Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The impact of Quaternary climate change on the historical population dynamics of Psammochloa villosa, a typical desert herb from northwest China

Jia Rui Jin
Qinghai Normal University
Yu Ping Liu
Qinghai Normal University
Xu Su
xusu8527972@126.com
Qinghai Normal University
Ting Lv
Qinghai University
Ying Hui Zheng
Qinghai Normal University
Qian Yang
Qinghai Normal University
Yang Lv
Qinghai Normal University
Xu Feng
Qinghai Normal University
Kai Yue Wei
Qinghai Normal University

Research Article

Keywords: Poaceae, Psammochloa villosa, genetic diversity, genetic differentiation, historical demography, gene flow

Posted Date: July 11th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-6516707/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Climatic oscillations and geographic barriers since the Quaternary have profoundly influenced plant distribution patterns, driving substantial species differentiation through range fragmentation. Investigating the interplay between genetic structure divergence and population dynamics under climatic forcing remains a central challenge in evolutionary biology. *Psammochloa villosa* (Poaceae), a dominant perennial grass endemic to the Inner Mongolian Plateau's desert steppe, serves critical functions in grassland restoration and livestock forage provision. Notably, its wind-pollination strategy and distant hybridization capacity establish this species as a model system for studying panmictic population equilibrium. This study employs a dual-marker approach, combining 10 low-copy nuclear gene loci (LCNG) with 13 SSR markers, to elucidate the population genetic architecture across 43 natural populations (n = 210 individuals) spanning the distribution range of *P. villosa*. The results showed that the populations of *P. villosa* were distinctly divided into two major branches, Group 1 and Group 2, separated by the Yinshan Mountains, with Group 2 exhibiting higher genetic diversity (H_d = 0.990, π = 0.00145). The populations showed significant diversity and phylogeographic structure, with the majority of genetic variation originating from differences between populations (F_{ST}(LCNG) = 0.766; F_{ST}(SSR) = 0.577). The Mantel test revealed a positive correlation between genetic distance and geographic distance. Subsequent correlation analysis of genetic differentiation (measured by F_{ST}) with climatic factors demonstrated an overall positive association, although only elevation showed a statistically significant correlation with F_{ST} values. Demographic history analyses revealed that both Group 1 and Group 2 exhibited smaller effective population sizes compared with N_A. Phylogenetic divergence analysis estimated that *P. villosa* and *Achnatherum splendens* diverged approximately 3.26 Ma, whereas the differentiation between Group 1 and Group 2 initiated around 0.38 Ma. Notably, bidirectional yet asymmetric gene flow was detected between the two groups. Geographical barrier analysis identified significant genetic discontinuities corresponding to major mountain ranges, including the Yinshan, Helan, Ordos Plateau, and Yabulai mountains. These findings collectively suggest that Quaternary uplift of the Tibetan Plateau, coupled with the divergent monsoon influences east and west of the Yinshan Mountains and progressive habitat fragmentation, have driven the observed genetic differentiation in *P. villosa* populations through environmental adaptation.

1 Introduction

Plant evolution is intricately linked to climatic and environmental dynamics, where these transformative forces not only drive species differentiation but also catalyze significant shifts in biogeographical distribution patterns[1]. The “Quaternary glaciation,” which began approximately 2.58 Ma, is a critical phase in the Earth’s environmental evolution, marked by multiple expansions and contractions of glaciers. The climate experienced dramatic fluctuations and frequent alternations between cold and warm periods. During this time, some plants were forced to adapt and evolve, driving changes in plant geographical distribution patterns and the evolution of ecosystems[2]. Over the past 30 years, many scholars have conducted biogeographical studies on various plant species worldwide, hypothesizing about the refugia of the Quaternary glaciation and the post-glacial expansion routes[3–4]. The molecular clock hypothesis suggests that the degree of difference between homologous genes of any two species should be approximately constant with their divergence time. In other words, the longer the time since the ancestral species began to diverge, the greater the molecular information difference between species[5]. Therefore, researchers often use population genetics methods, employing statistical tools to study changes in gene frequencies, gene flow, genetic drift, and other data within populations. Combining these tools with biogeographical methods, they explore how species respond to environmental changes, how they evolve through gene flow and isolation, and how their distribution patterns and historical evolution are shaped by geographical space[6–8].

In recent years, with the rapid development of molecular biology, DNA molecular markers have been extensively applied in biogeographical studies. Compared with morphological, cellular, and biochemical markers, molecular markers have the advantages of being numerous, highly polymorphic, predominantly "neutral" and co-dominant, and often expressed in DNA form, making them a relatively ideal new genetic marker. In eukaryotic research, plastid genes such as those from mitochondria and chloroplasts are widely used due to their easy acquisition, structural similarity, and sequence conservatism[9]. Ma et al.[10] studied the phylogenetic relationships of *Xanthopappus subacaulis* using the complete chloroplast genome. They found that *X. subacaulis* forms an independent and well-supported clade within the Asteraceae family. Chen et al.[11] studied the mitochondrial genome characteristics of *Elymus magellanicus* and reconstructed the phylogenetic relationships of Poaceae. The results indicated that *E. magellanicus* is closely related to *Hordeum vulgare*, *Aegilops speltoides*, *Triticum aestivum*, and *Triticum aestivum*. Park et al.[12] elucidated key features of the organelle structure, gene content, and DNA transfer dynamics based on the high-quality chloroplast and mitochondrial genome assemblies and the draft nuclear genome of *Cicuta virosa*. Their findings on gene loss, DNA transfer to the nucleus, and structural diversity among related species provide insights into the evolution of organelle genomes in the Apiaceae family.

However, chloroplast-based molecular markers also have limitations. Chloroplast genes are inherited in a strictly maternal manner, carry limited evolutionary information, and are difficult to use for revealing deep phylogenetic relationships within large population groups[13]. Compared with organelle genes, nuclear genes have more complex evolutionary patterns, including gene duplication and gene loss. Currently, commonly used DNA molecular markers include restriction fragment length polymorphisms (RFLP), random amplified polymorphic DNA (RAPD), amplified fragment length polymorphisms (AFLP), simple sequence repeats (SSR), and single-nucleotide polymorphisms (SNP), among others[14]. Low-copy nuclear genes (LCNG) with fewer copies in a given organism's genome, inherited from both parents and with abundant information sites, are less affected by recombination and are capable of preserving parental alleles over time. In addition, the evolutionary rate in intron regions is relatively fast, making them widely used to explore the phylogenetic relationships of plant hybrids and allopolyploid groups, population geographic distribution patterns, origins and evolution, and population historical dynamics[15–16]. SSR molecular markers are DNA sequences consisting of repeat units of approximately one to six nucleotides, offering advantages such as rich polymorphism, good repeatability, strong stability, co-dominance, simplicity in operation, and broad distribution[17]. They are among the most widely used molecular markers. Currently, SSR molecular markers are widely applied in plant genetic diversity detection, kinship analysis, and gene mapping, among other areas. A study on the genetic diversity and population history dynamics of the species *Camellia azalea* based on SSR molecular markers found that *C. azalea* has moderate to high levels of genetic diversity. However, its current ecological niche is very narrow, and it is predominantly distributed along stream banks.

Recent human activities have altered the gene flow of this species, which may have contributed to the strengthening of genetic structure in new individuals[18]. Using population transcriptomes, complete chloroplast genomes, genotyping sequencing, and microsatellite markers, Liu et al.[19] found that geographic, geological, and environmental factors are key contributors to shaping the cryptic intraspecific divergence and population dynamics of the cold-tolerant perennial herb *Notopterygium oviforme*.

Psammochloa villosa is the type species of the genus *Psammochloa*, belonging to the family Poaceae and the subfamily Stipeae, $2n = 2x = 46$[20]. *P. villosa* is a typical clonal plant distributed mainly in the sandy areas of the Inner Mongolia Plateau and its neighboring regions in China, including the Badain Jaran Desert, the Ulan Buh Desert, the Mu Us Sandy Land, the Hexi Corridor in Gansu, northern and central Ningxia, and the sandy regions of northern Shaanxi. It has well-developed rhizomes and a growth cycle that lasts more than 2 years. Like other early-growth and drought-tolerant plants, *P. villosa* demonstrates remarkable drought tolerance and adaptability to aeolian stresses, particularly sand burial and wind erosion. As a pioneer psammophyte, this species exhibits characteristic early-successional traits including efficient sand stabilization through rapid vertical accretion, extensive lateral spread via rhizomatous growth, and significant potential for initiating vegetative recovery in mobile dune ecosystems[21]. Furthermore, its inflorescences are relatively large and elongated, with many spikelets and large seeds, which have significant economic value[22]. On the one hand, it improves ground vegetation cover in highly heterogeneous degraded ecosystems; on the other hand, it provides abundant forage for livestock, thus reducing grazing areas and slowing down desertification caused by human activities[23]. Furthermore, *P. villosa* exhibits wind pollination and outcrossing characteristics, making it an ideal population for random mating in an ideal state[24].

Currently, research on *P. villosa* focuses primarily on morphological traits[25], ecological adaptability[26], genetic diversity[27], and germplasm resources[22]. Through the comparative observation of the roots, stems, leaves, and flowers of *P. villosa*, it was found that its clonal characteristics enable it to have a stronger survival ability in arid and barren environments. In addition, the pore density in the leaf epidermis was found to be positively correlated with the degree of environmental drought[25]. Li et al.[28] analyzed the genetic variation of seven *P. villosa* populations using inter-simple sequence repeat (ISSR) molecular markers, concluding that the ratio of genetic variation between populations to variation within populations was greater than 1.00. Liu et al.[29] analyzed transcriptomic data under different stress conditions, indicating that the *PvTIP41* gene is the most stable internal reference gene in leaves under drought and salt stress, as well as in stems under salt stress, providing important references for future genomic research on the plant's stress resistance mechanisms. Although Lv et al.[30] used AFLP molecular markers and spatial distribution models to predict the population formation history and distribution dynamics of *P. villosa*, AFLP is uniparental in inheritance and unable to distinguish between homozygous and heterozygous states, and the use of a single type of molecular genetic marker may lead to uneven coverage of polymorphisms.

Therefore, this study focuses on *P. villosa*, a unique desert herbaceous plant from the Inner Mongolia Plateau and its neighboring areas. Based on extensive field surveys and sampling, the study employs a combination of LCNG and SSR markers to examine the genetic variation distribution patterns of *P. villosa* populations and to determine their spatial genetic structure within populations, between populations, and across the entire geographical distribution range, comprehensively elucidating its modern distribution pattern and phylogenetic evolution history. The study aims to reveal how herbaceous plants in the arid desert areas of northwest China respond to Quaternary climate changes, providing a theoretical basis for the conservation and utilization of *P. villosa* germplasm resources.

2 Materials and Methods

2.1 Plant materials

We collected experimental accessions of 101 natural populations of Inner Mongolia and removed the redundant accessions. Finally, we obtained a diverse collection of 43 populations and 210 individuals of *P. villosa*. These accessions covered the main natural geographical distribution area of *P. villosa* in China (Fig. 1). Information about the accessions, including the collected location, habitat status, longitude, latitude, altitude, and number of individuals, is provided in Additional file 1: Table S1. We placed healthy leaves in silica gel to dry rapidly and used these for total DNA extraction. Young plants were frozen at − 80°C after processing in liquid nitrogen for transcriptome sequencing.

2.2 Transcriptome sequencing and data analyses

Total RNA was extracted separately from each tissue using the TaKaRa MiniBEST Plant RNA Extraction Kit. The sequencing platform PacBio was used for full-length transcriptome sequencing of mixed samples of roots, stems, and leaves of *P. villosa*. Raw sequencing data were processed, including read filtering, error correction, and quality control. We obtained unigenes by performing clustering and redundancy removal of assembled sequences using Cdhit software. Subsequently, using the transcriptome database of *Achnatherum splendens* as a reference (GenBank: SRP068186), we retrieved the homologous genes of *P. villosa* transcripts by OrthoMCL software. When the retrieved gene showed homology with *A. splendens* but not with its own sequence, it was considered a successful selection of LCNG in this species[31]. We designed primers and carried out experimental validation for 72 genes randomly selected from the screened LCNG of this species on the basis of primer software[32]. Meanwhile, we used MISA software to search, locate, and identify simple repetitive sequences in the obtained transcriptome data. From the identified SSR loci, we designed primers and screened out 13 pairs of SSR primers with high polymorphism, good repeatability, and strong representativeness. Finally, we conducted population genetics experiments on the screened 10 pairs of LCNG and 13 pairs of SSR primers using the DNA of 43 populations and 210 individuals as templates (Tables 1 and 2).

Table 1
Characteristics of ten low-copy nuclear gene primers developed in *P. villosa*

Locus	Primer name	Primer sequences (5'-3')	LCR (bp)	T_a (°C)
14031	Locus1	F: TTCAGTGGCTATCCTGTTCC	716	57.1
		R: CTGCACGTTAAACTACTCACAAG		
14004	Locus2	F: AGGTGAGGTGGGATTTGAGC	736	61.5
		R: AAGTGGGCGGATGGTTTC		
13593	Locus3	F: CCAGCTTAGTATGGCAGAGG	772	58.8
		R: CCACGCTTTGTTGATGTTTC		
14111	Locus4	F: CCTAACGGAAACGAGTGCC	1217	61.6
		R: CCTCATCAGTTCCCGCAGAC		
14075	Locus5	F: AGTTGCCAGTGGGTTTG	996	55.7
		R: CGAGCACATGGAGGTATCTT		
13604	Locus6	F: TCGCATTCCTCTTTCTCACCAAC	735	60.7
		R: CTATTCGACTTCCTCAGCTCCCT		
14192	Locus7	F: TGTCGGTAACCTGGATAGAGC	807	59.1
		R: TGATTAACTCCCTGACTGCTTC		
14928	Locus8	F: GAGCGGTTTGTCCCTGAT	918	58.4
		R: GTCCAGCTCTAGTGCCTGTT		
14224	Locus9	F: GCACTCAGTCACGGCAAAC	884	60.9
		R: AGCAACAGCCAGCTTCTCAC		
14612	Locus10	F: GATGCAGTGGCAGTGTTTG	887	57.1
		R: CTCTTACCGTCAGGCTCGT		

Note: AFL, Aligned full length; LCR, Length of coding region; T_a, Optimal annealing temperature.

