulations (K) was tested ranging from 1 to 10, with 10 repetitions for each K. The optimal K was determined using the Evanno method implemented in StructureSelector (Li & Liu, 2018). Population differentiation was quantified through molecular variance analysis (AMOVA) in Arlequin v.3.5, with significance assessed via 1,000 permutations (Excofer & Lischer, 2010). A UPGMA phylogenetic tree was constructed based on Nei's genetic distance using MEGA_X (Kumar *et al.*, 2018). The correlation between genetic and geographic distances was assessed using a Mantel test in GenAlEx, calculating pairwise $F_{ST}/(1-F_{ST})$ and the natural logarithm of geographic distances, with 1,000 permutations (Smouse & Long, 1992). Phylogeographic patterns were examined by comparing G_{ST} and R_{ST} values using SPAGeDi v.1.3 with 10,000 permutations based on the cpSSR dataset (Hardy & Vekemans, 2002). A median-joining (MJ) network was constructed from the cpSSR haplotypes using Network v.10.2 (Bandelt *et al.*, 1999).

Population demographic analysis: Genetic bottlenecks were tested using Bottleneck v.1.2.02 (Piry *et al.*, 1999), applying Wilcoxon's sign-rank test under the two-phased model (TPM) and a mode-shift test to detect recent population size reductions. Migrate-n v.5.0.4 was conducted to estimate the migration rates (M) among *B. chinense* populations in the Qinba Mountains (Beerli & Felsenstein, 1999). A Brownian motion mutation model with constant mutation rates and likelihood approach was implemented. Markov Chain Monte Carlo (MCMC) simulations were conducted with a long chain (increment = 100; genealogies sampled = 100,000, genealogies recorded per chain = 5000; burn-in = 10,000).

Results

Genetic diversity of *B. chinense* in the Qinba Mountains: Across the 19 populations of *Bupleurum chinense* in the Qinba Mountains, a total of 129 alleles were detected from the nSSR markers. The number of alleles per population ranged from 3 to 14, with a mean of 9.5 (Table 1). Genetic diversity within populations varied, with the number of observed alleles (*Na*) ranging from 1.500 in populations P11 and P17 to 2.833 in P7. The Shannon's information index (*I*) averaged 0.296, with a minimum value at 0.250 for P17 and a maximum value at 0.606 for P7. Observed

(n*Ho*) and expected (n*He*) heterozygosity averaged 0.186 and 0.235, respectively. Population P3 exhibited the lowest values for both n*Ho* and n*He*, coupled with a relatively high inbreeding index (F_{IS} = 0.756). Cultivated populations such as P2, P6, and P12 showed lower n*He* values, whereas P5 exhibited a relatively higher n*He*. The average *F* value across populations was 0.076, with values ranging from -0.123 in P14 to 0.783 in P3. The F_{IS} values deviated from zero for all the populations except for P14, which showed a negative F_{IS} indicating an excess of heterozygosity, all other populations with positive F_{IS} values exhibited a deficit of heterozygosity.

Table 1. Genetic parameters of the 19 *B. chinense* populations based on 6 nSSRs markers.

Population code	n*A*	*Na*	*I*	n*Ho*	n*He*	*F*	F_{IS}	*p* of TPM	Mode shift test
P1	9	2.167	0.457	0.167	0.267	0.307	0.413	0.875	L-shaped
P2	7	1.667	0.264	0.117	0.144	0.079	0.241	0.500	shifted
P3	7	2.000	0.266	0.036	0.137	0.783	0.756	0.125	L-shaped
P4	8	2.167	0.327	0.139	0.171	0.219	0.228	0.125	L-shaped
P5	11	2.000	0.517	0.233	0.328	0.279	0.319	0.125	L-shaped
P6	9	2.333	0.433	0.200	0.233	0.257	0.194	0.125	L-shaped
P7	12	2.833	0.606	0.267	0.327	0.130	0.217	0.156	L-shaped
P8	13	2.167	0.439	0.190	0.237	0.350	0.232	1.000	shifted
P9	14	2.333	0.430	0.189	0.235	0.090	0.23	0.313	L-shaped
P10	13	2.500	0.510	0.167	0.274	0.160	0.421	0.219	L-shaped
P11	3	1.500	0.303	0.167	0.208	0.333	0.333	0.625	shifted
P12	6	2.000	0.302	0.167	0.166	0.005	0.053	0.250	L-shaped
P13	10	2.000	0.372	0.189	0.210	-0.007	0.135	0.375	shifted
P14	7	1.833	0.411	0.315	0.278	-0.123	-0.075	1.000	L-shaped
P15	13	2.167	0.593	0.300	0.383	0.395	0.249	0.039	shifted
P16	9	1.833	0.281	0.095	0.143	0.167	0.368	0.500	L-shaped
P17	3	1.500	0.250	0.125	0.161	0.175	0.357	1.000	shifted
P18	14	2.333	0.484	0.233	0.284	0.089	0.213	0.844	L-shaped
P19	13	2.000	0.464	0.244	0.276	0.221	0.149	0.313	shifted
Mean	9.5	2.070	0.406	0.186	0.235	0.206	0.265		

nA Number of alleles, *Na* Number of observed alleles, *I* Shannon's information index, *nHo* Observed heterozygosity, *nHe* Expected heterozygosity, *F* Fixation index, *F* Inbreeding coefficient

Table 2. Genetic diversity indices of the 19 *B. chinense* populations based on 6 cpSSRs.