Table 2
13 pairs of SSR primer information in *P. villosa*

Locus	Primer name	Primer sequences (5'-3')	Motif	Allele size (bp)	T_m (°C)
Pvj1	Locus1	F: AAGAGCATTCCTTGCAGCA	$(TC)_{10}$	342–378	60.8
		R: ATGACTCCGGCATGTGGTT			
PvH4	Locus2	F: ATCCTTTGAGCATCCACCCCC	$(TCC)_5$	202–341	61.0
		R: GCTTGCTGAGGAGTCTCTGCGTCT			
Pv14	Locus3	F: GATCCTGGGTTTGGAAGTG	$(AGA)_6$	247–283	54.6
		R: CTTATAGTACTCTTCAAGCTCCTT			
Pv15	Locus4	F: AGCAGATTGGCAGAGAACTCAGG	$(GTG)_7$	309–342	63.8
		R: CTGCGCCTTCCTCCATCAT			
Pv21	Locus5	F: CCTCTACCGCCTCTTCGTCTCC	$(CGC)_6$	266–320	63.7
		R: GCGAACCAGAAGAGGTAGTACTC			
Pv33	Locus6	F: GCACCACGCCGGCAAGAAA	$(AGG)_6$	249–273	61.1
		R: CCGTGGACGACTTGCCGAAGA			
Pv39	Locus7	F: TCTGAGGTACAATATCCGAAGC	$(AGC)_7$	306–324	59.4
		R: ACTTGCGATTGAAGCGTGT			
Pv44	Locus8	F: AAGTAAGCCCAGAGCAGGAG	$(GGA)_6$	230–239	57.4
		R: AATCAGCCAATCATCAATCTCA			
Pv49	Locus9	F: GCCTTCGGGTGCGTGAGC	$(GCG)_5$	293–311	64.4
		R: CCTCGTCTTCCCCACGGTGA			
Pv50	Locus10	F: CGGCTGGACGAGGCGAAGGT	$(GCG)_6$	284–296	64.0
		R: CAGGAGGAGGGTGTTGGACGTGAC			
Pv51	Locus11	F: CGCCGCGTTCCCGCTGTA	$(CGG)_6$	216–258	64.0
		R: CCGTCCCCGCATTGCTGC			
Pv8	Locus12	F: TGGCAGTGGATGTGGAGCTGAT	$(AGC)_6$	306–345	65.4
		R: CACGGATGGATGGCACTGCTAC			
Pv23	Locus13	F: GACTAACCTCCTGGGCACT	$(CGG)_6$	234–252	63.7
		R: TGCCGCTGCATATGCCGAAG			

2.3 Population genetic structure

To infer the population genetic structure of *P. villosa*, this study employed three methods to analyze the geographical distribution patterns of SSR and low-copy nuclear gene data. (1) After converting the "01" matrix into a genetic distance matrix using GenAlEx v.6.5 software, we constructed UPGMA clustering and principal coordinate analysis (PCA) according to R scripts. (2) We analyzed the spatial analysis of molecular variation among different geographical groups of *P. villosa* by SAMOVA v.1.0 software[33] and deduced the optimal population combination from geographical and genetic information. The process set a range of K values from 2 to 10, with each K value undergoing 100 heuristic searches. The optimal population combination was determined by selecting the K value that maximized genetic differentiation between groups (F_{CT}), minimized genetic differentiation within groups (F_{SC}), or achieved stable F_{CT}. (3) We tested the different genetic clusters among species using Structure v.2.3 software[34]. The range of genetic clusters was set to $1 \le K \le 10$. The likelihood values for each cluster (K) were estimated during the operation, and it was assumed that there was correlation in allele frequencies between different clusters. Subsequently, the most likely number of genetic clusters (K) for the population was estimated according to the statistical method of Pritchard et al.[35] and the calculation method of ΔK[36].

2.4 Analysis of genetic diversity and genetic variation

Diversity parameters are important for measuring the level of population genetic diversity[37]. To estimate the genetic diversity level of *P. villosa*, this study calculated various genetic diversity parameters for SSR and low-copy nuclear gene data with GenAlEx v.6.5[38] and DnaSP v.5.1 software, including the number of individuals (N), percentage of polymorphic loci (PPL), polymorphic information content (PIC), observed number of alleles (N_a), effective number of alleles (N_e), Shannon's information index (I), observed heterozygosity (H_o), expected heterozygosity (H_e), haplotype diversity (H_d), nucleotide diversity (π),

average number of nucleotide differences (K), Nei's genetic diversity (h), total population diversity (H_t), genetic diversity within populations (H_s), inbreeding coefficient (F_{IS}), population differentiation (F_{ST}), genetic differentiation between populations (G_{ST}), and gene flow (N_m).

The classic measure of genetic variation is Wright's fixation index (F-statistic, F_{ST}), a measure of detecting the degree of genetic differences at the level of species populations[39]. The theory suggests the following guidelines for the interpretation of F_{ST} values: 0 ~ 0.05, little genetic variation; 0.05 ~ 0.15, moderate genetic differentiation; 0.15 ~ 0.25, great genetic differentiation; and above 0.25, very great genetic differentiation[40]. This study estimated the F_{ST} value using the molecular variance method in Arlequin v.3.5 software[41]. Two different grouping populations of *P. villosa* were subjected to analysis of molecular variance (AMOVA) with SSR and LCNG data. All populations were considered as one group to calculate the genetic variation between populations and within populations, and to estimate the levels of genetic variation between populations, within populations, among groups, and between groups based on the spatial AMOVA (SAMOVA) grouping results. Significance levels were tested through 1000 nonparametric permutations. In addition, we estimated the genetic differentiation coefficients G_{ST} and N_{ST} of populations in the SAMOVA grouping and the overall level of low-copy nuclear gene data by the PERMUT software. When $N_{ST} > G_{ST}$ ($P < 0.05$), it indicated the presence of significant phylogeographic structure among populations.

2.5 Correlation analysis, mismatch distribution, and neutrality tests

To assess whether the population's genetic structure was affected by spatial distance or other environmental factors, this study conducted a Mantel test using R scripts that examined the correlation between the genetic distance and the geographic distance. Moreover, we investigated the correlation between genetic diversity, genetic differentiation, and several ecological factors.

We conducted mismatch distribution analysis on the population of *P. villosa*, aiming to explore the historical dynamics of this species using Arlequin v.3.5 software. Mismatch distribution analysis is a method used to assess the fit between the expected value curve and the observed value curve based on the distribution of nucleotide differences among different haplotypes[42]. If the mismatch distribution curve is unimodal, it indicates that the species has recently experienced population expansion[43–44], whereas if the curve is multimodal, it suggests that the species has been in a phase of dynamic equilibrium or slow contraction for population size during a certain time period[43].

In addition, to determine whether the genetic locus conforms to the neutral evolution model. This study performed neutrality tests on the *P. villosa* population using DnaSP v.5.1 software. This theory assumes that the population maintains a mutation–drift equilibrium during the long-term evolution, and all nucleotides mutated with equal probability. A Tajima's D or Fu's Fs significantly greater than 0 indicates that the population has recently experienced a bottleneck effect or balancing selection. Conversely, a Tajima's D or Fu's Fs significantly less than 0 indicates that the population has undergone directional selection or an expansion event[45].

2.6 Demographic history estimation

To further investigate the historical dynamics of *P. villosa*, we simulated the population historical dynamics of 10 pairs of LCNG and 13 pairs of SSR primers based on an approximate Bayesian computation algorithm. This analysis aimed to determine the population divergence events, time, and changes in population size[46]. According to the population genetic structure of *P. villosa*, we inferred the migration process of this species. Each scenario was executed 10^6 times, and the best scenario model depends on the posterior distribution probabilities. Moreover, we inferred the goodness of fit for the best scenario through the difference between the observed value and the actual value of principal component analysis (PCA) in DIYABC v.2.0.4 software. The prior distribution parameters for the population history dynamics are presented in Table S2.

Furthermore, we simulated the effective population size, effective migration rate, and divergence time of the current and ancestral population for *P. villosa* using the isolation-with-migration model in the IMa software. Convergence of the runs was assessed by the effective sample sizes (ESS) (≥ 200) using Tracer v.1.7[47] and the peak graph conforming to a normal distribution. As there was a lack of fossil records of *Stipeae* and *Psammochloa*, estimates of the timescale of divergence were based on an evolutionary rate of $\mu = 6.1 \times 10^{-9}$ substitutions per site per year in Poaceae plants[48], and the generation time was set at 5 years[49].

2.7 Estimating divergence time

We selected the best nucleotide substitution model by jModelTest software[50], which included 11 fundamental nucleotide substitution models. Then, we estimated the divergence time of the population for *P. villosa* through BEAST software[51]. Due to the lack of fossil records or the rate of gene fragment evolution of *Psammochloa*, we performed the molecular clock calibration using the divergence time between *Aegilops tauschii* and *Triticum dicoccoides* (2.97 Ma, 2.61 ~ 5.27 Ma)[52], as well as the divergence time between *Aegilops tauschii* and *Brachypodium distachyon* (32.6 Ma, 27.8 ~ 37.0 Ma)[53]. The running parameters were set as follows. A relaxed molecular clock model was selected, and the Markov Chain Monte Carlo (MCMC) was run for 10^8 generations, with sampling every 1000 generations. The burn-in was set to 10%. Subsequently, we checked the posterior distribution probability and ensured that the ESS value was greater than 200 by Tracer v.1.6 software. If the ESS value was lower than 200, it was necessary to increase the number of iterations to continue the operation. Meanwhile, we displayed and edited the divergence time tree using Figtree v.1.4.3 software.

2.8 Population genetic barriers and gene flow

Population genetic barriers are caused by external environmental factors such as gene flow, genetic drift, and natural selection. Therefore, to infer the impact of population evolutionary history on population genetic structure, we detected genetic barriers among the population of *P. villosa* with the "Monmonier's maximum difference algorithm" in Barrier v.2.2 software[54].

When individuals migrate from one population to another, they bring their genes to the new population, resulting in gene flow, which can lead to population homogenization. Consequently, we estimated genetic parameters for SSR data and low-copy nuclear gene data based on Migrate-n v.4.4.3 software[55], including migration rates within and between all populations (*M*), effective population sizes (θ), and historical gene flow. We set the running parameters as follows: long-chains = 3, long-inc = 100, long-sample = 5000000, burn-in = 1000000, replicate = YES:5, heating = YES:1: {1.000000,1.500000,3.000000,1000000.000000}.

3 Results

3.1 Population genetic structure

Achnatherum splendens was used as an outgroup. A dendrogram of *P. villosa* was constructed on the basis of low-copy nuclear gene data depicting the genetic relationships among the 41 populations (Fig. 2a). The populations were divided into two main clusters, with a self-proclaimed differentiation rate of 56%. Group 1 included populations from the northeast of the Yinshan Mountains (P1-P12). Group 2 included populations from the Helan Mountains (P13-14、P16-19 and P24-28), the northeastern margin of the Qinghai-Tibet Plateau (P21-23), the Southern Yinshan mountain range (P29-43). Furthermore, the cluster results based on SSR data showed distinct genetic differentiation among populations from the northeast and southwest of the Yinshan Mountains (Fig. 2b). Because of the low self-proclaimed differentiation rate, the genetic structure for *P. villosa* was assessed using SAMOVA based on low-copy nuclear gene data.

The grouping is optimal when the value of F_{CT} reaches the maximum. The results indicated that this value was the largest (0.50155) with $K = 10$ (Table 3). However, there was a single population forming a group when *K* was greater than 5, resulting in the inability to detect an effective phylogeographic group. Meanwhile, due to the geographical proximity and similar habitat of P1 and P3, the best grouping was $K = 2$, for which the 41 populations were divided into two phylogeographic groups (Table 3). This result was completely consistent with the phylogenetic tree.