Population code	cp*A*	cp*P*	cp*Ne*	*Rh*	cp*H*	D^2sh
P1	2	1	1.180	0.333	0.051	1.778
P2	6	0	5.000	2.367	0.407	66.430
P3	4	2	2.649	1.432	0.128	4.051
P4	5	2	3.130	1.788	0.292	11.515
P5	5	0	3.689	1.885	0.344	50.210
P6	6	1	5.000	2.367	0.363	38.374
P7	3	1	1.744	0.952	0.207	29.460
P8	3	1	1.342	0.571	0.043	0.505
P9	4	0	3.169	1.658	0.265	36.571
P10	4	0	3.000	1.611	0.224	34.375
P11	1	0	1.000	0.000	0.000	0.000
P12	2	0	1.976	0.952	0.247	44.815
P13	5	2	2.922	1.595	0.298	33.816
P14	2	0	1.528	0.722	0.058	0.259
P15	3	3	1.316	0.533	0.041	0.190
P16	3	0	1.342	0.571	0.189	25.722
P17	1	0	1.000	0.000	0.000	0.000
P18	3	1	1.510	0.743	0.163	12.356
P19	6	1	3.947	1.968	0.378	70.997
Mean	3.6	0.8	2.444	1.160	0.195	24.285

cpA Number of haplotypes detected in each population, *cpP* Number of private haplotypes, *cpNe* Effective number of haplotypes, *Rh* Haplotypic richness, *cpHe* Genetic diversity, D^2sh Mean genetic distance between individuals
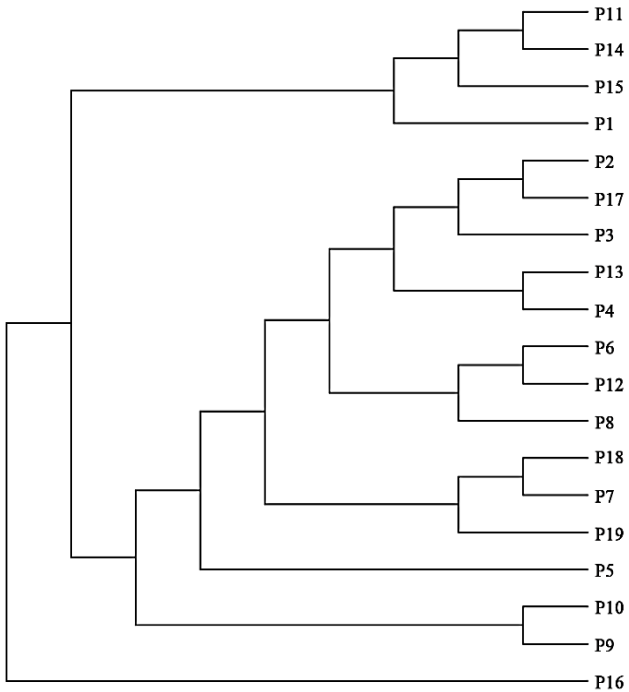

Fig. 3. UPGMA dendrogram depicting the relationships among 19 *Bupleurum chinense* populations based on nSSR markers.

A total of 28 chloroplast haplotypes were detected across the six cpSSR markers, 15 of which were private. The number of haplotypes per population varied from 1 (P11 and P17) to 6 (P19) (Table 2). Populations P2, P6, and P19 harbored the highest number of haplotypes (cp$A$ = 6), while P2 and P6 lacked private haplotypes. The highest number of private haplotypes was observed in P15 (cp$P$ = 3), followed by P3, P4 and P13 (cp$P$ = 2). The effective number of haplotypes (cpN_e) was highest in populations P2 and P6, followed by P19. Populations P11 and P17, with limited sample sizes (N = 4), showed the lowest values in cpNe, haplotype richness (Rh), and genetic diversity (cp$H$). The Rh values ranged from 0 to 2.367. The average cp$H$ was 0.075, with values ranging from 0 (P11 and P17) to 0.407 (P2). Notably, cultivated populations P2, P5, P6, and P12 showed high chloroplast genetic diversity. The D^2sh values varied greatly from 70.997 (P19) to 0 (P11 and P17).

Genetic structure of *B. chinense* in the Qinba Mountains: The 19 populations were clustered into 3 main groups in the PCoA plot based on the nSSR dataset (Fig. 2a). The first two principal components explained 36.77% and 22.75% of the total molecular variance, respectively. Cluster I consisted of 5 populations (P1, P11, P14, P15, and P18), while Cluster II included most of the populations (P2, P3, P4, P5, P6, P7, P8, P9, P10, P12, P13, P17, P18, and P19). Population P16 formed a distinct group, referred to as Cluster III. Meanwhile, the UPGMA tree revealed a similar clustering pattern, with the exception of population P18 (Fig. 3). The optimal number of genetic groups (K = 3) was supported by Bayesian structure analysis, which grouped the 232 individuals into three clusters, with noticeable individual admixture (Fig. 4). The AMOVA revealed that 22.34% of genetic variation occurred among populations based on the nSSR dataset (Table 3). Furthermore, the Mantel test showed no significant correlation between genetic and geographic distances (r = 0.038, P = 0.380) (Fig. 5a).

The cpSSR analysis revealed that the differentiation value for the G_{ST} (0.547) was slightly higher than that for the R_{ST} (0.520), suggesting that the populations did not exhibit a strong phylogeographic structure. This finding was consistent with the Mantel test results conducted using the cpSSR dataset, which revealed no significant correlation between genetic and geographical distances

(Fig. 5b). However, AMOVA revealed that 52.45% of the genetic variation was among populations, with the remaining 47.55% occurring within populations (Table 3). The PCoA result from cpSSRs clustered the 19 populations into two main groups (Fig. 2b). Group I was composed of 10 populations, most of which were distributed in the northern part of the Qinling Mountains (P1, P3, P4, P6, P14, P15, and P16), with the exception of P11, P17 and P18. Group II contained the remaining 9 populations, among which P2 and P5 were from the northern part of the Qinling Mountains, while P7, P8, P9, P10, P12, P13, and P19 were from the southern part. Notably, population P19 originated specifically from the Bashan Mountains. In the network analysis, H15 was the most abundant haplotype, occurring in 52.6% of the populations, followed by H21 and H27, respectively (Fig. 6). Notably, H15 occupied the central position within the network and was present in populations from both Group I and Group II. The populations in Group I exhibited the greatest haplotype diversity, with a total of 20 haplotypes.

Gene flow analysis: Wilcoxon's sign-rank test under the TPM model revealed no significant heterozygosity excess in all populations except for P15, indicating a lack of strong evidence for recent bottleneck events (Table 1). In contrast, the mode-shift test detected bottleneck events in populations P2, P8, P11, P13, P15, P17, and P19, as evidenced by a shifted allele frequency distribution (Table 1). The TPM is the most sensitive method for detecting bottlenecks occurring over approximately the last 2 Ne to 4 Ne generations. Additionally, the mode-shift indicator test, which is based on the allele frequency distribution, is most appropriate for detecting population declines that have occurred more recently, specifically over the last few dozen generations (Luikart & Cornuet, 1998; Yuan *et al.*, 2014).

Migration rates among populations were considerably higher in the nuclear dataset than in the chloroplast dataset (Fig. 7). The migration rates ranged from 0.87 (P15 to P16) to 9.33 (P16 to P17) for nSSRs, and from 0.18 (P8 to P6) to 9.36 (P10 to P9) for cpSSRs. Cultivated populations P2, P5, P6 and P12 exhibited higher migration rates in most comparisons. Clear asymmetric migration patterns were observed in populations P11 and P17, where outflow gene flow exceeded inflow gene flow.
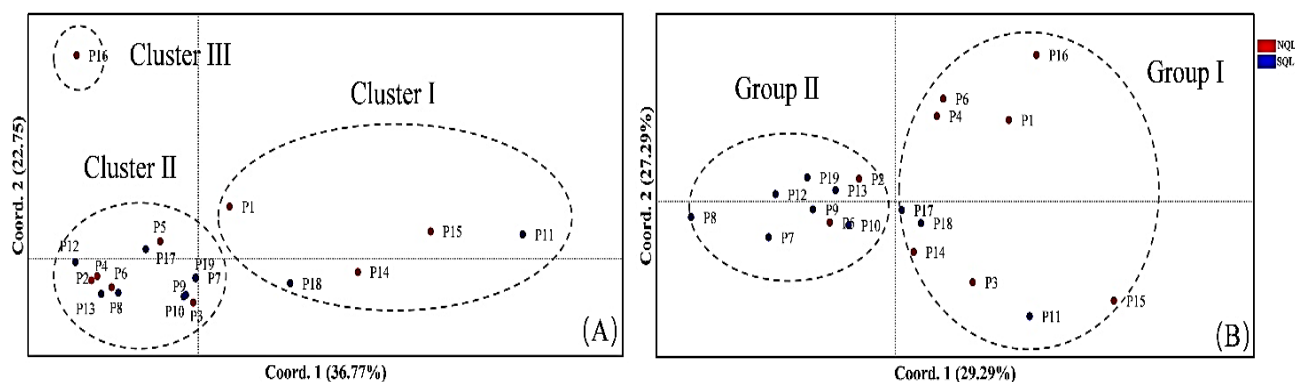

Fig. 2. PCoA analysis of 19 *Bupleurum chinense* populations based on (a) the nSSR dataset, (b) the cpSSR dataset.
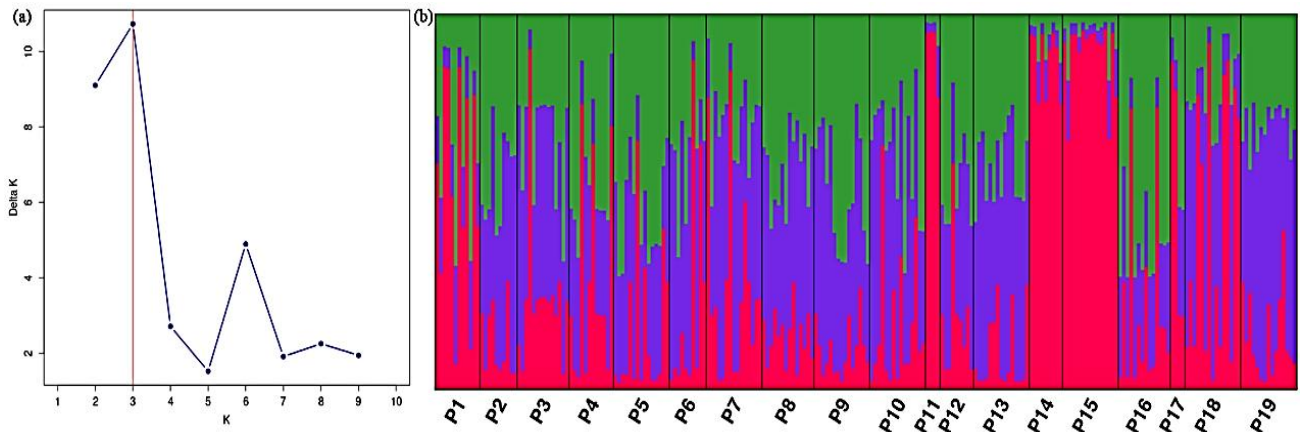
Fig. 4. Bayesian clustering results of 19 *Bupleurum chinense* populations estimated using STRUCTURE software based on the nSSR dataset. (a) The optimal *K* determined by *ΔK* likelihood; (b) Genetic structure with *K* = 3.
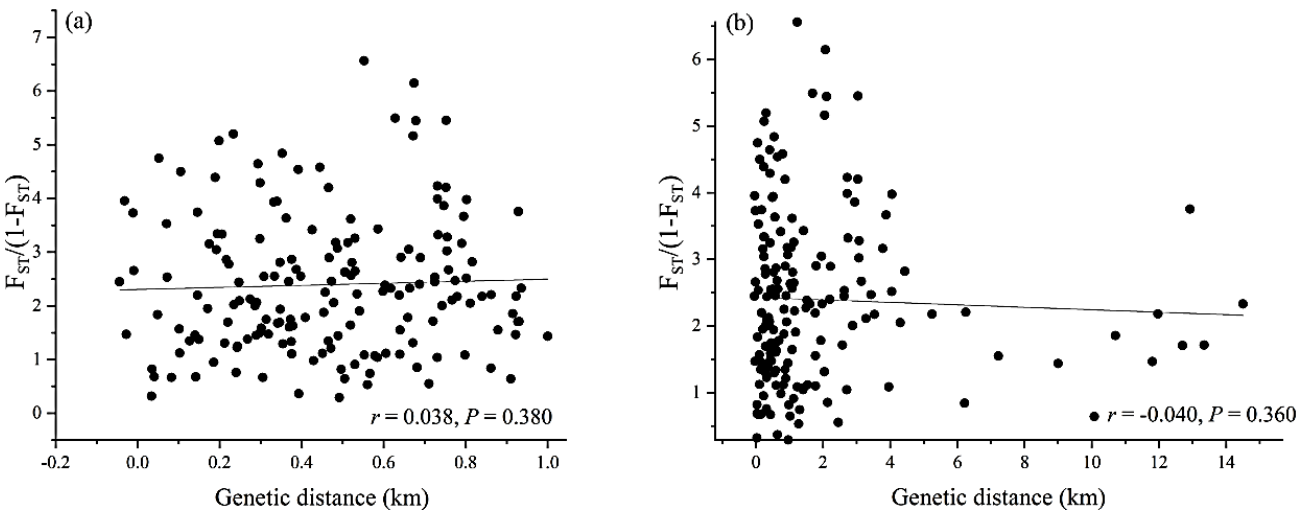