Table 3
The optimal numbers of population groups (*K*) inferred by SAMOVA algorithm based on low-copy nuclear gene datasets

K	Population grouping	F_{CT}	F_{ST}	F_{SC}	*P*
$K = 2$	(1, 2, 3, 4, 6, 7, 8, 9, 11, 12, 13, 14) (15, 16, 18, 19, 20, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45)	0.407	0.822	0.700	0.000
$K = 3$	(1, 4, 6, 7, 8, 11, 12, 13, 14) (2, 3, 9) (15, 16, 18, 19, 20, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45)	0.435	0.822	0.685	0.000
$K = 4$	(1, 4, 13) (2, 3, 9) (6, 7, 8, 11, 12, 14) (15, 16, 18, 19, 20, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45)	0.460	0.823	0.671	0.000
$K = 5$	(1, 4, 13) (2, 3, 9) (6, 7, 8, 11, 12) (14) (15, 16, 18, 19, 20, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45)	0.460	0.822	0.670	0.000
$K = 6$	(1, 4, 13) (2, 3, 9) (6, 7, 8, 11, 12) (14) (15, 16, 18, 19, 20, 21, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45) (29)	0.468	0.820	0.661	0.000
$K = 7$	(1, 4, 13) (2, 3, 9) (6, 7, 8, 11, 12) (14) (15, 16, 18, 20, 21, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45) (19) (29)	0.469	0.816	0.653	0.000
$K = 8$	(1, 4, 13) (2, 3, 9) (6, 7, 8, 11, 12, 14) (15, 16, 18, 28) (19, 20, 21, 23, 24, 25, 26, 27, 30, 31, 32, 34, 37, 38, 41, 44, 45) (29) (33, 36, 39, 40, 42, 43) (35)	0.483	0.794	0.602	0.000
$K = 9$	(1, 4, 13) (2, 3, 9) (6, 7) (8, 11, 12) (14) (15, 16, 18, 28) (19, 20, 21, 23, 24, 25, 26, 27, 30, 31, 32, 34, 35, 37, 38, 41, 44, 45) (29) (33, 36, 39, 40, 42, 43)	0.492	0.794	0.595	0.000
$K = 10$	(1, 4, 13) (2, 3, 9) (6, 7) (8, 11, 12) (14) (15, 16, 18, 28) (19, 20, 21, 23, 24, 25, 26, 27, 30, 31, 32, 34, 35, 37, 38, 41, 44, 45) (29) (33, 36, 39) (40, 42, 43)	0.502	0.793	0.585	0.000

The Bayesian clustering method of the STRUCTURE software was further applied to analyze the genetic structure of *P. villosa.* The most optimal number of population groups (*K*) was determined with the posterior probability LnP(D) and delta *K* (ΔK), setting the population size *K* as 2–15 and 2–10. The results indicated that ΔK was maximum when $K = 2$ (Fig. 3a,b,d,e), indicating that populations of this species could be divided into two subgroups. At the same time, there was gene flow between the two genetic units, according to the grouping results of the STRUCTURE program, where red gene pools represent the Group 1 population and blue gene pools represent the Group 2 population (Fig. 3c, f). The results of principal component analysis (PCoA) further confirmed the clustering results of STRUCTURE (Fig. 4). Overall, the population genetic structures of *P. villosa* obtained using different methods were almost identical.

3.2 Genetic diversity

Haplotype diversity index is an important indicator for measuring the degree of variation within a population. These spliced sequences of 10 primer pairs detected a total of 52 polymorphic sites in 174 samples, with the number of polymorphic loci (*S*) ranging from 2 to 11, and the number of haplotypes (*H*) ranging from 4 to 23 (Table 4). The haplotype diversity (H_d) ranged from 0.479 to 0.854, with an average of 0.658, in which the haplotype diversity was the

highest at site 14111 and the lowest at site 14192. These results showed that differences in gene expression contributed to the variation observed in *P. villosa*. Meanwhile, the nucleotide diversity (π) ranged from 0.00062 to 0.00414, with an average of 0.00180, indicating that 10 primer pairs exhibited high polymorphism. Based on the genetic diversity analysis of low-copy nuclear genes in the *P. villosa* populations, the results showed that H_d, π, and K respectively ranged from 0.000 to 1.000, 0.000 to 0.203, and 0.000 to 10.578 (Table S3). This indicated significant genetic diversity differences among *P. villosa* populations, in which populations located in the central Inner Mongolia Plateau exhibited higher genetic diversity. Simultaneously, this study also analyzed the genetic diversity of *P. villosa* populations based on the SAMOVA grouping results. Group 2 demonstrated a high level of genetic diversity, with H_d and π measuring 0.990 and 0.00145, respectively (Table 5).

Table 4
Statistics of genetic diversity for 10 pairs of low-copy nuclear genes of *P. villosa* in the present study

Locus	Number of polymorphic sites (S)	Number of Haplotypes (H)	Haplotype diversity (*H*d)	Nucleotide diversity (π)	Average number of nucleotide differences (*K*)	G + C	Rm
14031	3	4	0.526	0.00083	0.591	0.476	0
14004	3	4	0.569	0.00093	0.684	0.443	0
13593	5	6	0.713	0.00188	1.448	0.462	0
14111	11	23	0.854	0.00355	5.147	0.649	6
14075	5	6	0.512	0.00062	0.617	0.422	0
13604	2	4	0.737	0.00135	0.996	0.589	1
14192	3	4	0.479	0.00111	0.895	0.464	1
14928	4	8	0.533	0.00093	0.851	0.434	1
14224	9	13	0.838	0.00414	3.660	0.493	5
14612	7	10	0.818	0.00262	2.321	0.474	1
Mean	-	-	0.658	0.00180	1.721	0.491	-

Table 5
Genetic diversity for 41 populations of *P. villosa* based on low-copy nuclear genes in the present study

The parameter of genetic diversity	Populations of Group 1	Populations of Group 2	All populations
Number of polymorphic sites	31	43	52
Number of Haplotypes	47	110	157
Haplotype diversity	0.975	0.990	0.993
Nucleotide diversity	0.00141	0.00145	0.00189
Average number of nucleotide differences	12.251	12.528	16.378
G + C	0.495	0.496	0.496
Rm	13	18	26

An analysis was conducted on the polymorphic loci of 13 pairs of SSR primers using POPGEN 1.32 software, and the results are presented in Table 6. A total of 104 alleles were detected at the 13 microsatellite loci in the 195 samples of *P. villosa*, with the number of alleles per locus (N_a) ranging from 3 to 14 (average of 8). The primer with the highest number of alleles was Pv21, and the primer with the lowest number of alleles was Pv44. The expected heterozygosity (H_e) and observed heterozygosity (H_o) of the microsatellite loci ranged from 0.183 to 0.576 and 0.215 to 0.964, respectively. The polymorphic information content (PIC) generated ranged from 0.317 to 0.747, with a mean value of 0.511. the Shannon information index (*I*) ranged from 0.278 to 0.946, with a mean value of 0.567. These results indicate that the 13 pairs of SSR primer marker loci are suitable for genetic research on *P. villosa* owing to their high polymorphism and genetic diversity. At the population level, the mean values of H_e and H_o were 0.359 and 0.509, respectively (Table S3). The heterozygosity of P30 and P31 was greater than 0.5, indicating that these two populations have high genetic diversity and relatively low human selection factors. In addition, the fixation index (*F*) can reflect the extent of deviation of a species. If $F > 0$, it indicates a higher frequency of homozygotes within the population, whereas if $F < 0$, it suggests a higher frequency of heterozygotes within the population. In view of the above, P16 and P18 have positive values among the 39 populations, whereas the *F* values for the other populations are less than 0. This indicates a higher frequency of heterozygotes within the *P. villosa* population (Table S3).

Table 6
Summaries of genetic variation statistics for all SSR loci

Primer EcoR I/Mse I	N_a	N_e	F_{IS}	F_{ST}	I	H_o	H_e	N_m	PIC
j-F/j-R	10	1.692	-0.329	0.425	0.502	0.431	0.324	0.339	0.523
H4-F/H4-R	11	2.546	-0.674	0.264	0.946	0.964	0.576	0.697	0.747
44-F/44-R	3	1.362	-0.359	0.631	0.287	0.256	0.189	0.146	0.393
39-F/39-R	8	1.941	-0.497	0.484	0.622	0.595	0.397	0.266	0.737
33-F/33-R	6	1.484	-0.255	0.489	0.419	0.338	0.270	0.261	0.485
23-F/23-R	4	1.919	-0.849	0.187	0.662	0.862	0.467	1.090	0.482
15-F/15-R	9	1.509	-0.533	0.365	0.417	0.410	0.268	0.434	0.406
8-F/8-R	5	1.334	-0.176	0.721	0.278	0.215	0.183	0.097	0.594
14-F/14-R	8	2.245	-0.736	0.203	0.834	0.938	0.541	0.980	0.631
50-F/50-R	7	2.066	-0.053	0.113	0.779	0.518	0.492	1.956	0.474
51-F/51-R	8	1.432	-0.452	0.345	0.356	0.318	0.219	0.475	0.317
49-F/49-R	11	2.069	-0.046	0.202	0.789	0.492	0.471	0.989	0.530
21-F/21-R	14	1.482	0.004	0.193	0.482	0.277	0.278	1.047	0.329
Average	8	1.776	-0.381	0.356	0.567	0.509	0.359	0.675	0.511

Note: *N*a, observed number of alleles; *N*e, effective number of alleles; F_{IS}, inbreeding coefficient; F_{ST}, variance among populations; *I*, Shannon's information index; *H*o, observed heterozygosity; *H*e, expected heterozygosity; N_m, gene flow. PIC, polymorphic information content.

3.3 Genetic differentiation and genetic variation

Genetic variation analysis of *P. villosa* populations using low-copy nuclear gene data revealed that the proportion of genetic variation among populations was 76.6%, and that within populations was 23.4%, with F_{ST} = 0.766 (P< 0.01) (Table 7). These results showed that the genetic variation of this species existed primarily among populations. Meanwhile, this parameter of *P. villosa* population grouping revealed that 40.71% of the genetic variation was explained, and the genetic variation originated mainly from among populations accounting for 41.53% of the total genetic variation (Table 7). Genetic variation analysis of *P. villosa* populations using SSR data showed similar results.

Table 7
Results of analyses of molecular variance (AMOVAs) for *P. villosa* based on low-copy nuclear genes and SSR data

Data	Grouping	Source of variation	*df*	SS	VC	Percent variation (%)	Fixation index
LCNG	Total populations	Among populations	40	2242.692	6.384 Va	76.60	F_{ST} = 0.766**
		Within populations	307	598.808	1.951 Vb	23.40	
		Total	347	2841.500	8.334		
	SAMOVA groups	Among groups	1	697.447	4.396 Va	40.71	F_{CT} = 0.407**
		Among populations within groups	39	1555.148	4.485 Vb	41.53	F_{SC} = 0.700**
		Within populations	307	589.008	1.919 Vc	17.76	F_{ST} = 0.822**
		Total	347	2841.603	10.800		
SSR	Total populations	Among populations	38	1027.954		58	F_{ST} = 0.577**
		Within populations	156	539.600		42	
		Total	194	1567.554			

Note: *df*, Degrees of freedom; SS, Sum of squares; VC, Variance components; F_{CT}, Variance among groups relative to total variance; F_{SC}, Variance among populations within groups; F_{ST}, Variance among populations; Significant level: **P< 0.01.

Similarly, the genetic differentiation coefficient among populations based on low-copy nuclear gene data ranged from 0.533 to 0.977, with an average of 0.754, indicating a high level of genetic differentiation among populations (Table S4). The PERMUT analysis results showed that the total genetic diversity (H_t) of the population was 1.000, while the within-population genetic diversity (H_s) was 0.721, suggesting that the genetic diversity of the population existed primarily within populations (Table 8). The genetic differentiation coefficient (G_{ST}) among populations was 0.279, and the genetic differentiation coefficient within populations (N_{ST}) was 0.533 (Table 8). The coefficient of genetic differentiation in G_{ST} was significantly lower than that in N_{ST} (P< 0.05), which

indicated that there was obvious pedigree geographical structure in the *P. villosa* populations. Although there were differences in the estimation of genetic differentiation among populations by different primers or different gene fragments, they all showed that there was high genetic differentiation among populations. In addition, the study of genetic differentiation based on SSR data also supported this result, with F_{ST} = 0.356 (P < 0.01) (Table 6).

Table 8
Nei'S analysis of genetic diversity of *P. villosa* populations

Item	*Ht*	*Hs*	G_{ST}	N_{ST}
Mean	1.000	0.721	0.279	0.533
Standard error	0.002	0.034	0.034	0.035

Note: *Ht*, Total gene diversity; *Hs*, The average gene diversity within populations; G_{ST}, Gene differentiation coefficient between populations; N_{ST}, Gene differentiation coefficient within populations.

3.4 Factors influencing species differentiation

The Mantel test for IBD was conducted to estimate the correlation between the genetic distance and geographical distance of *P. villosa*. Positive correlation of genetic distance with geographical distance produced a pattern of isolation by distance. Applying the Mantel test, there was a significant positive correlation between genetic and geographical distances (r = 0.291, P = 0.010) for the entire sample (Fig. 5a). A positive correlation was also detected for Group 1 and Group 2 between genetic and geographical distances, respectively (Fig. 5b, c). Subsequently, correlation analysis between the genetic differentiation (F_{ST}) and climate factors consistently revealed a positive correlation, showing a significant correlation only with altitude (Fig. 5d). These findings confirmed that both geographical distance and altitude strongly influenced genetic differentiation.

3.5 Historical demography

The majority of genetic loci in *P. villosa* exhibited nonsignificant positive values for SSD and H_{Rag}, while those of 14075, 13604, 14192, and 14928 loci showed significantly positive values (Table 9). Estimation based on combined fragments showed that both parameters were positively correlated, with SSD and H_{Rag} values of 0.123 and 0.309, respectively (Table 9). We tentatively speculated that the species might undergo population expansion. Notably, mismatch distribution analysis for the entire populations generated nonsignificant values of the H_{Rag} and SSD indices. It produced bimodal curves, suggesting rapid population expansion (Fig. 6). Accordingly, the neutrality tests resulted in nonsignificant positive values of Tajima's *D*, Fu and Li's D^*, and Fu and Li's F^* for genetic loci except 14075 loci (Table 9). Conversely, the parameters estimated on the basis of combined fragments all showed highly significant positive values, with values of *D*, D^*, and F^* of 2.922, 2.425, and 3.198, respectively (Table 9). Overall, these findings confirmed the rejection of population expansion.

Table 9
The results of neutrality tests and mismatch distribution of *P. villosa* at ten nuclear loci

Locus	Neutrality test			Mismatch distribution	
	D	*D**	*F**	SSD	H_{Rag}
14031	0.395	0.736	0.742	0.008	0.075
14004	0.691	0.736	0.859	0.064	0.203
13593	1.539	0.938	1.370	0.077	0.281
14111	1.660	1.729**	2.057**	0.196	0.458
14075	-0.369	0.938	0.584	0.010*	0.046*
13604	2.775*	0.604	1.554	0.001*	0.014*
14192	1.362	0.736	1.124	0.012*	0.044*
14928	0.607	0.844	0.912	0.009*	0.050*
14224	2.966**	1.289	2.300**	0.146	0.391
14612	2.250*	1.097	1.809*	0.052	0.249
Combine	2.922**	2.425**	3.198**	0.123	0.309

Note: *P < 0.05; **P < 0.01.