Fig. 5. Mantel test of the correlation between $F_{ST}/(1-F_{ST})$ and geographic distances among the studied populations.

Table 3. Analysis of molecular variance (AMOVA) of *B. chinense* based on nuclear and chloroplast microsatellites.

	Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation (%)	*F*-statistics	*p*-value
nSSR	Among populations	18	102.277	0.20049 Va	22.34	F_{ST} = 0.22	0.000
	Within populations	445	359.881	0.80872 Vb	77.66		
cpSSR	Among populations	18	326.513	0.71942 Va	52.45	F_{ST} = 0.53	0.000
	Within populations	445	290.254	0.65226 Vb	47.55		

Discussion

Understanding genetic diversity and population structure of species is fundamental for utilizing breeding materials and enhancing breeding efficiency. The Qinba Mountains, recognized for the rich plant species diversity, are a notable biodiversity hotspot in China (Zhang *et al.*, 2022). Species endemic to or predominantly distributed in the region, such as *Akebia trifoliata* (Zhang *et al.*, 2021b) and *Actinidia chinensis* (Wang *et al.*, 2018), often show high genetic diversity. In this study, *B. chinense* populations in the Qinba Mountains exhibited low to moderate genetic diversity (nH_E = 0.235, cp$H$ = 0.195). This finding aligns with the central-marginal hypothesis, which suggests that smaller, isolated peripheral populations tend to have lower genetic diversity compared to central populations (Wei *et al.*, 2016).

However, factors such as population size, climate change, and human activities also contribute to genetic variation (Peng *et al.*, 2024). *B. chinense*, a moderately xerophytic species, is typically found in grasslands and shrub thickets (Sheh & Mark, 2005). Field observations and specimen data indicate that *B. chinense* populations are widespread across the Loess Plateau but less common in the southern Qinling Mountains. The distribution pattern suggests that the Qinba Mountains likely represent peripheral populations in China, as predicted by the Maxent model (Chen *et al.*, 2022). Populations from Shanxi and surrounding areas, mainly within the Loess Plateau, showed relatively high genetic diversity (average H_E = 0.66) (Song *et al.*, 2022). In the southeastern edge of the Qinling Mountains, specifically in Funiu Mountain, *B. chinense* populations exhibited low genetic diversity (Jiang *et al.*, 2022). This trend was mirrored in *B.*

rotundifolium populations in Central Europe, where habitat rarity led to lower genetic diversity (Brütting *et al.*, 2012). Although comparative data on the genetic diversity of *B. chinense* across its global range are limited, our findings suggest a low level of genetic diversity in the Qinba Mountains. The low genetic diversity could result from inbreeding, as most populations showed positive inbreeding coefficients (average $F_{IS} = 0.18$), indicating prevalent self-fertilization and biparental inbreeding (Salehi Shanjani *et al.*, 2010). Although *B. chinense* is a long-lived, predominantly outcrossing species, the limited availability of pollinators and small population sizes increase the probability of self-pollination, as observed in other Apiaceae species (East, 1940; Koul *et al.*, 1989; Wang *et al.*, 2015). Anthropogenic activities, such as urbanization, agriculture, and land restoration projects, have reduced suitable habitats and exacerbated population decline of *B. chinense* in the Qinba Mountains. This is starkly evident in populations P11 and P17, where no cpSSR diversity was detected. Overexploitation for medicinal use presents an additional threat.
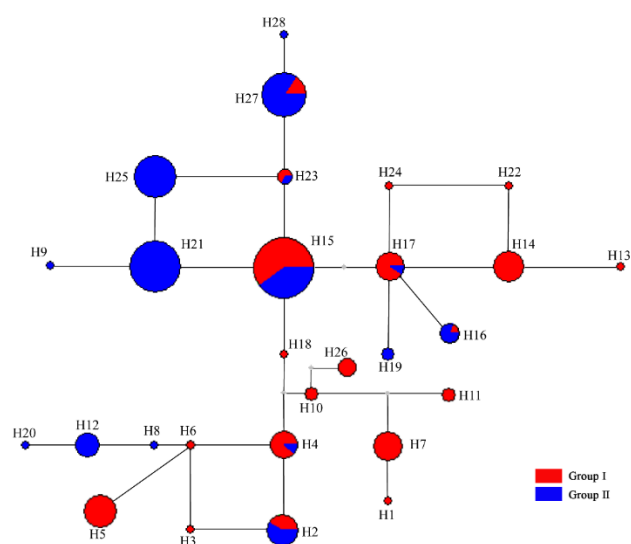

Fig. 6. Median-joing (MJ) network of chloroplast haplotypes. Circle size is proportional to haplotype frequency. Distinct colours represent different genetic groups inferred from PCoA results based on cpSSR markers.

Interestingly, despite the overall pattern, several populations (e.g., P2, P4, P5, P6, P9, P12, P13, P16, P19) displayed higher diversity for chloroplast markers (cp*H*) than for nuclear markers (n*H_e*). For maternally inherited species, nSSR markers typically reveal greater polymorphism than cpSSRs (Phumichai *et al.*, 2015; Zhu *et al.*, 2022). This increase in genetic diversity could be attributed to seed exchange between populations, which promotes gene flow. Studies on Chinese-cultivated apricot (Zhang *et al.*, 2021a) and other species (Bonnave *et al.*, 2014; Brusa & Holzapfel, 2018) have shown that seed exchange, often influenced by human activity, could increase genetic variation. In the case of *B. chinense*, seed exchange, primarily from Shaanxi and Gansu Provinces, likely facilitated genetic mixing across cultivation regions (Jiang *et al.*, 2022; Zhang *et al.*, 2022). The loss

of private haplotypes in several populations further supported this hypothesis. Specifically, no private haplotypes were found in P2, P9, and P12, and only one private haplotype was detected in P6 and P19. Seed exchange may thus play a role in driving higher genetic diversity in some populations, as observed with maternally inherited chloroplast markers.

Our analysis revealed significant genetic differentiation among the populations (n$F_{ST} = 0.22$, cpF_{ST} = 0.53), as defined by Wright (1978). Generally, peripheral populations are expected to exhibit greater differentiation (Dai & Fu, 2011; Islam *et al.*, 2014). PCoA, UPGMA, and STRUCTURE analysis based on the nSSR markers grouped 19 *B. chinense* populations into 3 clusters; however, no clear genetic structure was identified. STRUCTURE analysis revealed genetic admixtures within individuals among populations, and the Mantel test revealed no significant isolation-by-distance ($r = 0.038$, $p = 0.380$). The observed spatial pattern of low nuclear genetic differentiation and high chloroplast genetic differentiation may reflect gene flow through reproductive interactions and spatial movements among populations, as observed in other species (Parra-Rondinel *et al.*, 2021). Seed exchange, as previously discussed, likely played a significant role in genetic mixing. Furthermore, synchronized flowering periods, shared flower visitors, and potential cross-pollination may further facilitate gene flow between wild and cultivated populations, as seen in vegetatively propagated crops (Bonnave *et al.*, 2014; Bonnave *et al.*, 2016; Moscoe & Emshwiller, 2016; Parra-Rondinel *et al.*, 2021). Historical gene flow analysis indicated frequent gene flow exchanges among populations. Higher nuclear migration rates suggested that pollen played a more significant role in maintaining individual admixtures across populations. However, seed flow appeared to occur among populations, as indicated by the higher nuclear and chloroplast migration rates in cultivated populations P2, P5, P6, and P12. Chloroplast networks revealed shared haplotypes between cultivated and wild populations, further supporting this gene flow.

Theoretically, chloroplast SSR markers are useful for evaluating lineages of maternal origin, as they are dispersed by seeds and evolve at a slower rate, reflecting long-term historical gene flow (Phumichai *et al.*, 2015). The level of differentiation among populations based on cpSSR markers (cpF_{ST} = 0.53) was higher than that observed with nSSR markers, as expected in angiosperms (Hu & Ennos, 1997; Salehi Shanjani *et al.*, 2010). Although the Mantel test for cpSSR markers showed no significant correlation between genetic and geographical distances, the PCoA clustering reflected, to some extent, the geographical distribution patterns. Populations from the northern and southern parts of the Qinling Mountains formed two distinct groups, with the exceptions of P2, P5, P11, P17, and P18. Among these, P2 and P5 are cultivated populations, while P11, P17, and P18 are located near former plantations. Additionally, the multiple clusters of cultivated populations suggested that the cultivation of *B. chinense* in the Qinba Mountains might have multiple origins (Wills, 2006; Li *et al.*, 2024).

Studying genetic diversity and structure of a species provides a critical foundation for developing effective conservation strategies. Wild populations are important genetic resources, contributing genes that enhance resistance to biotic and abiotic stresses and can broaden the gene pool of cultivated species (Warschefsky *et al.*, 2014). Therefore, conserving wild populations across their entire geographic range is vital. On the basis of haplotype distribution data, wild populations of *B. chinense* in the Qinba Mountains harbor substantial private genetic resources. Evidence of a recent bottleneck in 7 of 19 *B.*