Subsequently, the historical demography of *P. villosa* was simulated using DIYABC v.2.0.4 software. According to the population genetic structure of *P. villosa*, we designed three scenario models (Fig. 7): (1) Group 1 and Group 2 diverged at time T_0, while Group 1 and ancestral species (N_A) diverged at time

T_1; (2) Group 2 and Group 1 diverged at time T_0, while Group 2 and ancestral species (N_A) diverged at time T_1; (3) the two groups were relatively independent, with a common ancestral species (N_A), and diverged in a radiating pattern at time T_1. By comparing the posterior distribution probabilities for three models, Scenario 2 was selected as the most accurate in model simulations with LCNG and SSR data (Table S5). Under Scenario 2, Group 2 diverged from N_A approximately 8.20 Ma (95% HPD: 2.81 ~ 14.2 Ma) and 4.42 Ma (95% HPD: 0.396 ~ 13.3 Ma), respectively (Table 10). Moreover, the timing of divergence between Group 2 and Group 1 was 3.96 Ma (95% HPD: 1.36 ~ 8.10 Ma) and 0.109 Ma (95% HPD: 0.03 ~ 0.275 Ma), respectively (Table 10). As shown in Table 10, parameter estimates obtained from the best model showed that the effective population size of N_A, Group 1, and Group 2 was 2.14×10^6 (95% HPD: 5.53×10^5~2.93×10^6), 3.84×10^5 (95% HPD: 1.40×10^5~7.51×10^5), and 4.65×10^5 (95% HPD: 1.53×10^5~9.15×10^5) based on LCNG, respectively. The effective population size of N_A, Group 1, and Group 2 was 2.01×10^6 (95% HPD: 4.27×10^5~2.91×10^6), 3.55×10^5 (95% HPD: 8.67×10^4~8.74×10^5), and 6.57×10^5 (95% HPD: 2.35×10^5~1.62×10^5) based on SSR data, respectively. Group 1 and Group 2 had a smaller effective population size than N_A, which indicated that there had been contraction events in the history of *P. villosa.* This finding was confirmed through the isolation with migration (IM) model. Surprisingly, the divergence times estimated for Group 1 and Group 2 based on the IM model differed from those based on the approximate Bayesian computation (ABC) model. Group 1 and Group 2 began to differentiate about 0.115 Ma (95% HPD: 0.0984 ~ 0.475 Ma) and 0.180 Ma (95% HPD: 0.148 ~ 0.328 Ma) with LCNG and SSR data (Table 11), respectively. In addition, the model strongly supported bidirectional and asymmetric gene flow between the two groups, and the migration rate from Group 1 to Group 2 ($M2 \rightarrow 1$ = 1.87 or 0.08) was lower than that in the opposite direction ($M1 \rightarrow 2$ = 4.99 or 0.19) (Table 11). Therefore, we inferred that this species occurred population contraction in the history, and the divergence between Group 1 and Group 2 occurred during the Pleistocene (0.0117 ~ 2.59 Ma).

Table 10
Estimates of the posterior distributions of demographic parameters revealed by Approximate Bayesian Computation (ABC) for the best Model (Scenario 2)

Data	Parameter	N_1	N_2	N_A	T_0	T_1
LCNG	Mean	4.03×10^5	4.90×10^5	1.99×10^6	4.22×10^6	8.30×10^6
	Median	**3.84×10^5**	**4.65×10^5**	**2.14×10^6**	**3.94×10^6**	**8.20×10^6**
	Mode	3.16×10^5	3.45×10^5	2.98×10^6	3.73×10^6	7.85×10^6
	HPD95% lower	1.40×10^5	1.53×10^5	5.53×10^5	1.36×10^6	2.81×10^6
	HPD95% upper	7.51×10^5	9.15×10^5	2.93×10^6	8.10×10^6	1.42×10^7
SSR	Mean	4.00×10^5	7.58×10^5	$1.88e\times10^6$	1.26×10^5	5.40×10^6
	Median	**3.55×10^5**	**6.57×10^5**	**2.01×10^6**	**1.09×10^5**	**4.42×10^6**
	Mode	2.31×10^5	5.63×10^5	2.94×10^6	6.90×10^4	1.15×10^5
	HPD95% lower	8.67×10^4	2.35×10^5	4.27×10^5	3.00×10^4	3.96×10^5
	HPD95% upper	8.74×10^5	1.62×10^6	2.91×10^6	2.75×10^5	1.33×10^7

N_1 = effective population size of Group 1 for *P. villosa*.

N_2 = effective population size of Group 2 for *P. villosa*.

N_A = effective population size of ancestral population.

T_0 = time since divergence between Group 1 and Group 2.

T_1 = time since divergence between ancestral population and Group 2.

Table 11
Demographic parameters scaled by the mutation rate from IMa multilocus analyses based on the group from SAMOVA

Data	Parameter	θ_1	θ_2	θ_A	m_1	m_2	t	N_1	N_2	N_A	T[year]	$2N_1m_1$	$2N_2m$
LCNG	Hipt	0.0341	0.0854	1.2951	4.5375	4.4125	0.0011	2.80×10^5	7.00×10^5	1.06×10^7	1.80×10^5	0.08	0.19
	HPD95% lower	0.0173	0.0425	0.357	1.4875	1.2875	0.0009	1.42×10^5	3.48×10^5	2.93×10^6	1.48×10^5	0.01	0.03
	HPD95% upper	0.1526	0.3178	13.6443	23.0875	21.9625	0.0020	1.25×10^6	2.60×10^6	1.12×10^8	3.28×10^5	1.76	3.49
SSR	Hipt	0.1875	0.7492	378.3522	9.995	19.990	0.0007	1.54×10^6	6.14×10^6	3.10×10^9	1.15×10^5	1.87	4.99
	HPD95% lower	0.1875	0.7492	260.9756	5.575	17.290	0.0006	1.54×10^6	6.14×10^6	2.14×10^9	9.84×10^4	1.05	4.32
	HPD95% upper	0.5626	0.7492	590.629	9.965	19.990	0.0029	4.61×10^6	6.14×10^6	4.84×10^9	4.75×10^5	1.87	4.99

θ_1, N_1 = effective population size of Group 1 for *P. villosa*.

θ_2, N_2 = effective population size of Group 2 for *P. villosa*.

θ_A, N_A = effective population size of ancestral population.

t, T = time since divergence between Group 1 and Group 2.

$N = \theta/4\mu G$, G = 5; $T = t/\mu$; $2Nm = \theta \times m/2$.

3.6 Divergence time estimation and gene flow

After concatenating 10 sets of nuclear gene fragments, this study selected the best nucleotide substitution model as the GTR + I + G by AIC calculation. As shown in Fig. 8, the topology of the maximum clade credibility (MCC) tree was consistent with that of the phylogenetic tree. Subsequent divergences within *P. villosa* occurred at various times throughout the Pleistocene, which was consistent with the results obtained based on IM model. The divergence time between *P. villosa* and *Achnatherum splendens* was 3.26 Ma (95% HPD: 2.30 ~ 4.63 Ma). Group 1 and Group 2 began to differentiate at about 0.38 Ma (95% HPD: 2.30 ~ 4.63 Ma). Within the 41 populations, divergence occurred at approximately 0.02–0.55 Ma.

The species had significant genetic barriers in the Yinshan Mountains, Helan Mountains, Ordos Plateau and Yabulai Mountains based on SSR data by Barrier v.2.2 software (Fig. 9). The direction of gene flow among populations was analyzed by Migrate-N software. This study set up four possible models: (1) gene flow is bidirectional between Group 1 and Group 2; (2) gene flow is unidirectional from Group 1 to Group 2; (3) gene flow is unidirectional from Group 2 to Group 1; (4) Group 1 and Group 2 are the same population. The best model was determined by comparing the *b* values in ln(Prob(D|Model)), with the model with the largest *b* value considered to be the best. After comparison, the best model was model 1. The results showed that the gene flow from Group 1 to Group 2 was 0.432 and 0.299 from the past to the present, respectively (Table 12), whereas the gene flow from Group 2 to Group 1 was 0.489 and 0.397, respectively (Table 12). This indicated that the gene flow from Group 2 to Group 1 was larger than that accepted by this group, which further confirmed the results based on the IM model.

Table 12
Historical gene flow among different geographic groups in *P. villosa*

Data	Model	ln(Prob(D|Model))	θ1	θ2	M1→2	M2→1
LCNG	Model 1	-3387.640 (1a)	0.463	0.514	0.432	0.489
		-1835.250 (1b)				
	Model 2	-3351.850 (1a)	0.580	0.471	-	0.564
		-1853.990 (1b)				
	Model 3	-3305.390 (1a)	0.541	0.593	0.519	-
		-1849.670 (1b)				
	Model 4	-4194.297 (1a)	0.551	-	-	-
		-2116.930 (1b)				
SSR	Model 1	-312291.460 (1a)	0.423	0.369	0.299	0.397
		-52370.370 (1b)				
	Model 2	-327473.070 (1a)	0.154	0.399	-	0.387
		-55079.070 (1b)				
	Model 3	-317875.900 (1a)	0.384	0.174	0.485	-
		-53363.530 (1b)				
	Model 4	-470987.850 (1a)	0.395	-	-	-
		-77562.400 (1b)				

4 Discussion

4.1 Genetic diversity and environmental adaptation

In the relationship between plants and the environment, the environment has an ecological effect on plants, influencing and changing their morphological structure, physiological, and biochemical characteristics. At the same time, plants exhibit adaptability to the environment, adjusting to external changes through their own variations[34]. Every plant species has its unique gene pool and genetic organization, and genetic diversity determines its ability to adapt to the environment. This is also the foundation for maintaining the long-term stability of ecosystems[56]. When the genetic diversity of a species is higher or its genetic variation is richer, its ability to adapt to the environment is stronger. Meanwhile, when genetic diversity decreases, the species' ability to adapt, reproduce, and resist diseases weakens[57]. Therefore, studying the genetic diversity of species not only helps to explore their adaptability and evolutionary history, but also provides an important theoretical basis for the conservation and rational utilization of species[58].

Based on 10 LCNG and 13 SSR primer pairs, this study explored the genetic diversity of *P. villosa*. The results showed that the 10 selected primer pairs exhibited high polymorphism, with values for H_d, π, and K ranging from 0.000 to 1.000, 0.000 to 0.203, and 0.000 to 10.578, respectively. At the population level, the analysis of polymorphic loci using 13 SSR primer pairs revealed that the values for PIC and I ranged from 0.317 to 0.747 and from 0.278 to 0.946, respectively. These findings collectively indicate significant genetic diversity among *P. villosa* populations. Moreover, the populations P30 and P31 located in Ningxia, China, showed relatively higher genetic diversity, consistent with previous results obtained from internal transcribed spacer (ITS) sequences and SSR molecular markers[59–60]. Furthermore, the genetic diversity of *P. villosa* showed similarities as well as differences when compared with other species in the Stipeae tribe. For example, Zhang et al.[61] studied the genetic diversity of seven populations of *Stipa grandis* distributed in the Inner Mongolia region using RAPD. The results showed that the genetic diversity within populations (H_s = 0.17) was higher than that between populations (H_t = 0.27). Current research suggests that the factors influencing species genetic diversity mainly include natural selection, pollination pathways, gene flow, genetic drift, and genetic mutations. The potential drivers of within-species genetic diversity include biological factors such as dispersal ability and life history, as well as environmental factors such as climate and human activities[62–63]. Accordingly, we infer that the genetic diversity of *P. villosa* might be rich because it is a long-lived perennial plant that can reproduce sexually through seeds and seedlings, and clonally under harsh environmental conditions. Moreover, *P. villosa* has strong reproductive and vegetation restoration abilities, and its adaptation to the growth environment also contributes to the formation of its genetic diversity.