chinense populations, mostly wild populations, highlights the need for targeted conservation efforts. With ongoing habitat loss, *ex situ* conservation strategies should be prioritized, especially for populations with high genetic diversity, which can serve as critical sources for future breeding efforts (Wang *et al.*, 2020). Furthermore, it is essential to establish a dedicated seed production base to minimize the mixing of seed resources, which could lead to instability in the quality of *Chaihu*. This study provides a genetic foundation for the classification, breeding, and sustainable utilization of *B. chinense* germplasm.
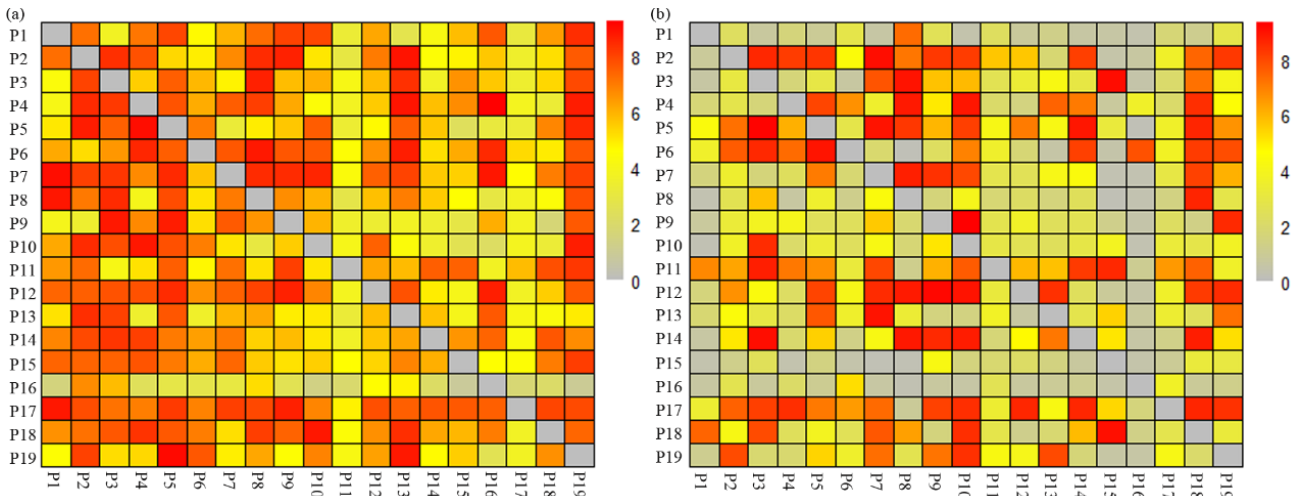

Fig. 7. Migration rates (M) among 19 *B. chinense* populations based on (a) the nSSR dataset, (b) the cpSSR dataset.

Conclusion

This study employed 6 nuclear SSR markers and 6 chloroplast SSR markers to assess the genetic diversity and structure of *B. chinense* populations in the Qinba Mountains of China. Our findings revealed a low to moderate level of genetic diversity and a high level of genetic differentiation among populations. These patterns primarily attributed to the combined effects of peripheral distribution, frequent seed exchange, and historical gene flow. The results provide valuable insights for the conservation and breeding management of *B. chinense* resources.

Authors Contribution: Wang P conceived and designed the experiments. Li HZ and Zhang B conducted the field sampling. Li HZ and Hu YJ performed the experiments. Li HZ and Liu P analyzed the data and drafted the manuscript. Wang P and Li HZ contributed to the language editing. All the authors approved the final manuscript for submission.

Conflict of Interest: The authors have no financial or proprietary interests in any material discussed in this article.

Acknowledgements

References

Bandelt, H.J., P. Forster and A. Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.*, 16(1): 37-48.

Beerli, P. and J. Felsenstein. 1999. Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. *Genetics*, 152(2): 763-773.

Bonnave, M., G. Bleeckx, J. Rojas Beltrán, P. Maughan, M.C. Flamand, F. Terrazas and P. Bertin. 2014. Farmers' unconscious incorporation of sexually-produced genotypes into the germplasm of a vegetatively-propagated crop (*Oxalis tuberosa* Mol.). *Genet. Resour. Crop Ev.*, 61: 721-740.

Bonnave, M., T. Bleeckx, F. Terrazas and P. Bertin. 2016. Effect of the management of seed flows and mode of propagation on the genetic diversity in an Andean farming system: the case of oca (*Oxalis tuberosa* Mol.). *Agr. Hum. Values*, 33: 673-688.

Brusa, A. and C. Holzapfel. 2018. Population structure of *Ailanthus altissima* (Simaroubaceae): The role of land-use history and management. *J. Torrey Bot. Soc.*, 145(1): 55-68.

Brütting, C., S. Meyer, P. Kühne, I. Hensen and K. Wesche. 2012. Spatial genetic structure and low diversity of the rare arable plant *Bupleurum rotundifolium* L. indicate fragmentation in Central Europe. *Agr. Ecosyst. Environ.*, 161: 70-77.

Chen, B.R., H. Zou, X.H. Meng, L.X. Wang, X.L. Hao, X. Kang, C. Wang and X.X. Zhang. 2022. Prediction of distribution pattern and change of suitable areas of *Bupleurum chinense* and *Bupleurum scorzonerifolium* under climate change in China. *Acta Ecol. Sin.*, 42(20): 8471-8482.

Dai, Q. and J. Fu. 2011. When central populations exhibit more genetic diversity than peripheral populations: A simulation study. *Chin. Sci. Bull.*, 56(24): 2531-2540.

Djedid, I.K., M. Terzaghi, G. Brundu, A. Cicatelli, M. Laouar, F. Guarino and S. Castiglione. 2021. Genetic diversity and differentiation of eleven *Medicago* species from Campania region revealed by nuclear and chloroplast microsatellites markers. *Genes*, 13(1): 97.

East, E.M. 1940. The distribution of self-sterility in the flowering plants. *Proc. Am. Philos. Soc.*, 82(4): 449-518.

Eliades, N. and D. Eliades. 2009. Haplotype analysis: software for analysis of haplotype data. Forest Genetics and Forest Tree Breeding, Georg-August University Goettingen, Goettingen, Germany. Available at: http://www.uni-goettingen.de/en/134935.html

Excoffier, L. and H.E.L. Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.*, 10(3): 564-567.

Fu, K.T. 1981. *Bupleurum*. In: Northwest Institute of Botany, Chinese Academy of Sciences. (Eds), *Flora Tsinlingensis*. Vol. I. Spermatophyta - Pars 3. Science Press, Beijing, pp. 390-395.

Goudet, J. 2003. FSTAT (version 2.9.4), a program (for Windows 95 and above) to estimate and test population genetics parameters. Department of Ecology & Evolution, Lausanne University, Switzerland. Available at: http://www.unil.ch/izea/softwares/fstat.html

Hameed, A., R. T. Khan, S. R. Abbas and S. Begum. 2023. Estimation of genetic diversity among *Linum usitatissimum* L. germplasm through molecular markers. *Pak. J. Bot.*, 55(1): 63-71.

Hardy, O.J. and X. Vekemans. 2002. SPAGeDi: A versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol. Ecol. Notes*, 2(4): 618-620.

Herden, T., M. Bönisch and N. Friesen. 2020. Genetic diversity of *Helosciadium repens* (Jacq.) W.D.J. Koch (Apiaceae) in Germany, a crop wild relative of celery. *Ecol. Evol.*, 10(2): 875-890.

Hu, X.S. and R.A. Ennos. 1997. On estimation of the ratio of pollen to seed flow among plant populations. *Heredity*, 79: 541-552.

Huang, R., Y. Liu, J. Chen, Z.Y. Lu, J.J. Wang, W. He, Z. Chao and E.W. Tian. 2022. Limited genetic diversity and high differentiation in *Angelica dahurica* resulted from domestication: Insights to breeding and conservation. *BMC Plant Biol.*, 22: 141.

Hulce, D., X. Li, T. Snyder-Leiby and J.C. Liu. 2011. GeneMarker® genotyping software: Tools to increase the statistical power of DNA fragment analysis. *JBT*, 22: S35-S36.

Islam, M.S., C. Lian, N. Kameyama and T. Hogetsu. 2014. Low genetic diversity and limited gene flow in a dominant mangrove tree species (*Rhizophora stylosa*) at its northern biogeographical limit across the chain of three Sakishima Islands of the Japanese archipelago as revealed by chloroplast and nuclear SSR analysis. *Plant Syst. Evol.*, 300: 1123-1136.

Jiang, M., S. Yan, W.C. Ren, N.N. Xing, H.Y. Li, M.Q. Zhang, M.Q. Liu, X.B. Liu and W. Ma. 2022. Genetic diversity and population structure of traditional Chinese herb *Radix bupleuri* resources using genome-wide SNPs through genotyping-by-sequencing. *Russ. J. Genet+.*, 58: 1485-1492.

Koul, P., A.K. Koul and I.A. Hamal. 1989. Reproductive biology of wild and cultivated carrot (*Daucus carota* L.). *New Phytol.*, 112(3): 437-443.

Kumar, S., G. Stecher, M. Li, C. Knyaz and K. Tamura. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.*, 35(6): 1547-1549.

Li, W.W., S.L. Liu, S.M. Wang, Y.H. Li, D.R. Kong and A.L. Wang. 2024. A single origin and high genetic diversity of cultivated medicinal herb *Glehnia littoralis* subsp. *littoralis* (Apiaceae) deciphered by SSR marker and phenotypic analysis. *PLoS ONE*, 19(8): e0308369.

Li, Y.L. and J.X. Liu. 2018. StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. *Mol. Ecol. Resour.*, 18(1): 176-177.

Liu, H.D., M.J. Li, T.Q. Li, L.Y. Wu and H. Zheng. 2024. Impacts of intensified human activity on vegetation dynamics in the Qinba Mountains, China. *Forests*, 15(9): 1561-1585.

Luikart, G. and J.M. Cornuet. 1998. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. *Conserv. Biol.*, 12(1): 228-237.

Ma, C.D., Y.C. Yang, H. Chang, E.H. Wang, S.L. Hui, Y. Sun, C.L. Wang and F. Liu. 2020. Textual research and thinking of industry development on Shaanxi Buleuri Radix. *Mod. Chin. Med.*, 22(4): 641-646.

Moscoe, L.J. and E. Emshwiller. 2016. Farmer perspectives on Oca (Oxalis tuberosa; Oxalidaceae) diversity conservation: values and threats. *J. Ethnobiol.*, 36(2): 235-256.

Parra-Rondinel, F., A. Casas, D. Begazo, A. Paco, E. Márquez, A. Cruz, J. Segovia, I. Torres-García, M. Zarazúa, L. Lizárraga and J. Torres-Guevara. 2021. Natural and cultural processes influencing gene flow among wild (atoq papa), weedy (araq papa and k'ipa papa), and crop potatoes in the Andean Region of southern Peru. *Front. Ecol. Evol.*, 9: 617969.

Peakall, R. and P.E. Smouse. 2012. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics*, 28(19): 2537-2539.

Peng, J., J. Xie, Y. Gu, H. Guo, S. Zhang, X. Huang, X. Luo, J. Qian, M. Liu, X. Wan and L. Chen. 2024. Assessing population genetic structure and diversity and their driving factors in *Phoebe zhennan* populations. *BMC Plant Biol.*, 24: 1091-1109.

Phumichai, C., T. Phumichai and A. Wongkaew. 2015. Novel chloroplast microsatellite (cpSSR) markers for genetic diversity assessment of cultivated and wild Hevea rubber. *Plant Mol. Biol. Rep.*, 33: 1486-1498.

Piry, S., G. Luikart and J. Cornuet. 1999. BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *J. Hered.*, 90(4): 502-503.

Pritchard, J.K., M. Stephens and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155(2): 945-959.

Powell, W., M. Morgante, C. Andre, M. Hanafey, J. Vogel, S. Tingey and A. Rafalski. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol. Breed.*, 2: 225-238.