4.2 Population genetic differentiation and genetic structure

The genetic diversity and structure of a species are long-term products of evolution[64]. Population genetic structure includes genetic differentiation between populations and genetic variation within populations, reflecting the effects of mutation, recombination, genetic drift, and selection during the evolutionary process[65] Factors influencing the genetic structure of a species mainly include life habits, evolutionary history, reproductive systems, seed dispersal mechanisms, and gene flow[66]. According to Wright's related theory and recent studies[39], when the F_{ST} value exceeds 0.25, there is significant genetic

differentiation between populations of a species. In this study, based on LCNG and SSR data, the F_{ST} of *P. villosa* were estimated to be 0.754 and 0.356, which may be related to its reproductive mode. Meanwhile, the gene flow size of *P. villosa* was 0.097, which is larger than that of self-pollinating plants (M_m = 0.065), but smaller than that of cross-pollinating plants (M_m = 5.38)[67]. Therefore, we hypothesize that sexual reproduction in *P. villosa* may occur primarily through self-fertilization, consistent with previous research findings. AMOVA results showed that the majority of the genetic variation in *P. villosa* populations is derived from among-population variation. Hensen et al.[68] using RAPD studied the genetic variation of *Stipa capillata* and found that genetic variation also comes primarily from among-population variation. Meanwhile, Lv et al.'s[30] AFLP-based study suggested that genetic variation in *P. villosa* originates primarily within populations. The clonal propagation characteristics of this species lend greater credibility to subsequent analyses employing LCNG and SSR molecular markers. Previous studies have shown that species of the *Stipa* genus reproduce primarily asexually. When there is a barrier to gene flow between populations, the proportion of genetic variation between populations significantly increases, while the level of genetic variation within populations decreases[69]. The results of this study also indicate that $G_{ST} < N_{ST}$ ($P < 0.05$), suggesting a significant correlation between the genetic variation of *P. villosa* populations and their geographic origin. Mantel tests further revealed a significant correlation between genetic distance and spatial distance ($r = 0.291$, $P = 0.01$), implying that spatial distance may be one of the important factors influencing genetic differentiation in *P. villosa* populations. At the same time, correlation analysis between the genetic diversity, F_{ST} values, and environmental factors indicated that altitude is also a major influencing factor in the genetic differentiation of *P. villosa*. However, there is uncertainty regarding the genetic structure of species populations and their geographic locations[70]. That is, populations distributed in different geographic locations may not exhibit genetic differentiation or genetic variation, while populations located in the same geographic area may also have distant phylogenetic relationships. Therefore, this study employed various analytical methods, including unweighted pair group method with arithmetic mean (UPGMA), SAMOVA, PCA, and Structure, to investigate the genetic structure of *P. villosa* populations. The results showed that the *P. villosa* populations were divided into two main groups, Group 1 and Group 2, with the Yinshan Mountains acting as a boundary. Genetic introgression was observed between the groups. This is because the Yinshan Mountains have a unique geographical location, with significant climate differences on either side due to varying altitude and mountain orientation, leading to distinct geographical differences in plant species distribution and vegetation. In addition, this study found that *P. villosa* populations that were geographically closer did not always cluster together, and individuals within the same population were not always grouped, especially in populations P13, P36, P42, and P43. We believe this may be due to the fact that these populations are distributed mainly in the Helan Mountains region. The Helan Mountains are situated in a complex and unique tectonic setting, with intense and frequent tectonic activity since the Paleozoic, exhibiting a "multicycle" characteristic[71]. This acts as a transitional zone between sandy land and grassland vegetation. Therefore, *P. villosa* populations in this region may have introduced new genes from other populations, while populations outside the Helan Mountains may have dispersed genes favorable for the local habitat. In summary, we believe that the limited gene flow and unique reproductive systems among *P. villosa* populations have led to a high level of genetic differentiation and variation in the species. Spatial distance and altitude are probably important factors influencing the genetic structure of *P. villosa* populations[72].

4.3 Speciation and population demographic history

The Quaternary glacial climate oscillations are considered to have played a significant role in shaping the contemporary geographic distribution and population dynamics of plants[73]. Repeated climate fluctuations led to varying degrees of population contraction and expansion, with some species surviving in small climate zones known as glacial refugia[74]. Due to the uplift of the Qinghai-Xizang Plateau, the arid climate and complex terrain of the northwestern desert regions of China, along with the predominantly east–west orientation of many mountain ranges, the Yinshan, Helan, Qilian, and Tianshan mountains have formed natural glacial refugia and geographical barriers[75]. Current research generally suggests that regions with high nucleotide diversity and haplotype polymorphism may be glacial refugia[76]. The *P. villosa* populations around the Yinshan and Helan mountain ranges tend to have higher nucleotide diversity and haplotype polymorphism. Among these, the P43 in Dalate Banner has the highest nucleotide polymorphism ($\pi = 0.203$) and haplotype diversity (H_d = 0.978), which indicates that the development of mountain ranges leads to spatial isolation, promoting differentiation between populations of the same species, thus influencing the species' geographic distribution pattern. Moreover, genetic flow barrier detection between *P. villosa* populations reveals that genetic flow predominantly moves from Group 2 to Group 1. The genetic barriers originate mainly from the Yin Mountains, Helan Mountains, Ordos Plateau, and Yablun Mountains. We believe these areas may have been the main refugia for *P. villosa* during the Quaternary glaciations[77].

Neutrality tests and mismatch distribution analysis results both suggest that the population dynamics of *P. villosa* did not experience a rapid expansion event in its history. BEAST and ABC model predictions indicate that *P. villosa* may have originated in the Late Miocene, with population divergence occurring during the Pleistocene. The divergence time between Group 1 and Group 2 is approximately 3.94 Ma and 0.11 Ma, while the divergence time between Group 2 and N_A is around 8.20 Ma and 4.42 Ma, which aligns with the previous divergence time (9 ~ 2.61 Ma) of species in the *Stipa* genus. Furthermore, the effective population sizes of Group 1 and Group 2 are both smaller than that of N_A. These findings confirm that *P. villosa* underwent contraction events in its history, probably due to spatial isolation caused by changes in local mountains (such as the Yinshan and Helan Mountains) and deserts (such as the Mu Us Desert, Tengger Desert, and Badain Jaran Desert), as well as population range contraction and habitat fragmentation caused by climatic oscillations. This is consistent with the historical dynamic processes observed in other desert plants such as *Gymnocarpos przewalskii* and *Helianthemum songaricum*[78–80]. Since the Quaternary glaciation, the third uplift of the Qinghai-Xizang Plateau around 8 Ma, the ongoing global cooling and drying, and the influence of the east–west monsoon climate of the Yin Mountains, which led to desert expansion and habitat fragmentation, are important factors involved in the genetic differentiation between populations of *P. villosa* in response to environmental changes.

5 Conclusion

This study investigated the genetic diversity, population genetic structure, and demographic history of 210 individuals from 43 natural populations of *P. villosa* using LCNG and SSR molecular markers. We found that the *P. villosa* populations were clearly divided into two groups, Group 1 and Group 2, by the

Yinshan Mountains, with populations in Group 2 showing higher genetic diversity. Significant diversity differences and phylogeographic structure were observed among the populations, with most genetic variation occurring between populations. Geographic distance and elevation had a strong influence on genetic differentiation. No significant population expansion was detected in the history of *P. villosa* populations, with divergence occurring around 0.38 Ma between Group 1 and Group 2. There was bidirectional and asymmetric gene flow between the two groups. The Yinshan Mountains, Helan Mountains, Ordos Plateau, and Yabulai Mountains serve as natural and significant genetic barriers. The uplift of the Qinghai-Xizang Plateau and the influence of the monsoonal climate of the Yinshan Mountains since the Quaternary period have been important drivers of the genetic differentiation between *P. villosa* populations, as they adapted to environmental changes.

Abbreviations

LCNG
low-copy nuclear genes
SSR
simple sequence repeats
RFLP
fragment length polymorphisms
RAPD
random amplified polymorphic DNA
AFLP
amplified fragment length polymorphisms
SNP
single-nucleotide polymorphisms
ISSR
inter-simple sequence repeat
PCA
principal coordinate analysis
N
the number of individuals
PPL
percentage of polymorphic loci
PIC
polymorphic information content
N_a
number of alleles
Ne
effective number of alleles
I
Shannon's information index
H_o
observed heterozygosity
H_e
expected heterozygosity
H_d
haplotype diversity
π
nucleotide diversity
K
average number of nucleotide differences
h
Nei's genetic diversity
H_t
total gene diversity
H_s
subpopulation gene diversity
F_{IS}
inbreeding coefficient
F_{ST}
Fixation Index
G_{ST}
genetic differentiation coefficient

*N*m
gene flow
AMOVA
analysis of molecular variance
SAMOVA
spatial AMOVA
MCMC
Markov Chain Monte Carlo
IM
isolation with migration
ABC
approximate Bayesian computation.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing Interests

The authors declare no competing interests.

Funding

This study was financially supported by the National Natural Science Foundation of China (Grant Number: 32160297, 41761009), the Program of Science and Technology International Cooperation Project of Qinghai Province (Grant Number: 2023-HZ-810), and the Qinghai Provincial Major Science and Technology Special Project (2023-SF-A5).

Author Contribution

J.R. and T. L. : Writing – review, editing & original draft. J.R. : Formal analysis. Y. P. & X. S. : Conceptualization. we greatly appreciated Y. H. , Q. Y. , Y. L. , X. F. and K. Y. performing the investigation & experiment, and the anonymous reviewers for their helpful comments.

Acknowledgements

This work was financially supported by the National Natural Science Foundation of China (No. 32160297, 41761009) made to Yuping Liu, the Program of Science and Technology International Cooperation Project of Qinghai Province (No. 2023-HZ-810) made to Yuping Liu, and the Qinghai Provincial Major Science and Technology Special Project (2023-SF-A5) made to Yuping Liu.

Data Availability

Genome assembly of P. villosa, Iso-seq and Hi-C data have been submitted to the DDBJ/EMBL/Genbank databases under BioProject number PRJ70313651: genome assembly–JABCND000000000; Iso-seq data–SRR11787985 and SRR11787359; RNA-seq data–SRR11775845, SRR11775821, SRR11775823, SRR11775824, SRR11775822, SRR11775544.

References

1. Ma C, Tang YJ, Ying JF. Volcanic phosphorus spikes associated with supercontinent assembly supported the evolution of land plants. Earth Sci Rev. 2022;232.
2. Bowen DQ. Questions of the Quaternary. Nature. 1984;311(5985):488.
3. Xu XW, Christina W, Michael FA, Mara LA, Sue L, Ge S, Wen J. Phylogeny and biogeography of the eastern Asian–North American disjunct wild-rice genus (*Zizania* L., Poaceae). Mol Phylogenet Evol. 2009;55(3):1008–17.
4. Li ZH, Chen J, Zhao GF, Guo YP, Kou YX, Ma YZ, Wang G, Ma XF. Response of a desert shrub to past geological and climatic change: A phylogeographic study of *Reaumuria soongarica* (Tamaricaceae) in western China. J Syst Evol. 2012;50(4):351–61.
5. Emile Z, Linus P. Molecules as documents of evolutionary history. J Theor Biol. 1965;8(2):357–66.

6. Ree RH, Moore BR, Webb CO, Michael JD. A likelihood framework for inferring the evolution of geographic range on phylogenetic trees. Evolution. 2005;59(11):2299–311.
7. Landis MJ, Matzke NJ, Moore BR, Huelsenbeck JP. Bayesian analysis of biogeography when the number of areas is large. Syst Biol. 2013;62(6):789–804.
8. Ye WQ, Yap ZY, Li P, Comes HP, Qiu YX. Plastome organization, genome-based phylogeny and evolution of plastid genes in Podophylloideae (Berberidaceae). Mol Phylogenet Evol. 2018;127:978–87.
9. Wang S, Wang YH, Sun JH, Cui XY, Li EZ, Wang RS, Li Q, Zhang PF, Dong WP, Guo LP, Huang LQ. Comparative chloroplast genome analyses provide new insights into molecular markers for distinguishing *Arnebiae Radix* and its substitutes (tribe Lithospermeae, Boraginaceae). Phytomedicine. 2025;136:156338–156338.
10. Ma ZL, Xu SS, Zhang Y, Marcos ACO, Su X, Ren ZM. Characterization and phylogenetic analysis of the complete chloroplast genome of *Xanthopappus subacaulis* (Asteraceae), an endemic medicinal plant from the Qinghai-Tibet Plateau in China. Cytol Genet. 2022;56(1):77–83.
11. Chen XD, Wu XJ, Zhang JL, Zhang M, You JC, Ru ZG. Characterization of the complete mitochondrial genome of *Elymus magellanicus* (É. Desv.) Á. Löve (Poaceae, Pooideae). Mitochondrial DNA Part B Resour. 2023;(8):795–8.
12. Park S, Hwang Y, Kim H, Choi KS. Insights into the nuclear-organelle DNA integration in *Cicuta virosa* (Apiaceae) provided by complete plastid and mitochondrial genomes. BMC Genomics. 2025;26(1):102.
13. Kim S, Lee YP, Lim H, Ahn Y, Sung SK. Identification of highly variable chloroplast sequences and development of cpDNA-based molecular markers that distinguish four cytoplasm types in radish (*Raphanus sativus* L). Theor Appl Genet. 2009;119(1):189–98.
14. Amiteye S. Basic concepts and methodologies of DNA marker systems in plant molecular breeding. Heliyon. 2021;7(10):e08093–08093.
15. Hernández GR, Mendoza CG, Magallón S. Low-copy nuclear genes reveal new evidence of incongruence in relationships within Malvaceae s. l. Syst Bot. 2021;46(4):1042–52.
16. Cheng L, Li M, Han QW, Qiao Z, Hao YL, Balbuena TS, Zhao YY. Phylogenomics resolves the phylogeny of Theaceae by using low-copy and multi-copy nuclear gene makers and uncovers a fast radiation event contributing to tea plants diversity. Biology. 2022;11(7):1007–1007.
17. Wu JW, Li YX, Zhong WR, Ran XY, Lyu G, Chen RJ, Zhao ZH, Tang DB, Wang JC, Lu HX, Zhang K. Genetic diversity evaluation and selection methods of sweet potato hybrid F1 population based on SSR markers and phenotypic detection. BMC Plant Biol. 2025;25(1):329–329.
18. Wang RQ, Xu Y, Zhang M, Hao G, Zhao QM, Liu XK, Liu XF, Yu B, Zhang WJ. Genetic structure and demographic analysis of a true single-population species, *Camellia azalea*. BMC Plant Biol. 2024;1:1272–1272.
19. Liu ML, Shang QH, Cheng YJ, Wang N, Sa W, Li BG, Li ZH. Drivers of intraspecific differentiation of an alpine cold-tolerant herb, *Notopterygium oviforme*:Roles of isolation by distance and ecological factors. J Syst Evol. 2022;(2):383–98.
20. Wang YN, Liu YP, Liu XL, Su X, Liu RJ. Chromosome number and karyotype analysis from different populations of *Psammochloa villosa* (Poaceae). Acta Bot Boreali-Occidentalia Sinica. 2021;41(9):1489–99.
21. Zhong LF, Feng XY, Zhao WZ. Fixing active sand dune by native grasses in the desert of Northwest China. Ecol Processes. 2024;13(1):77–77.
22. Lv T, Liu YP, Zhou YH, Liu T, Zhang XY, Su X. Germplasm collection and preliminary studies on genealogical differentiation of a desert species-*Psammochloa villosa*. Acta Agrestia Sinica. 2018;26(3):733–40.
23. Huang ZY, Dong M, Gutterman Y. Factors influencing seed dormancy and germination in sand, and seedling survival under desiccation, of *Psammochloa villosa* (Poaceae), inhabiting the moving sand dunes of Ordos, China. Plant Soil. 2004;259(1/2):231–41.
24. Yu FH, Dong M, Krüsi B. Clonal integration helps *Psammochloa villosa* survive sand burial in an inland dune. New Phytol. 2004;162(3):697–704.
25. Dong M. Effects of severing rhizome on clonal growth in rhizomatous grass species *Psammochloa villosa* and *Leymus secalinus*. Acta Bot Sinica. 1999;2:194–8.
26. Lv T, Liu YP, Zhou YH, Liu T, Zhang XY, Su X. Micromorphological characteristics of leaf epidermis and ecological adaptation of *Psammochloa villosa*, a desert plant from the Inner Mongolian Plateau. Bull Bot Res. 2018;38(03):330–7.
27. Liu YP, Liang RF, Lv T, Ma ZL, Liu T, Liu F, Su X, Harris AJ. Development of SSR markers for *Psammochloa villosa* (Trin.) Bor (Poaceae), a dominant species in the Inner Mongolian Plateau. Cytol Genet. 2021;55(6):576–82.
28. Li A, Ge S. Genetic variation and clonal diversity of *Psammochloa villosa* (Poaceae) detected by ISSR markers. Ann Botany. 2001;87:585–90.
29. Liu YP, Zhang Y, Liu F, Liu T, Chen JY, Fu G, Zheng CY, Su DD, Wang YN, Zhou HK, Su X, Harris AJ, Wang XM. Establishment of reference (housekeeping) genes via quantitative real-time PCR for investigation of the genomic basis of abiotic stress resistance in *Psammochloa villosa* (Poaceae). J Plant Physiol. 2021;268:153575–153575.
30. Lv T, Harris AJ, Liu YP, Liu T, Liang RF, Ma ZL, Su X. Population genetic structure and evolutionary history of *Psammochloa villosa* (Trin.) Bor (Poaceae) revealed by AFLP marker. Ecol Evol. 2021;11(15):10258–76.
31. Liu JT, Zhou YL, Luo CX, Xiang Y, An LZ. *De novo* transcriptome sequencing of desert herbaceous *Achnatherum splendens* (Achnatherum) seedlings and identification of salt tolerance genes. Genes. 2016;7(4):1–13.
32. Koressaar T, Remm M. Enhancements and modifications of primer design program Primer 3. Bioinformatics. 2007;23(10):1289–91.
33. Dupanloup I, Schneider S, Excoffier L. A simulated annealing approach to define the genetic structure of populations. Mol Ecol. 2002;11(12):2571–81.
34. Hubisz MJ, Falush D, Stephens M, Pritchard JK. Inferring weak population structure with the assistance of sample group information. Mol Ecol Resour. 2009;9(5):1322–32.
35. Pritchard JK, Stephens M, Donnelly P. Inferring of population structure using multilocus genotype data. Genetics. 2000;155(2):945–59.

36. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol. 2005;14(8):2611–20.
37. Zhang YH, Hou Y, Lou AR. Population genetic diversity of *Rhodiola dumulosa* in Northern China inferred from AFLP makers. Chin J Plant Ecol. 2010;34(9):1084–94.
38. Peakall R, Smouse P. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research–an update. Bioinformatics. 2012;28(19):2537–9.
39. Wright S. Evolution in Mendelian populations. Genetics. 1931;16:97–159.
40. Graves JE, McDowell JR. Stock structure of the world's istiophorid billfishes: a genetic perspective. Mar Freshw Res. 2003;54(4):287–98.
41. Excoffier L, Lischer H. Arlequin suite v3.5: a new series of programs to perform population genetic analyses under Linux and Windows. Mol Ecol Resour. 2010;10(3):564–7.
42. Slatkin M. A measure of population subdivision based on microsatellite allele frequencies. Genetics. 1995;1:457–62.
43. Liu Y, Dietrich CH, Wei C. Genetic divergence, population differentiation and phylogeography of the cicada *Subpsaltria yangi* based on molecular and acoustic data: an example of the early stage of speciation. BMC Ecol Evol. 2019;19:1–5.
44. Luo YF, Wu K, Fu D, Wang AP, Xu B. Population genetic structure of the Asian citrus psyllid, *diaphorina citri kuwayama* (Hemiptera: liviidae) in Jiangxi. Plant Prot. 2019;45(3):55–9.
45. Fu YX. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics. 1997;147:915–25.
46. Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, Marin JM, Estoup A. DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. Bioinformatics. 2014;30(8):187–1189.
47. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst Biol. 2018;67:901–4.
48. Gaut BS, Morton BR, McCaig BC, Clegg MT. Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene *Adh* parallel rate differences at the plastid gene *rbc*L. Proceedings of the National Academy of Sciences. 1996;93(19):10274–10279.
49. Xue ZQ, Greimler J, Paun O. The evolutionary history of New Zealand *Deschampsia* is marked by long distance dispersal, endemism, and hybridization. Biology. 2021;10(10):1–17.
50. Posada D. jModelTest: phylogenetic model averaging. Mol Biol Evol. 2008;25(7):1253–6.
51. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 2007;7(1):214.
52. Chen Q, Kang HY, Fan X, Wang Y, Sha LN, Zhang HQ, Zhong MY, Xu LL, Zeng J, Yang RW. Evolutionary history of *Triticum petropavlovskyi* Udacz. Et Migusch. inferred from the sequences of the 3-phosphoglycerate kinase gene. PLoS ONE. 2013;8(8):e71139.
53. Hedges SB, Marin J, Suleski M, Paymera M, Kumar S. Tree of life reveals clock-like speciation and diversification. Mol Biol Evol. 2015;32(4):835–45.
54. Manni F, Guerard E, Heyer E. Geographic patterns of (genetice, morphologic, lingustic) variation: how barriers can be detected by using Monmonier's algorithm. Hum Biol. 2004;76(2):173–90.
55. Beerli P. Migrate: documentation and program, part of LAMARC (Version 2.0). Seattle, WA, USA: Department of Genome Science, University of Washington; 2022.
56. Goessen R, Isabel N, Wehenkel C, Pavy N, Tischenko L, Touchette L, Giguère I, GrosLouis MC, Laroche J, Boyle B, Soolanayakanahally R, Mock K, Javier HV, SimentalRodriguez SL, Bousquet J, Porth IM. Coping with environmental constraints: Geographically divergent adaptive evolution and germination plasticity in the transcontinental *Populus tremuloides*. Plants, People, Planet. 2022;4(6):638–654.
57. Dillon S, McEvoy R, Baldwin DS, Rees GN, Parsons Y, Southerton S. Characterisation of adaptive genetic diversity in environmentally contrasted populations of *Eucalyptus camaldulensis* Dehnh. (river red gum). PLoS ONE. 2014;9(8):e103515.
58. Mugula BB, Omondi SF, Curto M, Kiboi SK, Kanya JI, Egeru A, Okullo P, Meimberg H. Microsatellites reveal divergence in population genetic diversity, and structure of *Osyris lanceolata* (Santalaceae) in Uganda and Kenya. BMC Ecol Evol. 2023;23(1):73–73.
59. Lv T. Studies on genetic diversity of *Psammochloa villosa*, a peculiar sandy plant. Qinghai Normal Univ. 2019;30–9.
60. Liang RF. Genetic structure and physiology of drought resistance in the population of *Psammochloa villosa*. Qinghai Normal Univ. 2021;35–6.
61. Zhang HM, Zhao ML, Li QF, Han B, Suo PF. Detection of genetic diversity of *Stipa grandis* in Inner Mongolia using RAPD markers. Acta Agrestia Sinica. 2003;2:170–7.
62. Wu Q, Dong S, Zhao YX, Yang L, Qi XJ, Ren ZH, Dong SB, Cheng J. 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. 2023;23(1):254–254.
63. Xu JW, Zou R, Chai SF, Jiang HL, Zhong WT, Tang JM, Wei X. Single nucleotide polymorphism markers based on hyper-seq revealed the genetic diversity and genetic relationship of *Horsfieldia hainanensis* in China. Forests. 2025;16(2):217–217.
64. Thaïs AB, Zdenek L, Petr K, Jacklyn MH, Silvia M, Jonah Y, Nicholas EM, Ken MJ. Assessing the impacts of adaptation to native-range habitats and contemporary founder effects on genetic diversity in an invasive fish. Evol Appl. 2024;17(10):e70006.
65. Jia SW, Wang Y, Li YC, Wu ZJ, Cai ZF, Shen J, Wang DR, Chen SQ. (2024) Population genetic structure of *Montipora digitata* coral-algal symbiosis in the South China Sea. Global Ecology and Conservation. 2024;55:e03231-e03231.
66. Xu R, Tatjana ND, Chen DL, Byambasuren M, Xu F, Guo XG. Phylogeography and ecological niche modeling of the Alashan Pit Viper (*Gloydius cognatus*; Reptilia, Viperidae) in Northwest China and Adjacent Areas. Animals. 2023;13(23):3726.

67. Peng JT, Liang CZ, Niu YM, Jiang W, Wang W, Wang LX. Moderate grazing promotes genetic diversity of *Stipa species* in the Inner Mongolian steppe. Landscape Ecol. 2015;30(9):1783–94.
68. Hensen I, Kilian C, Wagner V, Durka W, Pusch J, Wesche K. Low genetic variability and strong differentiation among isolated populations of the rare steppe grass *Stipa capillata* L. in Central Europe. Plant biology (Stuttgart, Germany). 2010;12(3):526–536.
69. Gagnaire PA, Yuki M, Saliha Z, Pierre V, Jun A, Patrick B. Within-population structure highlighted by differential introgression across semipermeable barriers to gene flow in *Anguilla marmorata*. Evolution; International Journal of Organic Evolution. 2011;65(12):3413–3427.
70. Miller MP, Bianchi CA, Mullins TD, Haig SM. Associations between forest fragmentation patterns and genetic structure in Pfrimer's Parakeet (*Pyrrhura pfrimeri*), an endangered endemic to central Brazil's dry forests. Conserv Genet. 2013;14(2):333–43.
71. Ma JH, He DF. Meso-Cenozoic tectonic events in the Helanshan tectonic belt and its adjacent areas: Constraints from unconformity and fission track data. Acta Petrologica Sinica. 2019;35(4):1121–42.
72. Guo J, Brown PJ, Rayburn AL, ButtsWilmsmeyer CJ, Boe A, Lee D. Genomic Variation shaped by environmental and geographical factors in prairie cordgrass natural populations collected across its native range in the USA. Genes. 2021;12(8):1240–1240.
73. Mäder G, Fregonezi JN, Lorenz LAP, Bonatto SL, Freitas LB. Geological and climatic changes in quaternary shaped the evolutionary history of *Calibrachoa heterophylla*, an endemic South-Atlantic species of petunia. BMC Evol Biol. 2013;13(1):178.
74. Bringloe TT, Fort A, Inaba M, Sulpice R, Ní GC, Agnes MM, Karen FD, Vieira C, Kawai H, Takeaki H, Dorte KJ, Birgit O, Samuel S, Heroen V. Whole genome population structure of North Atlantic kelp confirms high latitude glacial refugia. Mol Ecol. 2022;31(24):6473–88.
75. Guo P, Liu Q, Zhu F, Zhong GH, Che J, Wang P, Xie YL, Robert WM, Anita M. Multilocus phylogeography of the brown-spotted pitviper *Protobothrops mucrosquamatus* (Reptilia: Serpentes: Viperidae) sheds a new light on the diversification pattern in Asia. Mol Phylogenet Evol. 2019;13382–91.
76. Yang GS, Yang L, Shen SK, Wang YH, Wang YY. Unraveling the evolutionary tales of *Yunnanopilia longistaminea* (Opiliaceae): insights from genetic diversity, climate adaptation, and conservation strategies. Plants. 2025;14(5):706–706.
77. Yang YX, Wang ML, Liu ZL, Zhu J, Yan MY, Li ZH. Nucleotide polymorphism and phylogeographic history of an endangered conifer species *Pinus bungeana*. Biochem Syst Ecol. 2016;64:89–96.
78. Su ZH, Zhang ML, Sanderson SC. Chloroplast phylogeography of *Helianthemum songaricum* (Cistaceae) from northwestern China: implications for preservation of genetic diversity. Conserv Genet. 2011;12(6):1525–37.
79. Ma SM, Zhang ML, Sanderson SC. Phylogeography of the rare *Gymnocarpos przewalskii* (Caryophyllaceae): indications of multiple glacial refugia in north-western China. Aust J Bot. 2012;60(1):20–31.
80. Meng HH, Gao XY, Huang JF, Zhang ML. Plant phylogeography in arid Northwest China: Retrospectives and perspectives. J Syst Evol. 2015;53(1):33–46.

Figures

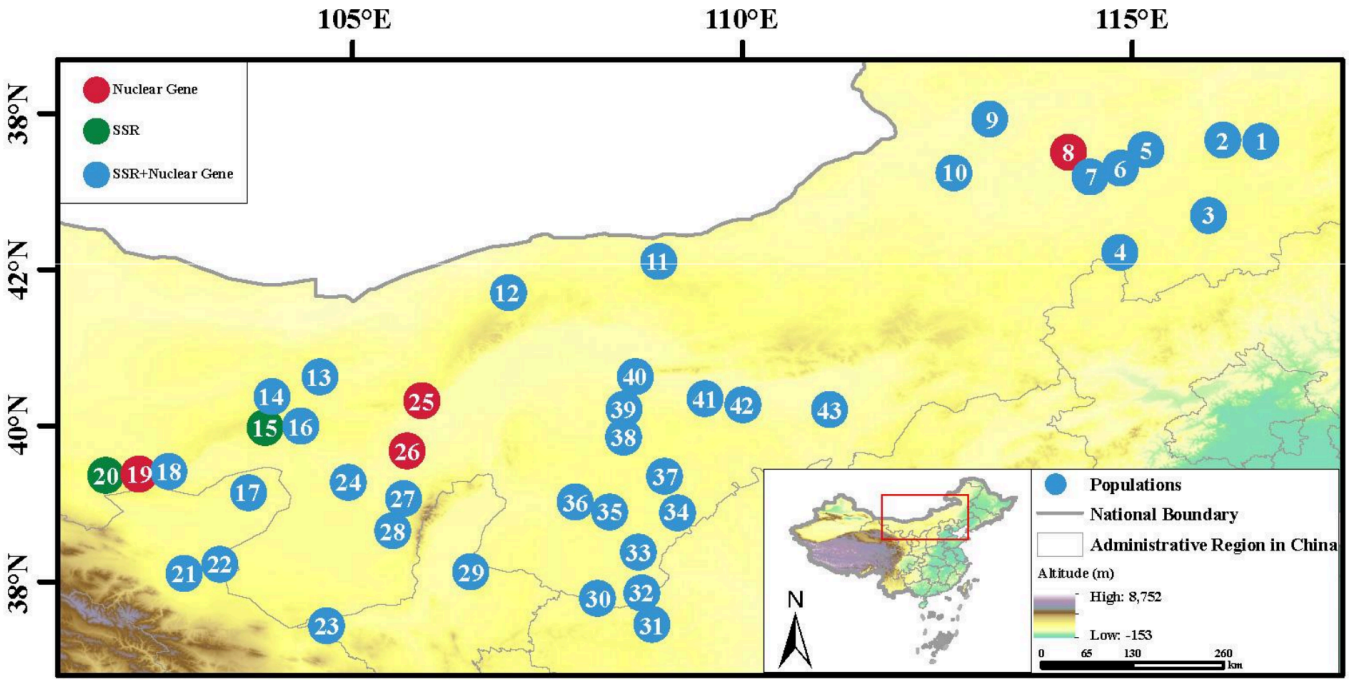


Figure 1 Localities of *P. villosa* surveyed in the Inner Mongolian Plateau.

Figure 1

See image above for figure legend

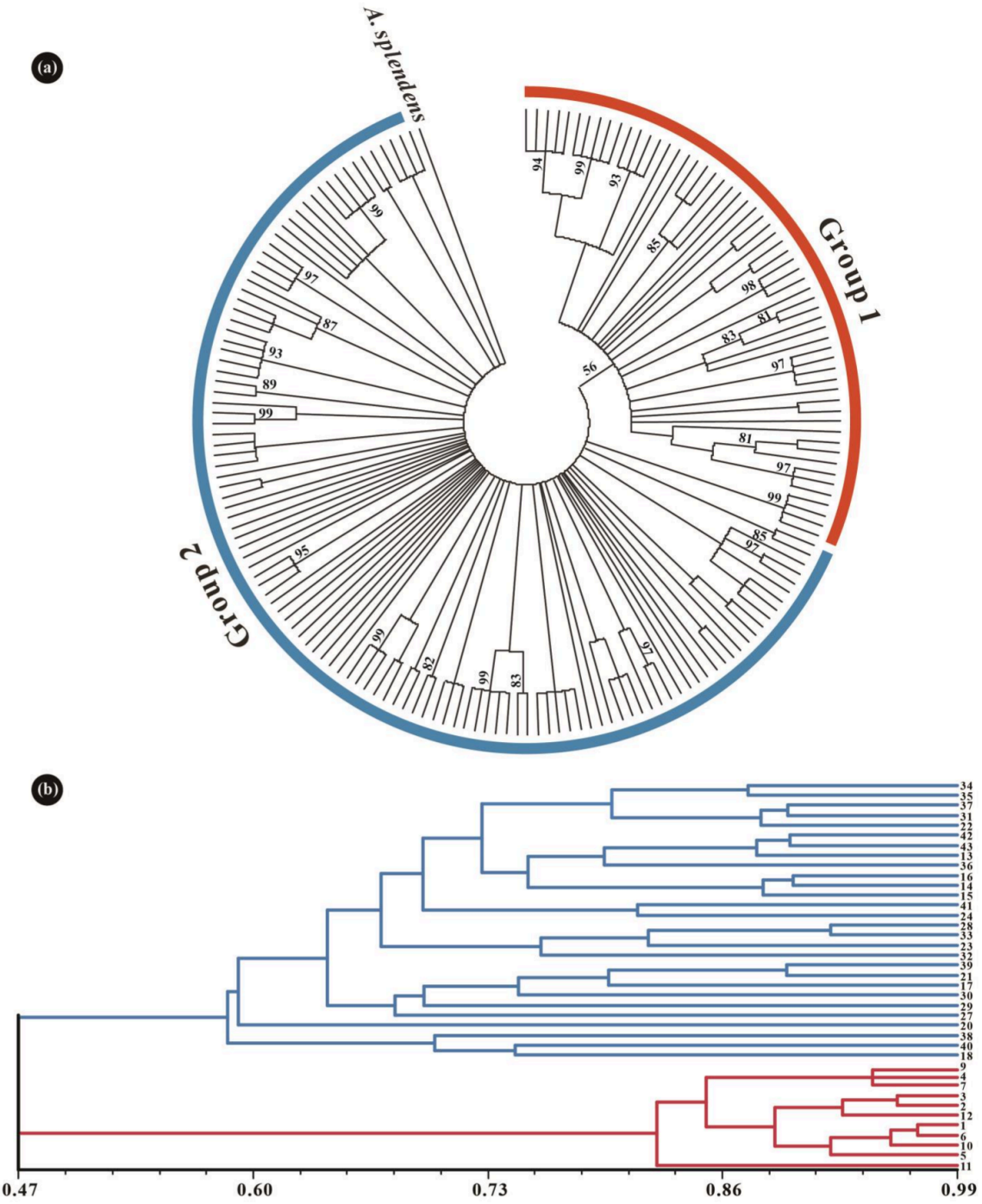


Figure 2 Phylogenetic tree constructed for *P. villosa* based on low-copy nuclear genes (a) and SSR data (b).

Figure 2

See image above for figure legend

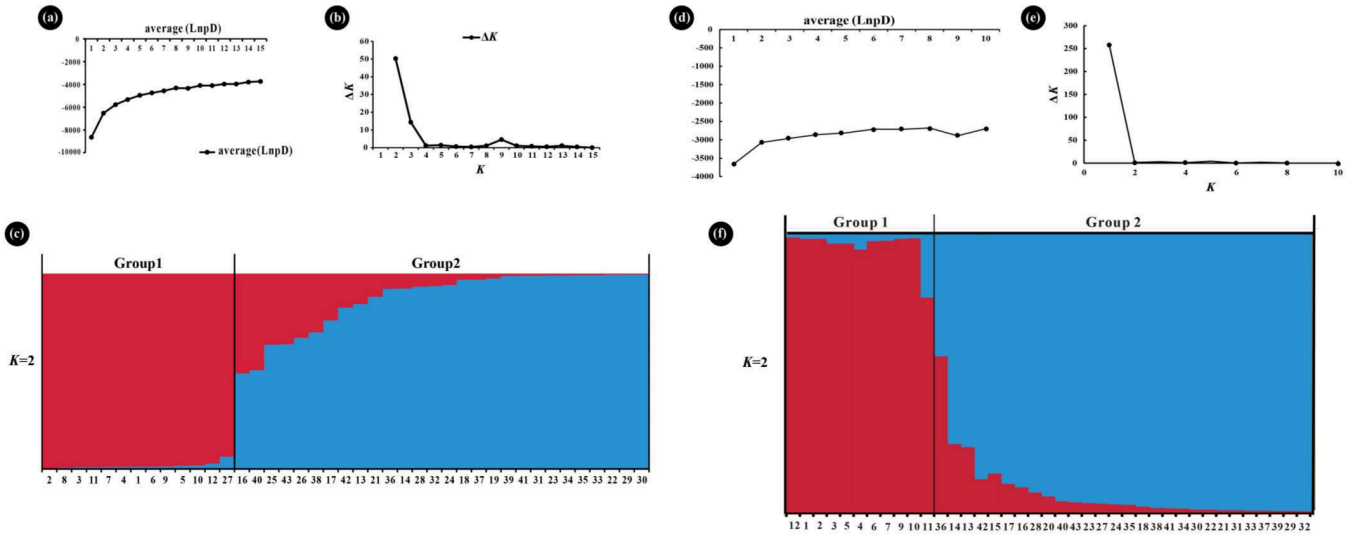


Figure 3 Structure analyses of *P. villosa* based on low-copy nuclear genes (a ~ c) and SSR data (d ~ f).

Note: (a) and (d), average Ln P(D) (±SD) over 15 and 10 runs for each *K* value; (b) and (e), the corresponding ΔK statistics calculated according to *K*; (c) and (f), histogram of genetic structure about the model with $K = 2$. The different colored bars (red and blue) refer to two different genetic pools (Group 1 and Group 2).

Figure 3

See image above for figure legend

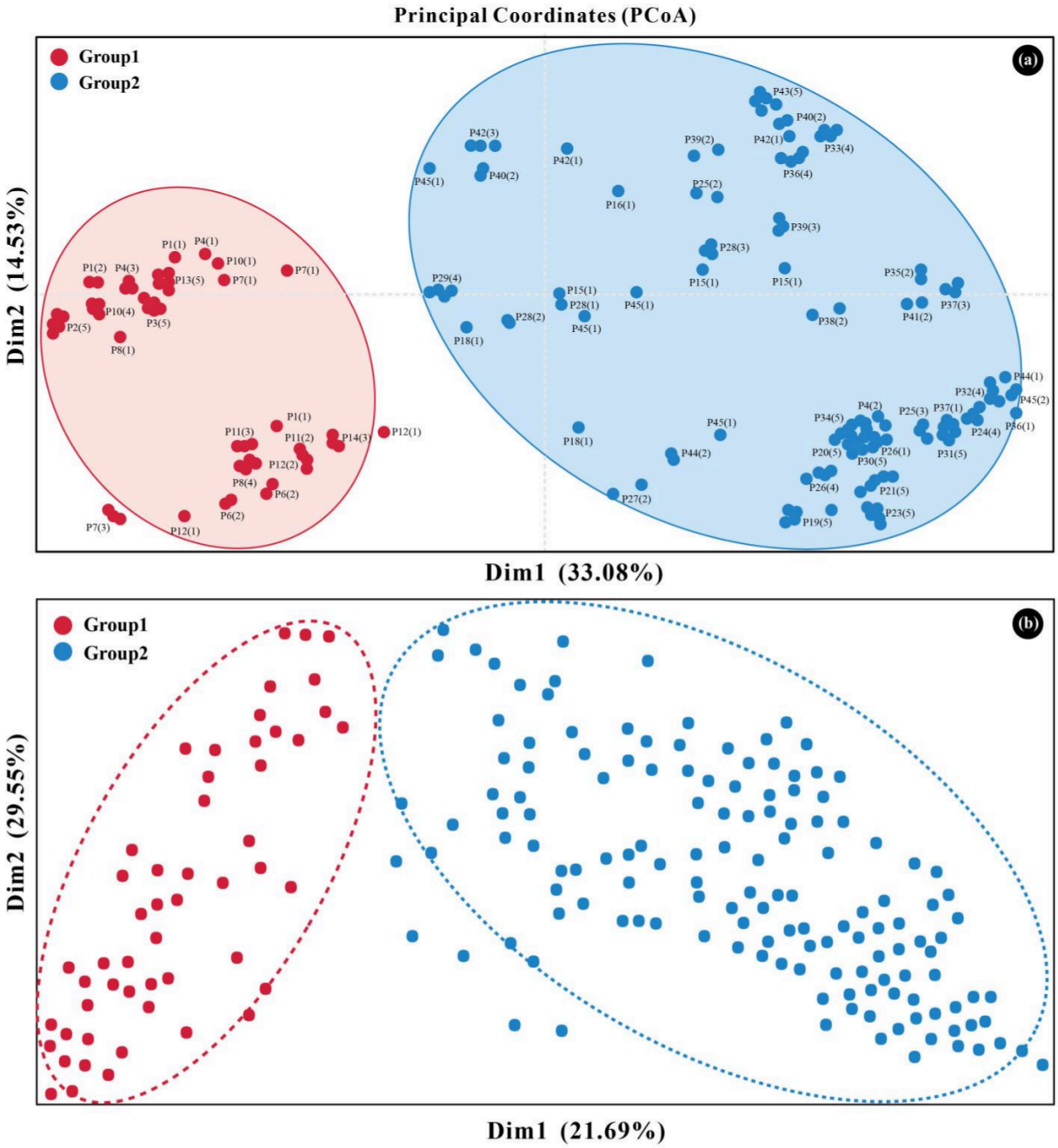


Figure 4 Principal coordinates analysis (PCoA) of *P. villosa* based on low-copy nuclear genes(a) and SSR data (b).