Salehi Shanjani, P., G.G. Vendramin and M. Calagari. 2010. Genetic diversity and differentiation of *Fagus orientalis* Lipsky in Hyrcanian forests revealed by nuclear and chloroplast microsatellite markers. *Conserv. Genet.*, 11: 2321-2331.

Smouse, P.E. and J.C. Long. 1992. Matrix correlation analysis in anthropology and genetics. *Am. J. Phys. Anthropol.*, 35(S15): 187-213.

Sheh, M.L. and F.W. Mark. 2005. *Bupleurum*. In: (Eds.): Wu, Z.Y., P.H. Raven and D.Y. Hong. *Flora of China*, Vol. 14. Missouri Botanical Garden Press, St. Louis, pp. 73.

Song, Y., X.R. Zhang, Z. Li, Z. Sun, A.X. Li, X.R. Du and Y.G. Qiao. 2022. Genetic diversity and genetic structure analysis of *Bupleurum chinense* DC. based on SSR molecular marker. *Acta Pharm. Sin.*, 57(4): 1193-1202.

Su, R., Y.Z. Su, S. Wang, J. Fan, Q.Q. Liu and M.F. Liu. 2024. A real-world study of the differences in Traditional Chinese Medicine diagnosis and treatment rules for coronavirus disease 2019 between Northern and Southern China. *J. Trad. Chin.*, 44(4): 822-829.

Wang, P., M. Zhang, J.J. Liu, L.R. Xu and W. Liu. 2015. Genetic diversity and structure of *Libanotis buchtormensis* (Fisch.) DC. in disjunct populations along the bilateral sides of deserts in northwestern China. *Plant Syst. Evol.*, 301: 2219-2230.

Wang, P., J.Q. Bai, X. Li, T.T. Liu, Y.M. Yan, Y.C. Yang and H.Z. Li. 2023. Phylogenetic relationship and comparative analysis of the main *Bupleuri Radix* species in China. *PeerJ*, 11: e15157.

Wang, X., X.Q. Liu, Y.Z. Ko, X.L. Jin, J.H. Sun, Z.Y. Zhao, Q.J. Yuan, Y.C. Chiang and L.Q. Huang. 2020. Genetic diversity and phylogeography of the important medical herb, cultivated Huang-Lian populations, and the wild relatives *Coptis* species in China. *Front. Genet.*, 11: 708.

Wang, Y.C., L. Liao and Z.Z. Li. 2018. Genetic differentiation of *Actinidia chinensis* and analysis of gene flow barriers in the Qinling Mountains, the species' northern distribution boundary. *Genet. Resour. Crop Evol.*, 65: 881-895.

Warschefsky, E., R.V. Penmetsa, D.R. Cook and E.J.B. Von Wettberg. 2014. Back to the wilds: Tapping evolutionary adaptations for resilient crops through systematic hybridization with crop wild relatives. *Amer. J. Bot.*, 101(10): 1791-1800.

Wei, X., V.L. Sork, H. Meng and M. Jiang. 2016. Genetic evidence for central-marginal hypothesis in a Cenozoic relict tree species across its distribution in China. *J. Biogeogr.*, 43(11): 2173-2185.

Wills, D.M. 2006. Chloroplast DNA variation confirms a single origin of domesticated sunflower (*Helianthus annuus* L.). *J. Hered.*, 97(4): 403-408.

Wright, S. 1978. *Evolution and the Genetics of Populations* (Volume 4). University of Chicago Press, Chicago.

Wu, Q., S. Dong, Y.X. Zhao, L. Yang, X.J. Qi, Z.H. Ren and S.B. Dong. 2023. Genetic diversity, population genetic structure and gene flow in the rare and endangered wild plant *Cypripedium macranthos* revealed by genotyping-by-sequencing. *BMC Plant Biol.*, 23: 254.

Yang, J., J. Luo, Q. Gan, L. Ke, F. Zhang, H. Guo, F. Zhao and Y. Wang. 2021. An ethnobotanical study of forage plants in Zhuxi County in the Qinba mountainous area of central China. *Plant Divers.*, 43(3): 239-247.

Yuan, N., Y. Sun, H.P. Comes, C.X. Fu and Y.X. Qiu. 2014. Understanding population structure and historical demography in a conservation context: Population genetics of the endangered *Kirengeshoma palmata* (Hydrangeaceae). *Amer. J. Bot.*, 101(3): 521-529.

Zhang, G.X., H. Wang, L.C. Shi, Y. Liu, R.Y. Yao, C. Sui, C.M. Yang, H.L. Ji, Q.L. Wang, and J.H. Wei. 2022. Identification of the original plants of cultivated *Bupleuri Radix* based on DNA barcoding and chloroplast genome analysis. *Peer J.*, 10: e13208.

Zhang, Q.P., D.Y. Zhang, K. Yu, J.J. Ji, N. Liu, Y.P. Zhang, M. Xu, Y.J. Zhang, X.X. Ma, S. Liu, W.H. Sun, X. Yu, W.Q. Hu, S.R. Lan and Z.J. Liu. 2021a. Frequent germplasm exchanges drive the high genetic diversity of Chinese-cultivated common apricot germplasm. *Horti Res.*, 8: 215.

Zhang, Z., Q. Yang, Y.N. Niu, Y.X. Zhang, S. Dong, W.W. Zhang and Z.Z. Wang. 2021b. Diversity analysis and establishment of core collection among *Akebia trifoliata* (Thunb.) Koidz. in Qinba mountain area of China using ISSR and SRAP markers. *Genet. Resour. Crop Ev.*, 68: 1085-1102.

Zhu, Y., W.N. An, J. Peng, J.W. Li, Y.J. Gu, B. Jiang, L.H. Chen, P. Zhu and H.B. Yang. 2022. Genetic diversity of *Phoebe zhennan* (S. Lee. et F. N. Wei) breeding population and extraction of core collection using nSSR, cpSSR and phenotypic markers. *Forests*, 13(8): 1320.