Figure 4

See image above for figure legend

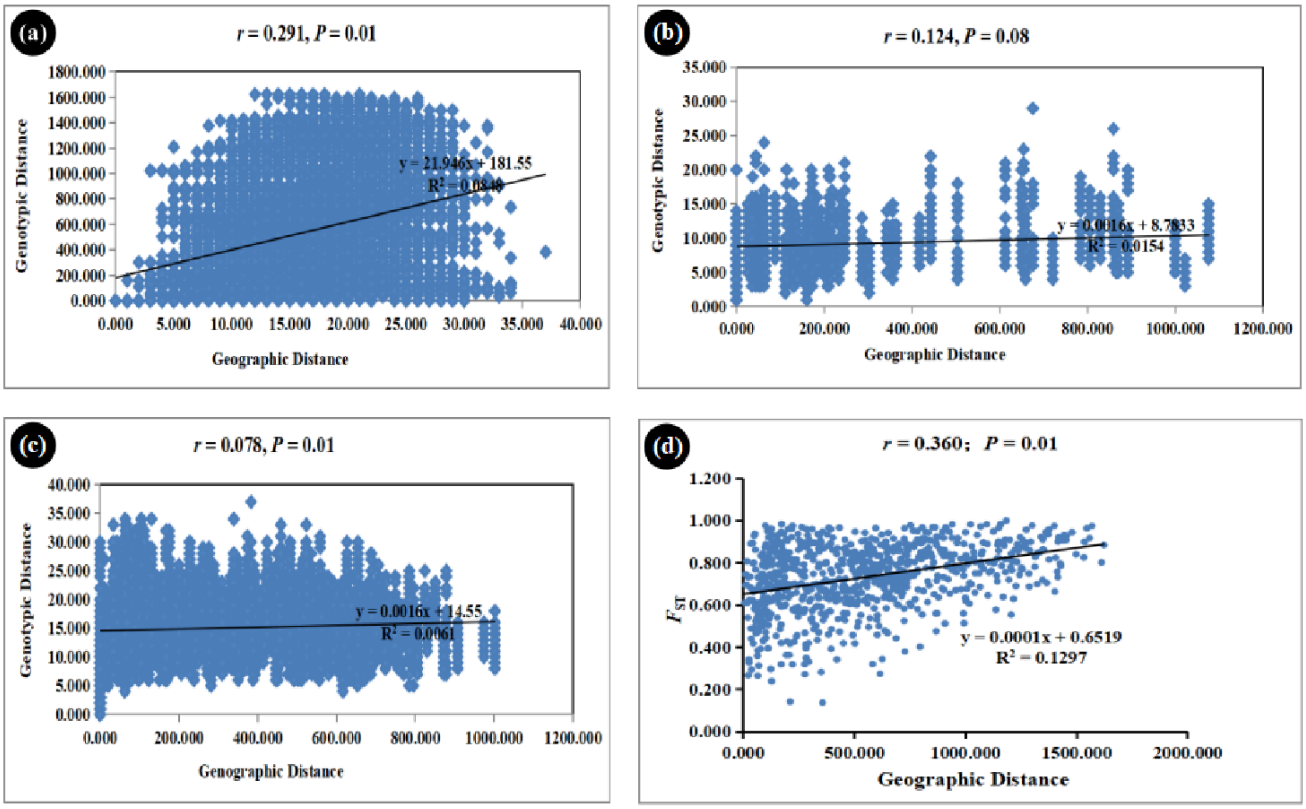


Figure 5 Correlation of mantel tests between geographic distance and genetic distance. (a) Mantel test from 39 populations of *P. villosa*; (b) Mantel test from populations of Group 1; (c) Mantel test from populations of Group 2; (d) The correlation between geographic distance and genetic distance based on low-copy nuclear genes.

Figure 5

See image above for figure legend

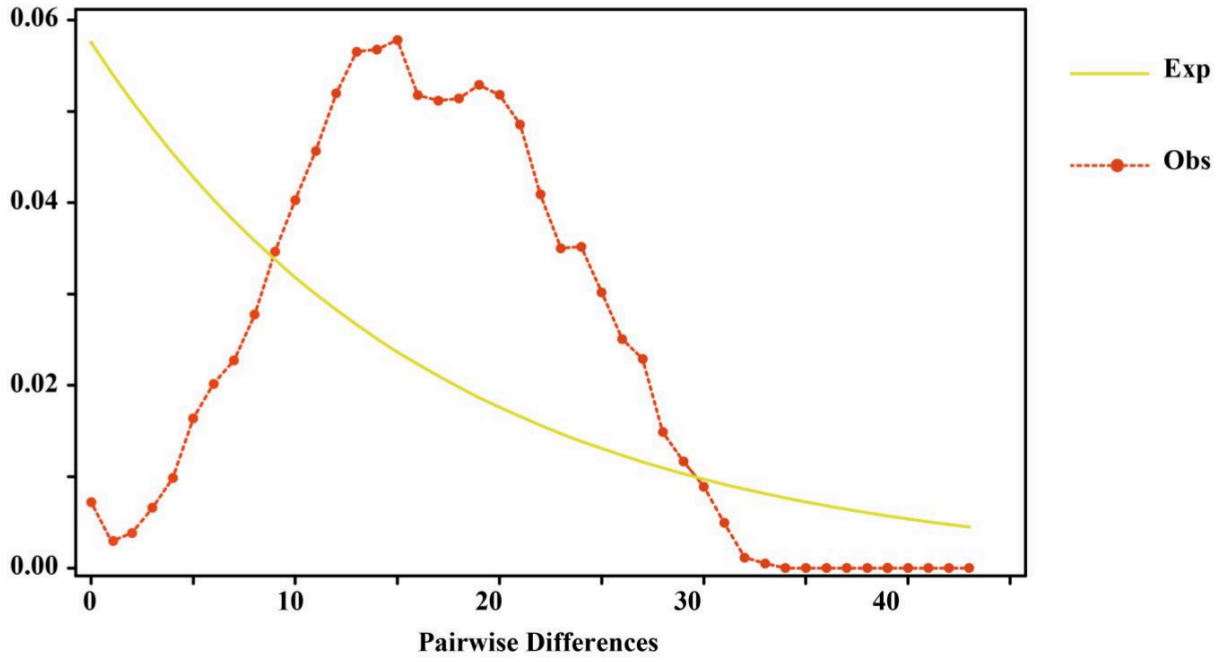


Figure 6 Mismatch distribution of *P. villosa* based on low-copy nuclear genes.

Figure 6

See image above for figure legend

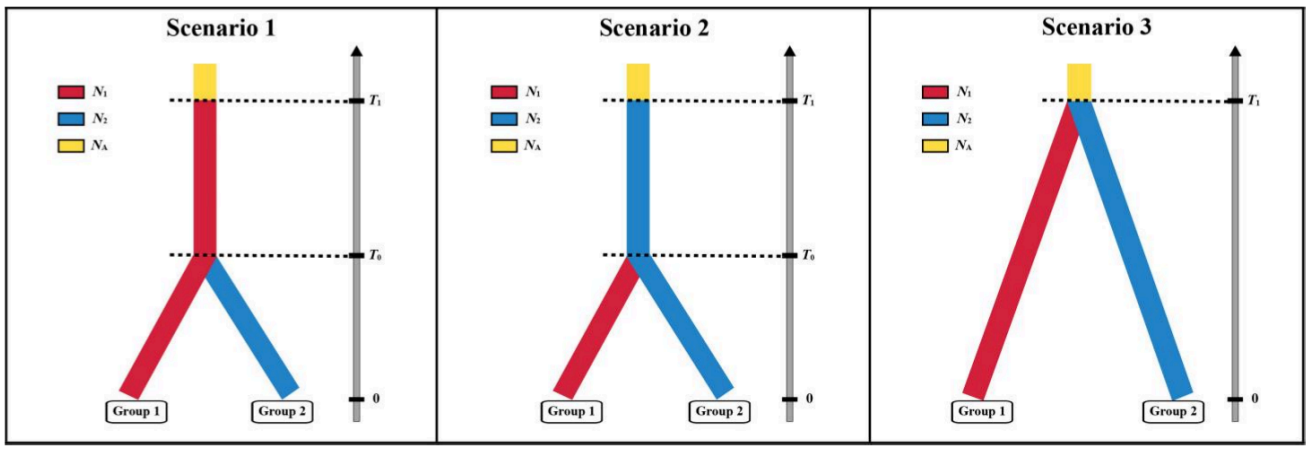


Figure 7 The three hypothetical evolutionary scenarios of *P. villosa*. N_1 and N_2, Long-term equilibrium effective population size of *P. villosa*; N_A, Effective population size of ancestral species; T_0 and T_1, Time of divergence between populations.

Figure 7

See image above for figure legend

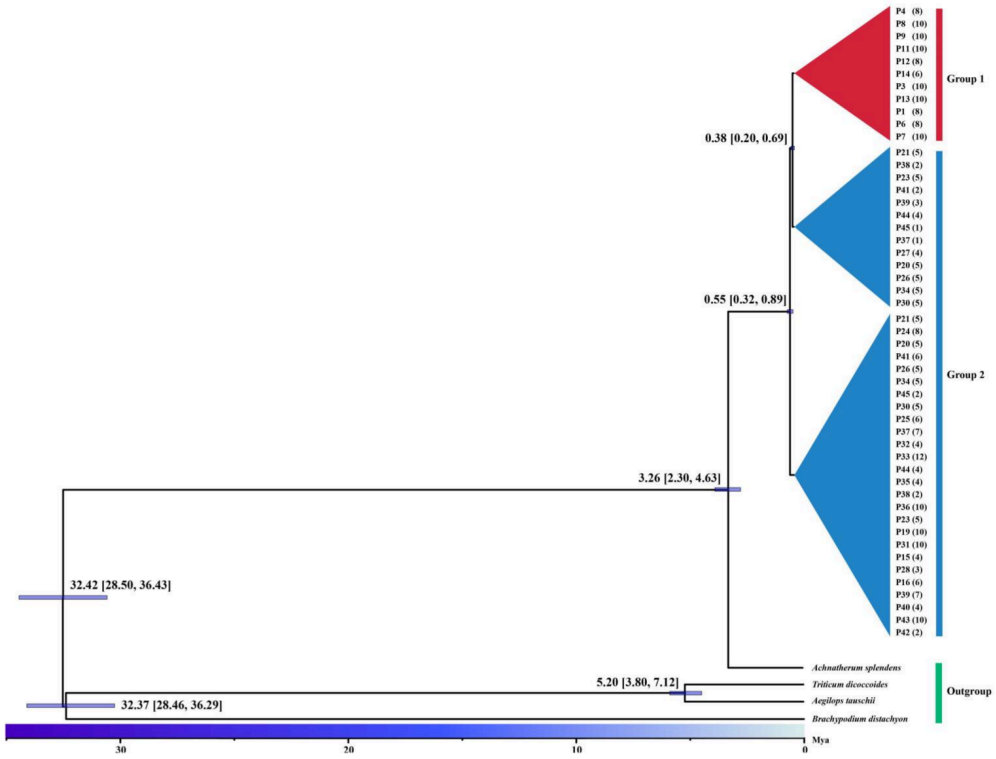


Figure 8 The divergence time of *P. villosa* estimated by BEAST. P stands for population; numbers in brackets represent the number of individuals; the numbers at each internal node of the evolutionary tree represent the divergence time and 95% confidence interval.

Figure 8

See image above for figure legend

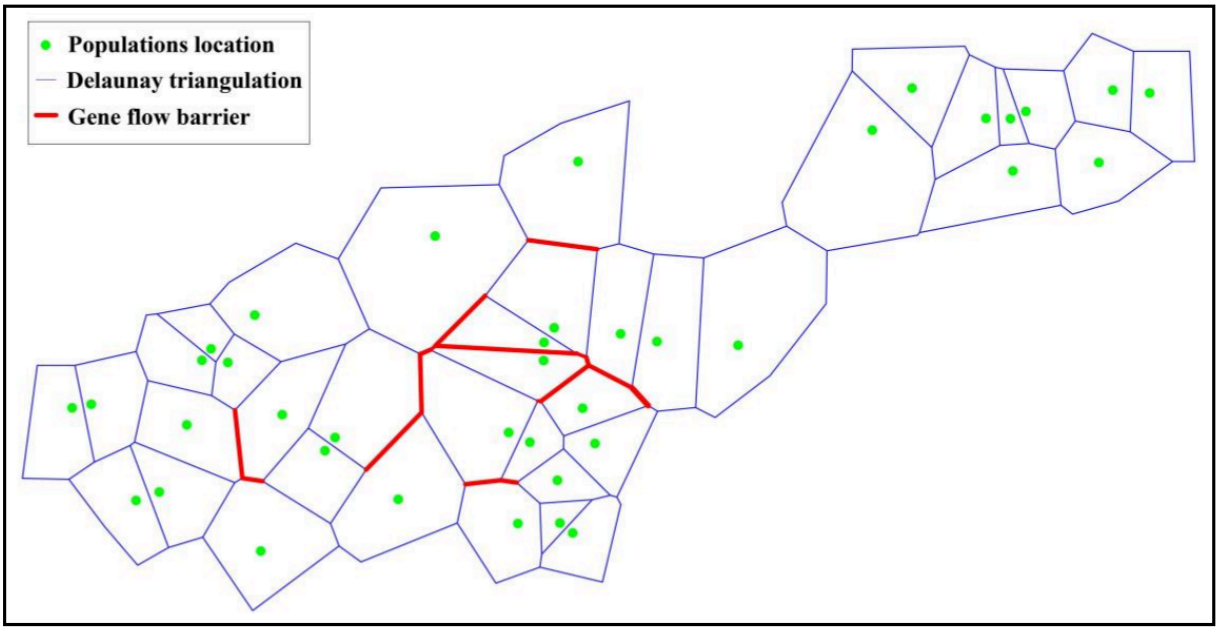


Figure 9 The analysis of isolation barriers of 39 populations of *P. villosa*.

Figure 9

See image above for figure legend

Supplementary Files

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

- PsaTableS1S5.doc